

**GIVING DISADVANTAGED ADOLESCENTS SKILLS TO FLOURISH:
RANDOM-CONTROL-TRIAL INTERVENTION INTEGRATING DEVELOPMENTAL
COACHING WITH OUTDOOR ADVENTURE EDUCATION**

Wendy Gelman Gwyn

AB (Int'l Rel), Stanford University

LLB, University of New South Wales

MAppSc (Psych of Coaching), The University of Sydney

Institute for Positive Psychology and Education

Faculty of Health Sciences

Australian Catholic University

Principal Supervisor: Professor Herbert W. Marsh

Co-Supervisors: Dr. Jiesi Guo and Dr. Theresa Dicke

Associate Supervisor (External): Dr. Michael Cavanagh, The University of Sydney

A thesis submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

January 2020

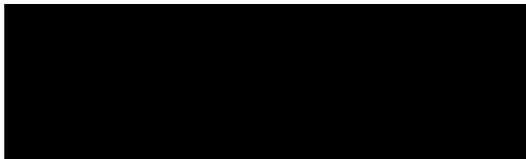
© Wendy G. Gwyn 2020

STATEMENT OF AUTHORSHIP AND SOURCES

This thesis contains no material that has been extracted in whole or in part from a thesis that I have submitted towards the award of any other degree or diploma in any other tertiary institution.

No other person's work has been used without due acknowledgment in the main text of the thesis.

All research procedures reported in the thesis received the approval of the relevant Ethics/Safety Committees (where required).



Wendy G. Gwyn

January 2020

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support and assistance of many people.

I am grateful to Professor Herbert Marsh for agreeing to work with me from a distance. Although my family's move to California was intended only to be for a year, he did not waiver in his support when a temporary move began to look more permanent. He helped me to advance my understanding of statistical methods and analysis and approached these areas with passion and enthusiasm that were contagious.

I am indebted to Dr. Michael Cavanagh who first introduced me to the work of Robert Kegan while I was studying coaching psychology at The University of Sydney. Acting as an external supervisor on this thesis, he was instrumental in my ability to undertake the qualitative research I proposed in addition to the quantitative study.

I would also like to thank my co-supervisor, Dr. Theresa Dicke, whose capacity for explaining complex concepts kept me sane. I am also grateful for her big-picture perspectives that often helped me to see the forest when I was stuck in the trees. I am thankful to Dr. Jiesi Guo who willingly stepped in to provide support while Dr. Dicke was on maternity leave. His expertise in R and Mplus was invaluable.

I am grateful also for the support of my colleagues and staff at the Institute for Positive Psychology and Education at the Australian Catholic University. I especially would like to thank Mathew Pfeiffer whose weekly Skype calls bridged the gap between California and Australia. I deeply appreciated his advice and support along the way.

A special mention goes to The Helmsman Project and its staff for implementing their novel program, for providing me with the opportunity to experience the program first-hand and research its effects, and for supporting the research process. Andrew Stainer was instrumental in establishing the program and encouraged me to join the research team. This research was also made possible by

the Australian Research Council, which provided grant funding through its Linkage Projects scheme.

The qualitative study would not have been achieved without the 13 students who shared their experiences with me and allowed me to push them to the edges of their meaning-making capacities. Their willingness to engage with me honestly and thoughtfully resulted in some of my most valuable learning in this entire process.

Sincere appreciation goes to my friend and neighbour, Laralyn Andrew Melvin, for agreeing to proofread this thesis in addition to her day job. Thank you also to my many friends, both in California and Australia, who cheered me on and provided a distraction when one was needed.

My parents sparked my love for the outdoors and adventure when they first sent me to summer camp at an early age. I am grateful to them also for instilling in me a passion for learning and a sense of determination, without which I could not have completed this work.

My children Jake, Ryan, and Alexa have been a constant source of support and inspiration for me. It has been gratifying to watch them develop a passion for the outdoors, and I am in awe of their own adventurous endeavours.

Finally, none of this would have been possible without my husband, Rhys. He is the one who pushes me to the edges of my meaning-making capacity. He also has been my fiercest supporter, believing that I could do this work even when I was most uncertain. I am grateful for his encouragement and advice, as well as the many cups of tea that kept me going.

TABLE OF CONTENTS

STATEMENT OF AUTHORSHIP AND SOURCES	III
ACKNOWLEDGEMENTS	V
LIST OF TABLES	XVIII
LIST OF FIGURES.....	XXI
ABBREVIATIONS AND SYMBOLS	XXIII
ABSTRACT.....	XXV
CHAPTER ONE	1
INTRODUCTION	1
CHAPTER TWO	7
OVERVIEW OF OUTDOOR ADVENTURE EDUCATION AND DEVELOPMENTAL COACHING: BACKGROUND, RELEVANT	
THEORETICAL CONCEPTS, EMPIRICAL SUPPORT, AND RESEARCH OPPORTUNITIES	7
<i>Introduction</i>	7
<i>Outdoor Adventure Education</i>	7
<i>Developmental Coaching</i>	8
<i>Central Theoretical Concepts</i>	12
Flourishing.....	12
Flourishing and outdoor adventure education.....	12
Flourishing and developmental coaching.	13
Constructivism and Experiential Learning Theory	13
Experiential learning theory and outdoor adventure education.	14
Experiential learning theory and developmental coaching.....	15
Goal Theory.....	16
Goal theory and outdoor adventure education.	18
Goal theory and developmental coaching.....	18
Challenge and Support	19
Challenge, support, and outdoor adventure education.....	20
Challenge, support, and developmental coaching.	21
Constructive-Developmental Theory	22
<i>Previous Research Findings and Existing Gaps in the Literature</i>	27
Outdoor Adventure Education.....	27
Previous research findings.....	27
Existing gaps in the literature.	30
Developmental Coaching	33
Previous research findings.....	33
Existing gaps in the literature.	36

<i>Integrating Developmental Coaching with Outdoor Adventure Education: Implications for the Present Investigation</i>	37
<i>Summary</i>	38
CHAPTER THREE	39
OVERVIEW OF QUANTITATIVE RESEARCH OUTCOMES: THEORETICAL CONSTRUCTS AND THEIR MEASUREMENT.....	39
<i>Introduction</i>	39
<i>Hope</i>	40
Theoretical Constructs.....	40
Snyder’s Hope Theory.....	40
Scheier and Carver’s Optimism.....	42
Self-Efficacy.....	43
Measuring Hope.....	45
Children’s Hope Scale (CHS).....	45
Life Orientation Test, Revised (LOT-R).....	45
<i>Self-Regulation</i>	45
Theoretical Construct.....	45
Measuring Self-Regulation: Adolescent Self-Regulatory Inventory.....	47
<i>Resilience</i>	48
Theoretical Constructs.....	48
Resilience.....	48
Grit.....	50
Measuring Resilience.....	51
Academic Resilience Scale (ARS).....	51
Life Resilience Scale (LRS).....	52
Short Grit Scale (Grit-S).....	52
<i>Motivation and Engagement</i>	52
Theoretical Construct.....	52
Measuring Motivation and Engagement: Motivation and Engagement Scale, Short (MES-S).....	53
<i>Wellbeing</i>	54
Theoretical Construct.....	54
Measuring Wellbeing.....	56
Satisfaction with Life Scale (SWLS).....	56
Warwick-Edinburgh Mental Well-Being Scale (WEMWBS).....	56
<i>Self-Concept</i>	57
Theoretical Construct.....	57
Measuring Self-Concept: Self-Description Questionnaire II-Short (SDQII-S).....	59
<i>Life Effectiveness</i>	60
Theoretical Construct.....	60
Measuring Life Effectiveness: Review of Personal Effectiveness with Locus of Control (ROPELOC).....	63
<i>Summary</i>	64

CHAPTER FOUR	65
GENERAL METHODOLOGY AND PROCEDURES	65
<i>Introduction</i>	65
<i>The Helmsman Project</i>	65
THP Program	66
Combined outdoor adventure education/coaching programs.	66
Coaching only programs.	67
Community project.	68
Coaches	68
Cohort Teachers	69
<i>Participants</i>	69
<i>Research Design</i>	70
Quantitative Design.....	71
Qualitative Design	71
<i>Survey Design and Format</i>	72
Demographic Information.....	72
Survey Items.....	72
Survey Format	75
<i>Research Procedures</i>	76
Ethics and Participant Consent	76
Survey Administration and Data Collection Procedures.....	77
Interview Procedures	78
<i>Quantitative Data Analysis</i>	78
Software	78
Treatment of Missing Data.....	78
Treatment of Response Bias and Negatively-Worded Items.....	80
Structural Equation Models (SEM).....	81
Calculating Factor Scores	82
Statistical Significance	82
Effect Size	83
<i>Qualitative Data Analysis</i>	84
<i>Summary</i>	84
CHAPTER FIVE	85
STUDY 1: EVALUATING THE PSYCHOMETRIC PROPERTIES OF THE MEASUREMENT INSTRUMENTS AND THEIR SCALES.....	85
<i>Introduction</i>	85
<i>Research Aims and Hypotheses</i>	88
Research Aims	88
Statement of Research Hypotheses.....	88
Research Hypothesis 1.1: Internal consistency reliability.	88

Research Hypothesis 1.2: Factor analysis	88
Research Hypothesis 1.3: Invariance analysis.	89
Research Hypothesis 1.4: Construct validity.	89
Research Hypothesis 1.4.1: Convergent validity	89
Research Hypothesis 1.4.2: Discriminant validity.	89
Summary	89
<i>Measurement Instruments</i>	89
Children’s Hope Scale (CHS).....	90
Life Orientation Test, Revised (LOT-R).....	92
Adolescent Self-Regulatory Inventory (ASRI)	94
Short Grit Scale (Grit-S).....	95
Motivation and Engagement Scale-Short (MES-S)	97
Academic Resilience Scale (ARS).....	99
Life Resilience Scale (LRS)	99
Satisfaction with Life Scale (SWLS)	100
Warwick-Edinburgh Mental Well-Being Scale (WEMWBS)	101
Self-Description Questionnaire II—Short (SDQII-S).....	104
Review of Personal Effectiveness with Locus of Control (ROPELOC)	106
Summary	108
<i>Methodology and Procedures</i>	109
Participants	109
Survey.....	109
Data Analysis	109
Treatment of missing data.....	109
Treatment of negatively-worded items.	110
Psychometric Analysis.....	110
Reliability.	110
Coefficient alpha.	110
Coefficient omega.	111
Threshold estimates.....	111
Factor analysis.	112
Confirmatory factor analysis.	113
Exploratory structural equation modelling.	114
Invariance testing.	116
Establishing goodness-of-fit.	117
Fit indices for factor analysis.....	118
Additional parameters of evaluation.	119
Fit indices for invariance testing.	119
Validity.	120
Summary	121
<i>Results</i>	122
Children’s Hope Scale (CHS).....	124

Results of Research Hypothesis 1.1: Internal consistency reliability of CHS.....	124
Results of Research Hypothesis 1.2: Factor analysis for CHS.....	126
Results of Research Hypothesis 1.3: Invariance analysis for CHS.....	129
Conclusion.....	130
Life Orientation Test-Revised (LOT-R).....	131
Results of Research Hypothesis 1.1: Internal consistency reliability of LOT-R.....	131
Results of Research Hypothesis 1.2: Factor analysis for LOT-R.....	132
Results of Research Hypothesis 1.3: Invariance analysis for LOT-R.....	134
Conclusion.....	135
Adolescent Self-Regulatory Inventory (ASRI).....	136
Results of Research Hypothesis 1.1: Internal consistency reliability of ASRI.....	136
Results of Research Hypothesis 1.2: Factor analysis for ASRI.....	136
Results of Research Hypothesis 1.3: Invariance analysis for ASRI-R.....	142
Conclusion.....	142
Short Grit Scale (Grit-S).....	143
Results of Research Hypothesis 1.1: Internal consistency reliability of Grit-S.....	143
Results of Research Hypothesis 1.2: Factor analysis for Grit-S.....	145
Results of Research Hypothesis 1.3: Invariance analysis for Grit-S.....	146
Conclusion.....	147
Motivation and Engagement Scale, Short (MES-S).....	148
Results of Research Hypothesis 1.1: Internal consistency reliability of MES-S.....	148
Results of Research Hypothesis 1.2: Factor analysis for MES-S.....	149
Results of Research Hypothesis 1.3: Invariance analysis for MES-S.....	153
Conclusion.....	154
Unidimensional Measures: ARS, LRS, SWLS, WEMWBS.....	155
Results of Research Hypothesis 1.1: Internal consistency reliability of unidimensional measures.....	155
Results of Research Hypothesis 1.2: Factor analysis for unidimensional measures.....	155
Results of Research Hypothesis 1.3: Invariance analysis for unidimensional measures.....	156
Conclusion.....	158
Self-Description Questionnaire II – Short (SDQII-S).....	158
Results of Research Hypothesis 1.1: Internal consistency reliability of SDQII-S.....	158
Results of Research Hypothesis 1.2: Factor analysis for SDQII-S.....	159
Results of Research Hypothesis 1.3: Invariance analysis for SDQII-S.....	161
Conclusion.....	163
Review of Personal Effectiveness with Locus of Control (ROPELOC).....	164
Results of Research Hypothesis 1.1: Internal consistency reliability of ROPELOC.....	164
Results of Research Hypothesis 1.2: Factor analysis for ROPELOC.....	165
Results of Research Hypothesis 1.3: Invariance analysis for ROPELOC.....	167
Conclusion.....	168
MTMM Analysis.....	169
Results of Research Hypothesis 1.4: Construct validity.....	169
Results of Research Hypothesis 1.4.1: Convergent validity.....	169

Results of Research Hypothesis 1.4.2: Discriminant validity	169
<i>Table 5.42</i>	176
Conclusion	177
Summary	177
<i>Discussion</i>	178
Strengths	178
Limitations	179
<i>Conclusion</i>	181
CHAPTER SIX	183
STUDY 2: QUANTITATIVE INVESTIGATION OF THP PROGRAM OUTCOMES	183
<i>Introduction</i>	183
<i>Research Aims, Hypotheses, and Questions</i>	185
Research Aims	185
Research Hypotheses and Questions: Statement and Rationale	186
Research Hypothesis 2.1: Short-term effects of THP programs on outcome variables	187
Research Hypothesis 2.1.1: Positive short-term effects of THP programs on all outcome variables	187
Research Hypothesis 2.1.2: Greater positive short-term effects of THP programs on outcome variables with most relevance to THP programs	187
Research Hypothesis 2.2: Long-term maintenance of positive effects of THP programs on outcome variables	188
Research Question 2.3: Follow-up effects of THP programs on outcome variables	188
Research Hypothesis 2.4: Aptitude-treatment interaction effects	189
Research Question 2.5: Differences in effects between Adventure Programs and Coaching Only Program	190
Research Question 2.6: Replication of RCT results with waitlist control group data	190
Summary	191
<i>Methodology and Procedures</i>	191
Participants	191
Measures	192
Research Design	193
Randomisation	194
Data Analysis	196
Treatment of missing data	196
Negatively-worded items	198
Scale relevance	198
Factor scores	200
Multiple regression and multiple-group analyses	200
Multiple regression analysis	200
Multiple-group analysis	201
Clustered data	201
Pre/Post intervention effects	202

Additional covariates.....	202
Data transformations.....	203
Interaction effects.....	203
Effect sizes.....	204
Summary.....	205
Results.....	205
Preliminary Descriptive Statistics and Analysis.....	206
Results of Research Hypotheses and Questions.....	209
Overview of Results.....	209
Results of Research Hypothesis 2.1: Short-term effects of THP programs on outcome variables.....	212
Results of Research Hypothesis 2.1.1: Positive short-term effects of THP programs on all outcome variables.....	215
Results of Research Hypothesis 2.1.2: Greater positive short-term effects of THP programs on outcome variables with most relevance to THP programs.....	216
Results of Research Hypothesis 2.2: Long-term maintenance of positive effects of THP programs on outcome variables.....	217
Results of Research Question 2.3: Follow-up effects of THP programs on outcome variables.....	221
Results of Research Hypothesis 2.4: Aptitude-treatment interaction effects.....	226
Example of aptitude-treatment interaction effect.....	231
Calculation of simple main effects.....	233
Aptitude-treatment interaction effects for the Short-Term Analysis (T1-T2).....	234
Aptitude-treatment interaction effects for the Long-Term Analysis (T1-T3).....	237
Conclusion.....	239
Results of Research Question 2.5: Differences in effects between Adventure Programs and Coaching Only Program.....	240
Results of Research Question 2.6: Replication of RCT results with waitlist control group data.....	243
Summary.....	251
Discussion.....	253
Strengths.....	253
Limitations.....	255
Conclusion.....	260
CHAPTER SEVEN.....	261
STUDY 3: QUALITATIVE EXAMINATION OF PARTICIPANT PROGRAM EXPERIENCES THROUGH THE LENS OF CONSTRUCTIVE-DEVELOPMENTAL THEORY AND INTERPRETATIVE PHENOMENOLOGICAL ANALYSIS.....	261
Introduction.....	261
Research Aims and Questions.....	265
Research Aims.....	265
Statement of Research Questions.....	265
Research Question 3.1: Assessment of constructive-developmental stage.....	266
Research Question 3.2: Constructive-developmental stage and program experience.....	266

Research Question 3.3: Constructive-developmental stage at THP program commencement and growth through program experience.	266
Research Question 3.4: Additional themes arising from participants' accounts of their program experiences.....	266
Summary.....	266
<i>Methodology and Procedures</i>	267
Participants	267
Research Design.....	269
Interpretative Phenomenological Analysis (IPA)	270
Constructive-Developmental Framework.....	271
The Instrumental knower.	272
The Socializing knower.	273
The Self-Authoring knower.....	274
The sub-phases of development between stages.....	275
Assessing constructive-developmental stage.	277
Interview Process.....	278
Ethics.....	281
Data Analysis.....	281
Recording and transcription.	281
Scoring constructive-developmental stage.	281
Coding.....	282
Potential Bias	282
Quotations from Interviews.....	283
Summary.....	283
<i>Results</i>	283
Introduction	283
Results of Research Questions.....	284
Results of Research Question 3.1: Assessment of constructive-developmental stage.	284
Constructive-Developmental stage at interview.	284
Instrumental-Socializing systems.....	285
Self-Authoring system.....	288
Conclusion.	289
Results of Research Question 3.2: Constructive-developmental stage and program experience.	289
Rules, boundaries, and structure.	289
Coaching sessions.....	291
Challenges and supports.	297
Community Project.....	300
Conclusion.	305
Results of Research Question 3.3: Constructive-developmental stage at THP program commencement and growth through program experience.....	306
Constructive-developmental stage at THP program commencement.	307
Constructive-developmental growth through program experience.	308

Instrumental to Socializing.....	308
Socializing to Self-Authoring.....	312
Conclusion.....	313
Results of Research Question 3.4: Additional themes arising from participants' accounts of their program experiences.....	314
Novelty, challenge and support.....	315
Program outcomes.....	319
Conclusion.....	323
Summary.....	324
<i>Discussion</i>	326
Strengths.....	326
Limitations.....	329
<i>Conclusion</i>	330
CHAPTER EIGHT.....	331
DISCUSSION AND CONCLUSION.....	331
<i>Introduction</i>	331
<i>Summary of Findings: Studies 1, 2, and 3</i>	332
Study 1.....	332
Study 2.....	334
Study 3.....	337
Summary.....	340
<i>Juxtaposing the Quantitative and Qualitative Findings</i>	341
Introduction.....	341
Overall Quantitative Program Outcomes.....	341
Adventure Programs.....	341
Coaching Only Program.....	343
Summary.....	345
Diverse Program Outcomes.....	345
Situational events.....	347
Program delivery.....	347
Timing of delivery.....	348
Coach training.....	348
Coach experience.....	348
Group dynamics.....	349
Initial grouping.....	350
Group experience.....	350
Post-program experiences.....	351
Post-program relationships.....	351
Post-program project.....	352
Other post-program opportunities.....	352
Individual developmental differences.....	353

Summary.....	353
Consistent Program Outcomes across Studies	354
Global positive self-beliefs.....	354
Quantitative outcomes.....	355
Qualitative outcomes.....	357
Social competence.....	359
Quantitative outcomes.....	359
Social Effectiveness.....	360
Cooperative Teamwork.....	361
Opposite-Sex Relationships Self-Concept.....	361
Qualitative outcomes.....	362
Perspective-taking capacity.....	364
Quantitative outcomes.....	364
Qualitative outcomes.....	365
Goal Self-Regulation.....	366
Summary.....	367
Primary Program Elements.....	367
Challenge and support.....	369
Experiential learning framework.....	369
Developmental coaching.....	370
Summary.....	371
Summary.....	371
<i>Strengths, Limitations, and Directions for Future Research.....</i>	<i>372</i>
Strengths and Directions for Future Research.....	372
Methodological contributions.....	372
Robust outcome measures.....	373
Rigorous research design and statistical methods.....	373
Mixed-methods approach.....	374
Theoretical contributions.....	374
Broad range of outcomes.....	375
OAE content versus processes.....	377
Adventure-based versus non-adventure-based programming.....	377
Outcomes with greater effects for those most in need.....	378
OAE with intact groups.....	378
Qualitative findings.....	379
Limitations and Directions for Future Research.....	380
Survey instrument.....	380
Quantitative outcomes.....	381
Research sample.....	381
RCT design.....	382
Qualitative design.....	384
<i>Implications for Educational Policy and OAE Research, Design, and Practice</i>	<i>385</i>

<i>Conclusion</i>	387
REFERENCES.....	389
APPENDICES.....	449
APPENDIX A Survey Items: Cross-References and Wording	449
APPENDIX B Comparison Overview of Adventure Programs	460
APPENDIX C Comparison of Adventure Programs and Coaching Only Program	462
APPENDIX D Recruitment Protocol	468
APPENDIX E Self-Reported Demographic Details of THP Program Participants	472
APPENDIX F Survey: Participant Consent and Demographic Questions (Qualtrics Online Version) ..	474
APPENDIX G THP Survey: Qualtrics Online Version	480
APPENDIX H Ethics Approval Report	500
APPENDIX I THP Application and Consent Form	501
APPENDIX J Research Interview Consent Form.....	511
APPENDIX K THP Program/Survey Dates and Community Project Details	512
APPENDIX L List of Statistical R Packages	515
APPENDIX M Coefficient Alpha for Each Scale	516
APPENDIX N Supplemental Research Hypothesis on Gender Invariance	517
APPENDIX O Factor Analytic Models: Chi-Square Test Statistics	522
APPENDIX P Longitudinal Invariance Models: Chi-Square Test Statistics	523
APPENDIX Q Factor Loadings and Correlations for SDQII-S ESEM/CFA	524
APPENDIX R Ropeloc Factor Loadings and Correlations for CFA	526
APPENDIX S Descriptive Statistics	527
APPENDIX T Baseline (T1) Group Differences	538
APPENDIX U Waitlist Control Group Extended Baseline Comparisons.....	539
APPENDIX V Intervention Group Within-Subjects Pre-Post Analysis.....	541
APPENDIX W Multivariate Omnibus Tests of High Relevance Scales	542
APPENDIX X Gender-Treatment Interactions: Significant Results	543
APPENDIX Y Subject-Object Interview Prompts.....	548
APPENDIX Z Participant Interview: Introduction to the Interview	551

LIST OF TABLES

<i>Table 2.1 Kegan’s Five Stages of Meaning-Making</i>	<i>25</i>
<i>Table 4.1 Summary of Survey Scales and Sample Items</i>	<i>74</i>
<i>Table 4.2 Percentage of Participants Missing All Data for a Wave: T1-T5</i>	<i>79</i>
<i>Table 4.3 Guidelines for Categorising Effect Sizes</i>	<i>83</i>
<i>Table 5.1 SDQII-S Sample Items</i>	<i>106</i>
<i>Table 5.2 ROPELOC Sample Items</i>	<i>108</i>
<i>Table 5.3 Overview of Results for Psychometric Analyses of Measurement Instruments and Scales</i>	<i>123</i>
<i>Table 5.4 Reliability Estimates (Omega Coefficients) for the CHS and its Scales using Long Form Data</i>	<i>125</i>
<i>Table 5.5 Inter-Item Correlations for the CHS Items using Long Form Data</i>	<i>126</i>
<i>Table 5.6 Fit Statistics for the Hypothesised CHS Factor Models using Long Form Data</i>	<i>127</i>
<i>Table 5.7 Standardised Factor Loadings for the CHS Two-Factor CFA and Two-Factor ESEM</i>	<i>128</i>
<i>Table 5.8 Longitudinal Invariance Models Based on CHS Two-Factor CFA-M at T1 and T2: Change in Fit Statistics</i>	<i>130</i>
<i>Table 5.9 Reliability Estimates (Omega Coefficients) for the LOT-R and its Scales using Long Form Data ...</i>	<i>132</i>
<i>Table 5.10 Inter-Item Correlations for the LOT-R Items using Long Form Data</i>	<i>132</i>
<i>Table 5.11 Fit Statistics for the Hypothesised LOT-R Factor Models using Long Form Data</i>	<i>133</i>
<i>Table 5.12 Standardised Factor Loadings for the LOT-R Two-Factor CFA and Two-Factor ESEM</i>	<i>134</i>
<i>Table 5.13 Longitudinal Invariance Models Based on LOT-R Two-Factor CFA at T1 and T2: Change in Fit Statistics</i>	<i>135</i>
<i>Table 5.14 Reliability Estimates (Omega Coefficients) for the ASRI and its Scales using Long Form Data ...</i>	<i>136</i>
<i>Table 5.15 Fit Statistics for the Hypothesised ASRI Factor Models using Long Form Data</i>	<i>137</i>
<i>Table 5.16 Standardised Factor Loadings for the ASRI Two-Factor CFA and Two-Factor ESEM</i>	<i>138</i>
<i>Table 5.17 Fit Statistics for the Hypothesised ASRI-R Factor Models using Long Form Data</i>	<i>140</i>
<i>Table 5.18 Standardised Factor Loadings for the ASRI-R Three-Factor CFA and Three-Factor ESEM</i>	<i>141</i>
<i>Table 5.19 Longitudinal Invariance Models Based on ASRI-R Three-Factor ESEM at T1 and T2: Change in Fit Statistics</i>	<i>142</i>
<i>Table 5.20 Reliability Estimates (Omega Coefficients) for the Grit-S and its Scales using Long Form Data .</i>	<i>144</i>
<i>Table 5.21 Inter-Item Correlations for the Grit-S Items using Long Form Data</i>	<i>145</i>
<i>Table 5.22 Fit Statistics for the Hypothesised Grit-S Factor Models using Long Form Data</i>	<i>145</i>
<i>Table 5.23 Standardised Factor Loadings for the Grit-S Two-Factor CFA and Two-Factor ESEM</i>	<i>146</i>
<i>Table 5.24 Longitudinal Invariance Models Based on Grit-S Two-Factor ESEM at T1 and T2: Change in Fit Statistics</i>	<i>147</i>
<i>Table 5.25 Reliability Estimates (Omega Coefficients) for the MES-S and its Scales using Long Form Data</i>	<i>148</i>
<i>Table 5.26 Inter-Item Correlations for the MES-S Items using Long Form Data</i>	<i>149</i>
<i>Table 5.27 Fit Statistics from EFA for the Hypothesised MES-S Factor Models using Long Form Data</i>	<i>150</i>
<i>Table 5.28 Geomin Rotated Loadings for MES-S EFA with Three Factors</i>	<i>151</i>

<i>Table 5.29 Standardised Factor Loadings and Correlations for the MES-S Three-Factor ESEM.....</i>	<i>152</i>
<i>Table 5.30 Longitudinal Invariance Models Based on MES-S Three-Factor ESEM at T1 and T2: Change in Fit Statistics</i>	<i>154</i>
<i>Table 5.31 Reliability Estimates (Omega Coefficients) for the Unidimensional Measures using Long Form Data.....</i>	<i>155</i>
<i>Table 5.32 Fit Statistics for the Unidimensional Measures using One-Factor CFAs and Long Form Data</i>	<i>156</i>
<i>Table 5.33 Longitudinal Invariance Models for Unidimensional Measures Based on One-Factor CFAs at T1 and T2: Change in Fit Statistics</i>	<i>157</i>
<i>Table 5.34 Reliability Estimates (Omega Coefficients) for the SDQII-S and its Scales using Long Form Data</i>	<i>159</i>
<i>Table 5.35 Fit Statistics for the Hypothesised SDQII-S Factor Models using Long Form Data.....</i>	<i>160</i>
<i>Table 5.36 Two Longitudinal Invariance Models Based on SDQII-S ESEM/CFA at T1 and T2 using (A) Wide Format Data and (B) Long Format Data with a Grouping Variable.....</i>	<i>162</i>
<i>Table 5.37 Reliability Estimates (Omega Coefficients) for the ROPELOC and its Scales using Long Form Data</i>	<i>165</i>
<i>Table 5.38 Fit Statistics for the Hypothesised ROPELOC Factor Models using Long Form Data</i>	<i>166</i>
<i>Table 5.39 Two Longitudinal Invariance Models Based on ROPELOC CFA at T1 and T2 using (A) Wide Format Data and (B) Long Format Data with a Grouping Variable</i>	<i>167</i>
<i>Table 5.40 Monotrait-Heteromethod Correlations (Highlighted on the Diagonal) and Heterotrait-Heteromethod Correlations (All Other Cells)</i>	<i>171</i>
<i>Table 5.41 Heterotrait-Monomethod Correlations at T1</i>	<i>173</i>
<i>Table 5.42 Heterotrait-Monomethod Correlations at T2</i>	<i>175</i>
<i>Table 6.1 THP Program Overview by Program Mode</i>	<i>192</i>
<i>Table 6.2 Breakdown of Participants by Group and Gender</i>	<i>195</i>
<i>Table 6.3 Percentage of Participants Missing All Data at each Wave: T1 to T4.....</i>	<i>197</i>
<i>Table 6.4 Differences at Baseline (T1) between Groups</i>	<i>208</i>
<i>Table 6.5 Overview of Results for Short-Term, Long-Term, and Follow-Up Analyses of THP Program Effects</i>	<i>210</i>
<i>Table 6.6 Short-Term Program Effects (T1-T2)</i>	<i>213</i>
<i>Table 6.7 Long-Term Program Effects (T1-T3)</i>	<i>218</i>
<i>Table 6.8 Follow-Up Program Effects (T2-T3)</i>	<i>223</i>
<i>Table 6.9 Aptitude-Treatment Interaction Effects: Parameter Coefficients for Significant Effects.....</i>	<i>227</i>
<i>Table 6.10 Summary Results of Effects of THP Programs including Aptitude-Treatment Interaction Effects</i>	<i>230</i>
<i>Table 6.11 Significant Aptitude-Treatment Interactions on T2 Cooperative Teamwork for Arctos Adventure Program Participants</i>	<i>232</i>
<i>Table 6.12 Aptitude-Treatment Interaction Effects for Short-Term Analysis (T1-T2).....</i>	<i>235</i>
<i>Table 6.13 Aptitude-Treatment Interaction Effects for Long-Term Analysis (T1-T3).....</i>	<i>238</i>

<i>Table 6.14 Comparison of Adventure Programs and Coaching Only Program Effects for all Waves</i>	<i>241</i>
<i>Table 6.15 Within-Subjects Effect Sizes at T4 (from Average Extended Baseline) for Waitlist Control Group as Program Participants</i>	<i>245</i>
<i>Table 7.1 THP Program Overview by Program Mode</i>	<i>268</i>
<i>Table 7.2 Interview Participant Details by THP Program and Gender (N=13)</i>	<i>268</i>
<i>Table 7.3 Constructive-Developmental Sub-phases between Kegan’s Instrumental and Socializing Ways of Knowing.....</i>	<i>276</i>
<i>Table 7.4 Interview Participants and their Constructive-Developmental Stage.....</i>	<i>285</i>
<i>Table 8.1 Short-Term Effect Sizes for Global Positive Self-Belief Outcomes</i>	<i>356</i>
<i>Table 8.2 Short-Term and Long-Term Effect Sizes for Social Competence Outcomes</i>	<i>360</i>
<i>Table 8.3 Short-Term Effect Sizes for Open Thinking Outcomes.....</i>	<i>364</i>

LIST OF FIGURES

<i>Figure 2.1. Developmental trajectory. The figure illustrates the differing courses of assimilative and accommodative change. Adapted from Cavanagh, 2016.</i>	<i>11</i>
<i>Figure 2.2. David A. Kolb’s (1984) experiential learning cycle. This figure outlines the learning process central to David A. Kolb’s experiential learning theory. Adapted from https://www.simplypsychology.org/learning-kolb.html.</i>	<i>14</i>
<i>Figure 5.1. Revised ASRI three-factor structure (ASRI-R).</i>	<i>139</i>
<i>Figure 5.2. Modified MES-S three-factor structure.</i>	<i>153</i>
<i>Figure 6.1. Research design for Study 2.</i>	<i>194</i>
<i>Figure 6.2. THP program effect sizes with 95% confidence intervals for Short-Term Analysis (T1-T2) with scales grouped by relevance.</i>	<i>214</i>
<i>Figure 6.3. THP program effect sizes with 95% confidence intervals for Long-Term Analysis (T1-T3) with scales grouped by relevance.</i>	<i>219</i>
<i>Figure 6.4. THP program effect sizes with 95% confidence intervals for Follow-Up Analysis (T2-T3) with scales grouped by relevance.</i>	<i>224</i>
<i>Figure 6.5. Factor scores for T2 Cooperative Teamwork at different T1 Cooperative Teamwork baseline levels for the waitlist control and Arctos Adventure Program groups.</i>	<i>233</i>
<i>Figure 6.6. THP program effect sizes with 95% confidence intervals for control group (as program participants) within-subjects pre-post analysis (Average T1, T2, T3 to T4) with scales grouped by relevance.</i>	<i>247</i>
<i>Figure 7.1. Histogram reflecting distribution of interview participants’ time (years) from completion of THP program to interview date (N = 13).</i>	<i>269</i>
<i>Figure 8.1. Diagram reflecting the relationship between primary program elements, participant meaning-making capacity, and program outcomes.</i>	<i>368</i>

ABBREVIATIONS AND SYMBOLS

Abbreviation/Symbol	Description
α	Cronbach's alpha or coefficient alpha
ω	coefficient omega
χ^2	chi square
Δ	delta (change)
Adv or Adventure Programs	each of the THP programs which includes both coaching and an adventure component: Arctos, James Craig, and Outward Bound
Arctos or Actos Adventure Program	the small-sized group, sailing-based Adventure Program on a yacht by the same name
ARS	Academic Resilience Scale
ASRI	Adolescent Self-Regulatory Inventory
CFA	confirmatory factor analysis
CFI	comparative fit index
CHS	Children's Hope Scale
CI	confidence interval
Coaching Only Program	the THP program which includes coaching and provides skills training in lieu of an adventure component
Cohen's d or d	an effect size used to indicate the standardised difference between two means
Cohort Teacher	a teacher for a school assigned by the school to support a group of its students participating in a THP program
Community Project	optional project undertaken by the program participants to make a positive difference in their community
df	degrees of freedom
ELT	experiential learning theory
ES	effect size
ESEM	exploratory structural equation model or modelling
FIML	full information maximum likelihood
Grit-S	Short Grit Scale
HTHM	heterotrait-heteromethod
HTMM	heterotrait-monomethod
JC, James Craig, or James Craig Adventure Program	the large-sized group, sailing-based Adventure Program on the tall ship by the same name

Abbreviation/Symbol	Description
LOT-R	Life Orientation Test-Revised
LRS	Life Resilience Scale
MCAR MAR MNAR	missing completely at random, missing at random, missing not at random
MES-HS and MES-S	Motivation and Engagement Scale-High School and Motivation and Engagement Scale-Short
ML MLR	maximum likelihood estimation and maximum likelihood estimation with robust standard errors
MTHM	monotrait-heteromethod
MTMM	multitrait-multimethod
OAE	outdoor adventure education
OECD	Organisation for Economic Cooperation and Development
OB, Outward Bound, or Outward Bound Adventure Program	Outward Bound Australia and the medium-sized group, land-based Adventure Program run by Outward Bound Australia
r	the Pearson correlation coefficient
RCT	randomised controlled trial
RMSEA	root mean square error of approximation
ROPELOC	Review of Personal Effectiveness with Locus of Control
SD, SE	standard deviation, standard error
SDQII-S	Self-Description Questionnaire II-Short
SEM	structural equation model or modelling
SOI	subject-object interview
Survey	the Likert-style self-report survey in paper and online format completed by the participants for this research
SWLS and SWL	Satisfaction with Life Scale and Satisfaction with Life
T1, T2, T3, T4, T5	assessment wave 1, assessment wave 2, etc.
THP	The Helmsman Project
THP program(s)	the novel program provided by THP or together all of the modalities through which the program is offered
TLI	Tucker-Lewis index
WEMWBS and SWEMWBS	Warwick-Edinburgh Mental Well-Being Scale and its short version

ABSTRACT

The Helmsman Project brings a novel extracurricular program to high schools located in areas of socioeconomic disadvantage. Integrating a series of structured developmental coaching sessions with outdoor adventure experiences, this program aims to positively influence participants' personal and social development through a range of outcomes, with a particular focus on building hope, resilience, and self-regulation.

Outdoor adventure education (OAE) research has found the quality of program facilitation to have a strong connection with program outcomes. As coaching is a form of facilitated development, it was hypothesised that incorporating coaching into OAE would contribute to the experiential learning process and enhance program impact. Current research supports the effectiveness of both OAE and coaching for the development of important life skills. However, research in both fields is limited and what exists is undermined by criticisms of a lack of methodological rigour.

This thesis investigates the effectiveness of The Helmsman Project program using a randomised controlled trial with 362 high school students from 11 schools located in Western Sydney, an area of socioeconomic disadvantage. The research outcomes include 41 scales from 11 measurement instruments, covering the broad constructs of hope, self-regulation, resilience, motivation and engagement, wellbeing, multiple facets of self-concept, and various life effectiveness skills.

Study 1 evaluates the psychometric properties of the outcome measures by considering internal consistency reliability, factor structure and its invariance, and construct validity. In Study 2, data from the waitlist control group is used as a basis of comparison with the intervention group data from one pre-test and two post-tests, applying multiple regression analysis. The waitlist control group data also serves as an extended baseline against which to compare later results for the control group when they subsequently experienced a program.

The adventure component of The Helmsman Project program is provided through three different modalities, two water-based and one land-based. The distinct effects of these programs are considered in addition to the overall effects of

these programs as a whole. Moreover, a separate coaching program (without the adventure experiences) was offered, providing the opportunity to test the incremental benefits of the adventure component for The Helmsman Project program.

This thesis also includes a qualitative study. Study 3 examines how 13 program participants made meaning of their experiences in The Helmsman Project program, using semi-structured interviews and the lens of constructive-developmental theory. A mixed-methods approach provides the most complete picture of the program effects and helps to further the evidence base for the benefits that OAE and coaching programs have to offer disadvantaged adolescents.

Results from this research demonstrate some significant positive effects, particularly in developing hope, positive global self-beliefs, wellbeing, and other life effectiveness skills, including social effectiveness and open thinking. While follow-up analysis indicated general stability in the outcomes during the three-month period following program completion, some significant short-term positive effects were not retained over the long term. However, several new significant positive effects were evidenced, with some significant effects found in the follow-up period. In general, those scales of most relevance to the design and aims of The Helmsman Project program demonstrated the greatest effects. Moreover, aptitude-treatment interaction analysis suggests that for some outcomes these programs may provide greater benefits to those most in need, reinforcing the value in providing school-based OAE programs to disadvantaged students. However, the varied effects across programs also underline the complexity inherent in delivering these types of programs, and the difficulty of isolating the various program elements and their differential impact on outcomes.

The qualitative results provide support for constructive-developmental capacity as a relevant individual difference influencing OAE program experiences and outcomes. This finding has implications for the design and implementation of OAE. By understanding these developmental differences, OAE providers can match processes and expectations more closely to developmental capacities and provide a better holding environment for learning and developmental growth. There is also an indication that for some participants the positive program outcomes occurred

only with further experiences that prompted reflection on the program learnings and provided further opportunity to apply those learnings. Consequently, some program effects may take time.

The findings in this thesis evidence the value in a mixed-methods approach to OAE research. Taken together, the quantitative and qualitative studies provide a more complete and holistic understanding of the relationship among OAE program elements, participants, and outcomes. Following a juxtaposition of Study 2 and Study 3, the strengths and limitations of this research are highlighted, as well as directions for future research and implications for educational policy and OAE research, design, and practice. It is hoped that the methodologically rigorous results from this multidimensional investigation contribute to literature, research, and practice in the fields of OAE and coaching psychology, and provide a platform for future study.

CHAPTER ONE

INTRODUCTION

“High school experiences provide an essential platform for academic and non-academic accomplishments, psychological development, further education, and future life.” (Marsh & Kleitman, 2002, p. 468)

In 2016, the Organisation for Economic Co-operation and Development (OECD; 2016) published a report on education that recognised that what makes people flourish in their professional and private lives are the qualities of character and skills developed through school. However, research confirms that throughout the OECD member countries, students of lower socioeconomic status continue to experience inequity in educational outcomes when compared with their more advantaged peers (Castejón & Zancajo, 2015; Parker, Jerrim, Schoon, & Marsh, 2016). This educational disparity leads not only to inequalities in employment opportunities and earnings, but also to inequalities in health and overall wellbeing (Cutler & Lleras-Muney, 2010; Organisation for Economic Cooperation and Development, 2016; Richter, 2006).

Such broad-ranging inequities result in enormous social and financial costs for our society. Researchers in the United States estimate the costs associated with child poverty in that country at \$500 billion per year, which is 4 percent of the US Gross Domestic Product (Coley & Baker, 2013). Governments in many countries are focused on reducing such imbalance. While the economic and social disadvantages that generate these inequalities need to be addressed directly through social and economic policies, issues as complex as this subject cannot be met by one focus or approach alone (see, for example, Duncan & Murnane, 2011). Consequently, research into effective strategies for improving outcomes for disadvantaged students is essential.

While the imbalance in educational outcomes is impacted by a range of influences that extend beyond school, it is now well-established that factors such as

academic self-concept, school engagement, aspirations, goal setting, and identity development play a part in these unequal outcomes (Dietrich, Parker, & Salmela-Aro, 2012; Heaven & Ciarrochi, 2008; Parker et al., 2016; Parker et al., 2012). Even when academic ability is equal, able but disadvantaged youth are still more likely to have less ambitious educational aspirations and be more disengaged from school (Parker et al., 2016). Accordingly, skills necessary for successful outcomes are not limited to cognitive skills; non-cognitive skills such as motivation, persistence, resilience, self-regulation, and goal strategies are also essential (Heckman, 2006). Beyond the impact non-cognitive skills may have on one's cognitive abilities, some scholars suggest that these skills are a core determinant of a person's developmental and life potential (Garcia, 2015; Gutman & Schoon, 2013).

Participation in extracurricular activities has been associated with the development of non-cognitive skills, as well as other positive outcomes, including academic achievement, school and civic engagement, and educational attainment (Covay & Carbonaro, 2010; Eccles, Barber, Stone, & Hunt, 2003; Marsh, 1992a; Marsh & Kleitman, 2002; Richmond, Sibthorp, Gookin, Annarella, & Ferri, 2018; Zaff, Moore, Papillo, & Williams, 2003). In addition, the benefits of extracurricular activities that are school-based appear to be greater for disadvantaged than advantaged students, thereby reducing inequality (Marsh, 1992a; Marsh & Kleitman, 2002).

The Helmsman Project (THP) brings a novel extracurricular program to high schools located in areas of socioeconomic disadvantage. Integrating a series of structured developmental coaching sessions with outdoor adventure experiences, this program aims to positively influence participants' personal and social development through a range of outcomes, with a particular focus on building hope, resilience, and self-regulation. A further program objective is for participants to develop broader perspective-taking capacities, enabling them to perceive different and bigger possibilities than before. Through these outcomes, the ultimate goals of the programs are to improve participants' educational engagement and wellbeing, thereby assisting participants to flourish and reach their full potential.

Both outdoor adventure activities and coaching have been applied independently in school settings with a focus on personal growth and the

development of important life skills (e.g., Green, Grant, & Rynsaardt, 2007; Green, Oades, & Grant, 2006; Hattie, Marsh, Neill, & Richards, 1997; Richmond et al., 2018). While the research findings on the separate application of outdoor adventure education and coaching in school and other settings have been largely positive, more robust research design and analysis is required in both areas, as well as a better understanding of the relationship between program variables, participants, and outcomes (see Grant, 2012a, 2016b; Grant & Cavanagh, 2007; Grant & Cavanagh, 2011; Hans, 2000; Neill, 2008; Richmond et al., 2018; Sheard & Golby, 2006; Sproule et al., 2013). To date, no research has been found which has tested the effectiveness of the integration of both.

The design and delivery of the THP program is based on existing evidence from research, theory, and practice in outdoor adventure education, coaching psychology, and psychology more broadly. In outdoor adventure activities, one program variable that research has found to have a strong connection with program outcomes is the quality of the program facilitation (e.g., Hattie et al., 1997; McKenzie, 2000, 2003; Paisley, Furman, Sibthorp, & Gookin, 2008). As coaching is a form of facilitated growth and development, it is hypothesised that incorporating coaching into the programs will contribute to the experiential learning process and enhance program impact.

This research project aims to investigate the effectiveness of the THP program using a methodologically robust design and statistically rigorous analysis. By testing the effectiveness of THP's novel program, we can assess a new approach while also advancing the research base in outdoor adventure education and coaching psychology, with the potential to improve outcomes in both fields. The THP program provides the adventure experiences through three alternative modes, two of which are water-based and one land-based. This research also provides an opportunity to consider the impact of these different types of adventure experiences. Moreover, a separate program offering only the coaching element of the program (replacing the adventure experiences with skills-based workshops) has been implemented in some of the groups for the purpose of examining the incremental benefits of the outdoor adventure experience for the THP program. This is important given the additional costs and complexities involved in providing

the adventure experiences. Finally, the present investigation incorporates both quantitative and qualitative studies. It is anticipated that the application of a mixed-methods approach will provide the most complete picture of the THP program effects and help to further the evidence base for the benefits that outdoor adventure education and coaching programs have to offer disadvantaged adolescents.

This thesis comprises eight chapters. Following this introductory chapter, Chapter Two and Chapter Three review the literature relevant to this thesis. Chapter Two provides an overview of the theories central to the THP program and how those theories are relevant to outdoor adventure education and coaching. Some of the existing empirical literature on both outdoor adventure education and coaching psychology is reviewed, as well as gaps in the existing literature. A case is then made for the incorporation of coaching into outdoor adventure education. Given that this research incorporates 41 scales from 11 different measurement instruments, which purport to measure a range of constructs, Chapter Three provides an overview of the outcomes for this research, including a detailed review of the instruments used to measure those outcomes and the theoretical constructs underlying them.

Following the literature review, Chapter Four gives a detailed account of the overarching methodology and procedures for the research. More specific elements of the methodology and procedures for each study are included in the chapter devoted to that study. Each of the next three chapters provides a self-contained report for one of the three studies in this thesis: Chapter Five covers the psychometric analysis of the 41 measurement scales used in the research and the data derived therefrom; Chapter Six details the quantitative investigation of the THP program outcomes using a randomised controlled design and factor scores derived from the psychometric analysis; and Chapter Seven documents the qualitative exploration of a subset of THP program participants using a semi-structured interview technique to assess the way in which those participants made meaning of their experiences in the program. Each of these chapters is intended to stand alone with its own introduction, research aims, hypotheses, and questions, methodology and procedures, results, discussion, and summary. Consequently,

there is some overlap across these chapters. Chapter Eight provides an opportunity to draw together the findings from these studies and offer some concluding thoughts in relation to the research aims, as well as any implications of the present research for the future.

CHAPTER TWO

OVERVIEW OF OUTDOOR ADVENTURE EDUCATION AND DEVELOPMENTAL COACHING: BACKGROUND, RELEVANT THEORETICAL CONCEPTS, EMPIRICAL SUPPORT, AND RESEARCH OPPORTUNITIES

Introduction

Both the coaching and outdoor adventure components of the THP program seek to integrate evidence-based practice from strong theoretical backgrounds in the fields of educational, developmental, and positive psychology. This chapter begins with a brief background on outdoor adventure education and coaching, and then reflects on theories central to the THP program and how those theories have been applied in practice and research relevant to outdoor adventure education and coaching. It concludes by outlining some of the previous research findings in these areas, as well as some of the gaps in the existing literature, and how some of those gaps will be addressed by the current research. A detailed analysis of the outcomes measured in this research and the theoretical constructs and measurement instruments underlying those outcomes is provided in Chapter Three.

Outdoor Adventure Education

In 1941, Kurt Hahn developed the first Outward Bound program (Hahn, 1957). This month-long course, designed to develop initiative, fitness, self-reliance, and resourcefulness, is said to be the origin of current outdoor adventure education (Hattie et al., 1997; Neill, 2008). Some of the essential features of Outward Bound and other outdoor adventure programs today include immersion in an unfamiliar environment, challenging activities, goals, a supportive environment, and opportunities for discovery and reflection (Neill, 2008; Schafermeyer, 1978; Sibthorp, Furman, Paisley, Gookin, & Schumann, 2011).

At the heart of outdoor adventure education (OAE) is the context of a natural environment. There are well-established links between being in nature and

improvements in personal health and wellbeing (Carpenter & Harper, 2016; Dillon, 2012). Louv (2005) used the term “nature-deficit disorder” (p. 34) to refer to our modern disconnection from nature and the negative impact this has on our physical, psychological, and social health and wellbeing (see also Maller et al., 2008). It is suggested, however, that these negative effects can be rectified through re-exposure to natural environments. The beneficial outcomes found from exposure to nature are many, including improvements in mood, anxiety, attention, physical health and health-oriented behaviour, self-esteem, self-concept, and life satisfaction (Neill, 2008).

While today’s OAE programs come in many forms and serve a variety of purposes, the term *outdoor adventure education* (or OAE) in this thesis, refers more specifically to organised programs for small groups of adolescents that apply an experiential learning model. Goal setting and goal striving also sit at the core of these experiential learning programs (Neill, 2008). By including *adventure*, it is also intended that these programs involve a component of adventure that occurs in the outdoor environment and is challenging for the participants. Outdoor education, more generally, can have a range of aims, including physical, personal, social, educational, therapeutic, and environmental (Dillon, 2012; Neill, 2008). However, the outdoor adventure education on which this thesis is focused, has as its primary aim, the personal and social development of its participants through a wide variety of outcomes, including hope, resilience, and self-regulation (for more detail, see the section below headed “Flourishing and outdoor adventure education”).

Developmental Coaching

The coaching undertaken within the THP program fits within the field of psychological coaching, rather than sports or educational coaching. It draws on, and contributes to, established psychological theories, principles, and approaches and the evidence underpinning these approaches (Grant & Cavanagh, 2011; Green, Oades, & Robinson, 2012; Madden, Green, & Grant, 2011; Palmer & Whybrow, 2008; Whybrow, 2008). While psychological coaching is a relatively new field, the general practice of coaching has been around for much longer, with sports coaching going back to ancient Greece (Allen, 2016). The formal study of psychology as it applies to

coaching is said to date from the 1920s, when Coleman Griffith conducted research into athletics and coaching (Whybrow, 2008). The term *coaching psychology* first appeared in Elisha Curtiss Gaylord's 1967 book of the same name, on the topic of psychology as it related to sports coaching (Whybrow, 2008).

At the same time as psychology was finding its way into sports coaching, business coaching began to take hold through the use of psychologists and organisational development teams tasked with increasing productivity (Allen, 2016). Alongside this growth in sports and business coaching, the humanist movement of the 1960s gave impetus to a broader form of coaching focused on health, wellbeing, and life in general (Palmer & Whybrow, 2008). When applied in an educational environment, this form of coaching is distinguished from tutoring or other support in that setting specifically aimed at improving academic performance (Green, Grant, et al., 2007).

Coaching can be understood generally as a collaborative and solution-focused process in which the coach facilitates the enhancement of a person's goal attainment and wellbeing through improvements in self-directed learning and personal growth (Green, Grant, et al., 2007). However, coaching is not a singular activity. A finer understanding of the type of coaching undertaken in the THP program can be gained by focusing on the primary aims of the coaching interventions. One taxonomy of coaching breaks coaching down into four main aims or types: skills coaching, performance coaching, developmental coaching, and remedial coaching (Standards Australia, 2011). Skills coaching is about building capability through skills acquisition, while performance coaching focuses on using existing skills more effectively. Remedial coaching is aimed at modifying problematic attitudes and behaviours impeding one's life goals. Developmental coaching is aimed at helping people develop a more complex understanding of themselves, others, and the systems in which they live, in order to improve their ability to meet life's challenges more effectively (Standards Australia, 2011). Typically, any coaching intervention combines elements of all four coaching types, but is weighted toward one type or another depending on its aim (Grant, Passmore, Cavanagh, & Parker, 2010; Standards Australia, 2011).

While the THP program focuses on a mix of skills acquisition, performance enhancement, and correction of maladaptive responses, it is more than this. It involves developing a new relationship to oneself and the challenges one faces. Consequently, the coaching used in the THP program is primarily developmental coaching, and, as such, it is aimed at helping participants see themselves, others, and the systems in which they live in more nuanced and complex ways, so as to open new possibilities for responding to the challenges life presents them.

Piaget (1976; 1983) distinguished between two types of development or change: *assimilation* and *accommodation*. Assimilation (also referred to as horizontal growth; Cook-Greuter, 2004) involves learning new skills or acquiring new knowledge that fits within one's current cognitive structure. On the other hand, accommodation (also referred to as vertical transformation; Cook-Greuter, 2004) occurs when one's current ways of seeing the world fail to interpret an experience satisfactorily, resulting in a need to adjust one's cognitive structure. When this occurs, a person doesn't just know more, they know differently (Kegan, 1994). Such transformative change is a process that typically occurs in marked transitions preceded by a period of confusion and instability that can be uncomfortable (Cavanagh, 2016). This process is quite different to assimilative development, where the changes often occur quickly and are noticeable, such as with the acquisition of a new skill (Cavanagh, 2016). The course of these two types of change is depicted below in Figure 2.1, which reflects an increasing upward trend for assimilative change marked by a sharp decline followed by a sharp incline for accommodative change.

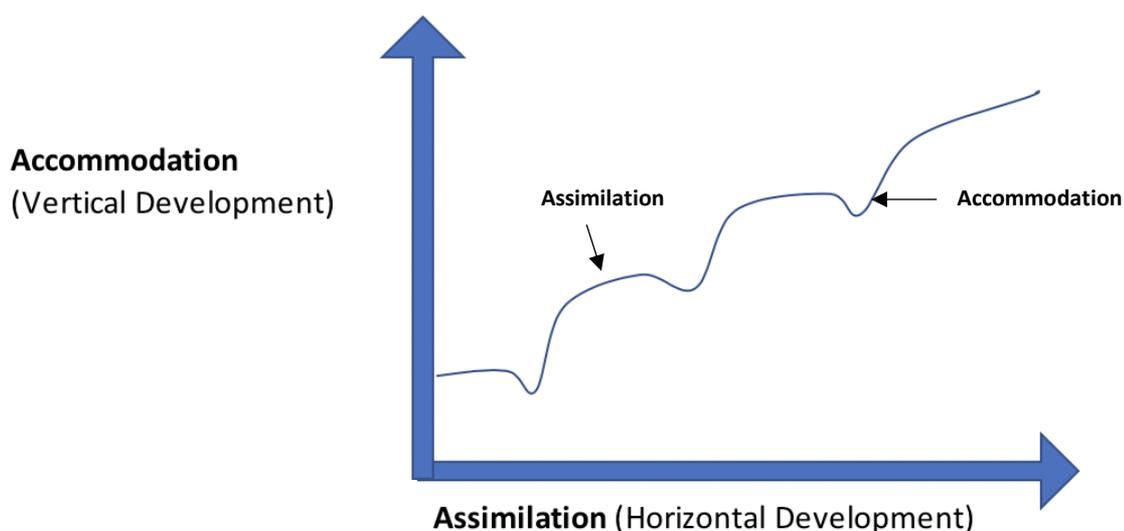


Figure 2.1. Developmental trajectory. The figure illustrates the differing courses of assimilative and accommodative change. Adapted from Cavanagh, 2016.

The THP program seeks to stimulate not only horizontal growth but vertical growth as well. Central to building hope, resilience, and self-regulation, and the skills related to those outcomes, is the capacity to take broader perspectives, which often requires a developmental shift. Therefore, while both types of change are relevant to the THP program, accommodative change, and therefore developmental coaching, are key to the program's outcomes. Accordingly, in this thesis the evidence-based, psychological coaching that forms part of the THP program is referred to as *developmental coaching*.

The THP program aims to foster developmental growth in its participants so that they are better equipped to meet the challenges they face and more engaged in their world, ultimately allowing them to flourish and reach their full potential. The next section describes in more detail what is meant by *flourishing*, an umbrella term used in the field of positive psychology which encompasses many of the outcomes relevant to this research (more detail on these outcomes and their theoretical underpinnings can be found in Chapter Three). The following sections then cover a number of other theoretical concepts that are central to the philosophies and processes underlying OAE and developmental coaching: constructivism and experiential learning theory, goal theory, and the interaction between challenge and support. These theories ground the processes underlying the THP program.

However, it is suggested that how the program participants make sense of those processes and their capacities for sense-making more generally also have an impact. Therefore, the final section of the theoretical review focuses on constructive-developmental theory, a theory of the development in the way one constructs meaning of his or her experiences.

Central Theoretical Concepts

Flourishing

The field of positive psychology focuses on understanding what makes life worth living and how to nurture it (Seligman & Csikszentmihalyi, 2000). Positive psychologists approach this task from the premise that what makes life most worth living is not merely the absence of distress and disorder, or even its opposite; rather, what makes life most worth living is something distinct in its own right (Ciarrochi, Kashdan, & Harris, 2013; Keyes, 2003). This distinct something has been referred to as *flourishing* (Huppert & So, 2013; Keyes, 2002; Seligman, 2011). Corey Keyes is credited with being the first in the field of psychology to use the term flourishing (Hone, Jarden, Schofield, & Duncan, 2014). Keyes (2002) developed a model of mental health that proceeds on a continuum from languishing to flourishing. Flourishing individuals feel good about life and are functioning well psychologically and socially (Keyes & Annas, 2009). There is considerable empirical support for wide-ranging benefits to individuals and our communities associated with flourishing (Howell, 2009; Huppert, 2004; Huppert, 2009; Keyes, 2002; Keyes & Annas, 2009; Vella-Brodrick, Park, & Peterson, 2009). This notion of flourishing is a broad aim of the THP program, to be derived through a range of positive outcomes focused on personal and social development.

Flourishing and outdoor adventure education. Much of the literature on OAE has as a primary outcome, the personal and social development of its participants (Sibthorp, 2000; Stott, Allison, Felter, & Beames, 2015). This concept is also commonly considered an important outcome in the field of education (sometimes referred to as non-cognitive skills or social and emotional learning; e.g., Gutman & Schoon, 2013). Barret and Greenaway (1995) include within this term the enhancement of aspects of self-concept, increases in the internalisation of locus of

control, development of interpersonal relationships and cooperative teamwork, and improvements in character strengths (e.g., humour, patience, vitality, optimism, self-confidence, self-regulation, spirituality). Buck and Inman (1998) note that the concept of personal and social development refers both to the processes of development and the outcomes of that development. As an outcome, they suggest that personal and social development can be understood as the development of interrelated knowledge and understanding, skills, and attitudes needed to enable people to sustain values within their lived experience. These categories cover concepts such as values and strengths knowledge, self-esteem, self-confidence, self-regulation, communication, cooperation, leadership, and autonomy (Inman & Buck, 1998). Marsh, Richards, and Barnes (1986b) also focus on engaging in goal setting, developing resilience, and building hope and wellbeing as important developmental outcomes of OAE. All of these elements of personal and social development are closely aligned with flourishing.

Flourishing and developmental coaching. It has been suggested that having a high level of intentional goal striving is also an important element of flourishing (Grant, 2007). As a human change methodology, coaching by its nature is a goal-directed activity (Grant, 2012b). Research has found that even when coaching may be explicitly focused on goal attainment, it nevertheless often enhances wellbeing (Grant, 2003; Green et al., 2006). Consequently, beyond the focus coaching may bring to outcomes directly related to flourishing, such as wellbeing, hope, and resilience (Grant, 2017; Grant & Cavanagh, 2011; Green, Grant, et al., 2007; Green et al., 2006), the emphasis coaching places on the goal-setting and goal-striving process also has potential to influence flourishing. The importance of goal pursuit is discussed further below (see the section headed “Goal Theory”).

Constructivism and Experiential Learning Theory

Constructivism is a philosophical theory about how people learn. Although the concept of constructivism encompasses a wide range of perspectives, central to this theory is the view that knowledge is something that is actively constructed in the interaction with one’s environment, rather than an objective reality that is passively received.

Experiential learning theory (ELT) is based on constructivist thinking (David A. Kolb, 1984; 2014). Drawing from a group of “foundational scholars”,¹ David A. Kolb’s ELT aims to explain how experience is transformed into learning in a way that integrates experience, perception, cognition, and behaviour (David A. Kolb, 1984; 2014). David A. Kolb is concerned not with any experience, but with our subjective, conscious, and intentional experience. Although Kolb’s ELT has its critics (see, e.g., Bergsteiner, Avery, & Neumann, 2010; Desmond & Jowitt, 2012; Fenwick, 2001; Kayes, 2002; Miettinen, 2000; Ord & Leather, 2011; Seaman, 2008; Vince, 1998), it remains influential in learning theory and practice (Beard & Wilson, 2013; Kayes, 2002). Kolb’s experiential learning cycle (see Figure 2.2) provides a learning framework for the THP program.

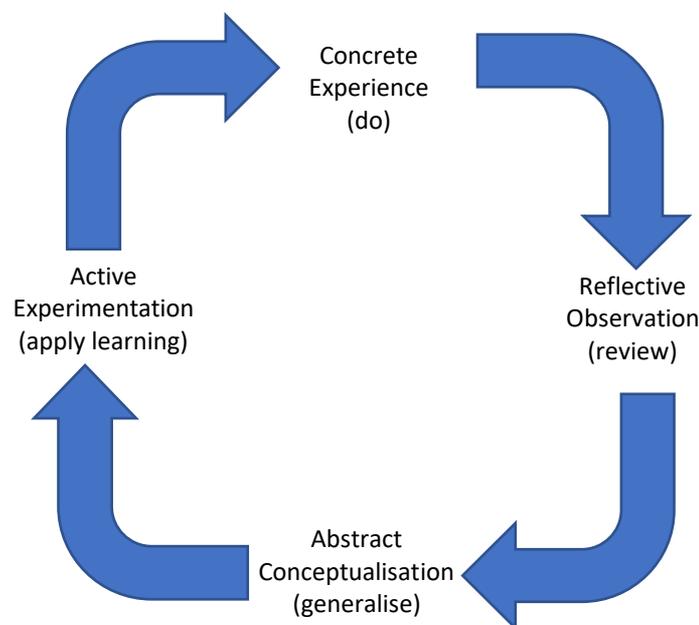


Figure 2.2. David A. Kolb’s (1984) experiential learning cycle. This figure outlines the learning process central to David A. Kolb’s experiential learning theory. Adapted from <https://www.simplypsychology.org/learning-kolb.html>.

Experiential learning theory and outdoor adventure education. A key element of ELT is the act of engaging in an experience. OAE programs provide a

¹ David A. Kolb (2014) acknowledges that his theory is grounded in the work of a group that he refers to as the “Foundational Scholars of Experiential Learning: William James, Kurt Lewin, John Dewey, Jean Piaget, Lev Vygotsky, Carl Jung, Carl Rogers, Paulo Freire, and Mary Parker Follett” (p. xvii).

rich environment of experiences perfectly suited to ELT, and research suggests that experiences engaged through ELT can produce positive outcomes (Hattie et al., 1997). Accordingly, many OAE programs draw on the principles associated with this learning theory, although which principles and models they embrace and the extent to which they do so, varies among programs (Neill, 2008). OAE and experiential learning theory function well together because OAE programs can provide not only the opportunity for experience and experimentation, but also the space for reflection and analysis. The Outward Bound solo, in particular, has been found to be a component of OAE that has a significant impact on awareness of self, others, and the natural environment (Kalisch, Bobilya, & Daniel, 2011; cf. Henderson, 2009). However, as Neill (2008) has indicated, there still exists a need for more detailed consideration of precisely how the experiential learning principles and processes interact with other theoretical components to account for outcomes in OAE.

Experiential learning theory and developmental coaching.

Experiential learning is a process, such that more than the act of engaging in experience is required. This process includes both a making sense, and transformation, of experience. More specifically, effective learning through experience requires not only being immersed in a concrete experience, but also being able to observe and reflect on that experience, and then being able to analyse those reflections into abstract concepts and generalisations that are used to inform future experiences (David A. Kolb, 1984; 2014). This learning is then applied to a new experience, beginning the next loop of this iterative process.

While an experiential learning environment (like that found in OAE) can stimulate and facilitate learning, adolescents often need to be supported in this process. Research on the Outward Bound solo, for example, has found that the developmental level of adolescents may impede their ability to use time alone for reflection effectively (Kalisch et al., 2011; Maxted, 2005). Coaches can build on experiences and add value to the ELT process by scaffolding experimentation, reflection, and abstraction. In particular, Kemp (2008) suggests that the facilitated process of reflection that coaches provide results in broader and deeper reflection,

leading to greater insight and awareness. It is this facilitated process, he argues, that makes the use of an experiential learning framework so effective within a coaching context.

Goal Theory

Goals have been described as “internal representations of desired states” (Austin & Vancouver, 1996, p. 338). Goals provide a frame of reference for making sense of our behaviour, as well as a template for the creation of meaningful life narratives that help us to satisfy our need for purpose (Crescioni & Baumeister, 2013). Incorporating goal setting into the THP program ensures there is something at stake for the participants, making them more than mere passengers on an experiential learning journey. It provides a context for reflection.

The premise of goal theory is that goals energise and direct people’s activities in organised ways and ultimately assist in giving meaning to people’s lives (Carver & Scheier, 1998; Latham & Locke, 1979; Locke, 1996; Locke & Latham, 1990; Locke & Latham, 2002; Ryan, 1970; cf. Ordóñez, Schweitzer, Galinsky, & Bazerman, 2009 and reply Locke & Latham, 2009). Setting goals is about ensuring that our activities will be personally relevant and challenging. Setting and striving for goals provide opportunities for self-regulatory development (Spence, Stout-Rostron, Van Reenen, & Glashoff, 2019).

The focus that goals bring to action and effort makes feedback a critical component of the goal-performance relationship (Locke, 1996). For goals to be effective, people need feedback about how they are progressing vis-à-vis their goal (Locke, 1996; Locke & Latham, 2002). Such feedback allows people to adjust the level or direction of their effort or their performance strategies to better meet the goal (Locke & Latham, 2002). It also can be helpful in assessing when to disengage from a goal, which is equally important to the process (see, e.g., Wrosch, Scheier, Carver, & Schulz, 2003).

Goals also affect persistence, with more difficult goals leading to prolonged effort when there is control over the time spent on a task (Locke & Latham, 2002). Further, goals indirectly impact action by triggering the use of task-relevant or

related knowledge and skills, or deliberate planning to develop such knowledge or skills where none exist (Locke & Latham, 2002).

Goals are, at once, both an outcome to strive for and a standard for judging satisfaction: “the more goal successes one has, the higher one’s total satisfaction” (Locke & Latham, 2002, p. 710). Accordingly, it is suggested that achieving one’s goals can improve wellbeing, provided that the goal is self-concordant (Sheldon & Elliot, 1999; Sheldon & Kasser, 1998). Self-concordance is about the degree to which goals are internally determined (Sheldon & Elliot, 1999). Research has found that people who pursue goals that they “own” persist longer in their efforts toward goal attainment and, therefore, are more successful at achieving their goals (Sheldon & Elliot, 1999; Sheldon et al., 2004; Sheldon & Houser-Marko, 2001). In addition, the nature of these goals as self-concordant makes it more likely that when they are attained, they will afford the experiences of autonomy, competence, and relatedness essential to wellbeing (Sheldon & Elliot, 1999; Sheldon et al., 2004; Sheldon & Houser-Marko, 2001).

The framing of goals, however, is important. Research evidence suggests that framing goals in a positive, rather than negative, way can have psychological benefits (Coats, Janoff-Bulman, & Alpert, 1996). More specifically, using approach goals, which focus on trying to attain a desired outcome, as opposed to avoidance goals, which instead focus on trying to avoid an outcome or type of behaviour, has been associated with greater academic performance and wellbeing (Elliot & McGregor, 2001). The same desired end state can often be described both as an approach or an avoidance of some outcome (Coats et al., 1996).

Further, the type of goal set has an influence both on self-efficacy and satisfaction (Bandura, 2013; Latham & Brown, 2006). The literature on goal theory has drawn a distinction between the effects of setting a learning goal versus an outcome goal, particularly when a person must acquire knowledge or skills in order to perform a task (Latham & Brown, 2006). In this situation, it has been found that setting a specific, difficult learning goal results in higher performance and higher self-efficacy than setting a specific, high, distal outcome goal, or urging the person to do their best (Latham & Brown, 2006; Winters & Latham, 1996). Moreover, it has been argued that learning motivates and invigorates people, “as they connect with

their innate capabilities and see possibilities for the future” (Spence et al., 2019, p. 4).

Goal theory and outdoor adventure education. In considering what aspects of OAE programs lead to positive outcomes, research has found that setting and pursuing goals are critical components of such programs (Marsh, Richards, & Barnes, 1986a; McKenzie, 2000). Hattie et al.’s (1997) leading review on the impact of OAE proposed that difficult goals were a key element of effective OAE programs. Moreover, OAE provides a rich environment for setting specific and challenging goals (Crane, Hattie, & Houghton, 1997; Neill, 2008). Research on goal setting in the context of OAE also has found that specific goals were more effective than vague goals, however, without any direction only a third of goals set were specific (Crane et al., 1997). Accordingly, goal setting needs to be expressly built into OAE programs in order to ensure that effective goals are established. Of further importance to successful outcomes is the goal striving process, for which feedback and support are essential (Locke, 1996). OAE programs provide many opportunities for feedback, both from the natural environment in which they are based, and from others involved in the program, including the instructors and other participants (Hattie et al., 1997; Neill, 2008). Moreover, feedback within the microcosm of the OAE environment is often immediate.

Goal theory and developmental coaching. Both the process of goal setting and the steps for making progress toward goal attainment, in a positively-focused and supportive environment, are considered core components of effective coaching (Grant, 2006, 2007, 2012b; Grant & Cavanagh, 2007). In a recent within-subjects study comparing four aspects of the coach-coachee relationship, Grant (2014) found that a goal-focused relationship was a statistically unique and significantly more powerful predictor of coaching success than the other relationship aspects (satisfaction with the relationship, autonomy support, and proximity to an “ideal” relationship).

Grant (2012b) has developed a coaching approach that strongly links goal pursuit with self-regulation and experiential learning. Using this approach, the coach facilitates the coachee’s movement through a cycle in which the individual:

(a) sets a goal; (b) develops a plan of action; (c) sets the action in motion; (d) monitors performance; (e) uses a standard of comparison by which to evaluate performance; and (f) changes any actions that will further enhance the person's ability to better reach the goal. Coaches facilitate goal attainment through this process by helping people to identify self-concordant goals, increase motivation through development of character strengths and self-efficacy, identify resources, and manage the monitoring and evaluation of the self-regulatory cycle (Grant et al., 2010). Research illustrates how coaching may positively impact goal self-concordance, alignment with personal values, and goal commitment and suggests that these interactions may be some of the mechanisms through which coaching is effective as a support for goal attainment (Burke & Linley, 2007).

Challenge and Support

From a social constructivist perspective, learning depends on interactions with others in an environment that involves both challenge and support. Vygotsky (1978), a leading social constructivist, referred to this environment as the *zone of proximal development*. In this challenging learning space, the learner requires the support of a *more knowledgeable other*. Being in such an environment is a powerful formula for growth (Sanford, 1962, 1966; Vygotsky, 1978; see also Rathunde, 1996).

The experience of engaging in challenging activities through proximal goal pursuit has been described as being in *flow* (Nakamura & Csikszentmihalyi, 2014). Research has found that engaging in challenging tasks is beneficial not only for learning, but more generally for a person's wellbeing (Csikszentmihalyi, 1992). The level of challenge, however, is a balancing act. If a task is too simple, it can lead to boredom (Csikszentmihalyi, 1992; Nakamura & Csikszentmihalyi, 2014); activities that offer no challenge are unlikely to be motivating or to bolster one's sense of efficacy upon completion (Crescioni & Baumeister, 2013). On the other hand, an overly challenging task can create anxiety (Csikszentmihalyi, 1992; Nakamura & Csikszentmihalyi, 2014), and tasks that are so difficult as to assure failure may have a negative impact on one's self-efficacy and self-esteem (Crescioni & Baumeister, 2013). Therefore, finding one's level of optimal experience is likely to be important to positive outcomes. Optimal experience results from the opportunity to engage in

challenging activities together with possession of the capacity (through some combination of skill and support) to meet those challenges.

While support can be provided in a variety of ways, the concept of support applied in the THP program is human support (or scaffolding) in which dialogue and the active development of a shared understanding are the primary mechanisms that allow for growth (for a fuller description of scaffolding and its historical context, see Shvarts & Bakker, 2019). An appropriate balance of challenge and support requires one to provide a “careful calibration of the support” based on a person’s current level of competence and in a way that is continually adapted to their ongoing development (Stone, 1998, p. 349). Also essential is for such support to be faded over time, thereby resulting in transfer of responsibility for the task or skill as the person develops (Van de Pol, Volman, & Beishuizen, 2010). Accordingly, support must be fluid, rather than routine.

Challenge, support, and outdoor adventure education. While research indicates that a range of OAE program activities can lead to positive outcomes, it has been suggested that the qualities of the activities, rather than the activities themselves, are essential to these outcomes (Kemp, 2006; McKenzie, 2000, 2003; Walsh & Golins, 1976). One quality, in particular, is the controlled exposure to challenge that is involved in these activities (Martin & Leberman, 2005; McKenzie, 2000, 2003; Sheard & Golby, 2006; Walsh & Golins, 1976). Such challenges take participants outside their comfort zone with the potential to generate new perceptions of what is possible (Carpenter & Harper, 2016; Taniguchi & Freeman, 2004) The flexible nature of OAE programs allows for modifications at the individual level to provide experiences that are challenging for diverse participants (Sibthorp et al., 2015).

Much of the philosophy underlying OAE emphasises the need to be at the edge of one’s “physical and psychological possibilities” in order to stimulate growth (Neill & Dias, 2001, p. 1). It is maintained that the challenge in OAE creates constructive anxiety or a state of dissonance, and using one’s mental, emotional, and physical resources to successfully traverse the challenge and overcome the dissonance is what can lead to development (see the research summarised by

McKenzie, 2000; see also Hofer & Pintrich, 1997). Priest (1993) has provided a more detailed model of the interaction between perceptions of risk/challenge and competence, performance, feedback loops, and attribution, which demonstrates the complex relationship between challenge and growth. Importantly, it is the subjective perception of challenge and one's capacity to meet the challenge, rather than an objective one, that determines the quality of these experiences (Nakamura & Csikszentmihalyi, 2014). As a consequence, a person's capacity to make meaning of challenging experiences and their self-efficacy for meeting those challenges will influence the outcomes of OAE programs.

What seems to be essential to learning through challenge, is not only having the right level of challenge, but also the need for support in achieving the appropriate level. Research undertaken by Neill and Dias (2001) on an Outward Bound program, has found empirical support for the benefits of challenge, specifically where there was social support for the challenge (see also, Ewert & Sibthorp, 2009; Sibthorp & Jostad, 2014). Together with difficult goals and feedback, challenge and support are key elements of effective OAE programs (Hattie et al., 1997; Martin & Leberman, 2005).

Challenge, support, and developmental coaching. While an OAE program provides the opportunity for growth, it is the process of engagement in that program that is key to the developmental outcomes (Kemp, 2006). Recent research in OAE suggests that structured facilitation, focused on relevant psychological theory, may have a key role to play in effective developmental outcomes (O'Brien & Lomas, 2017). Skilled coaches can influence both the challenge and the support that form part of the development process. Accommodative (or vertical) development often requires that a person be held in a space of tension, paradox, and ambiguity in order to see the limitations of their current ways of understanding, and to allow more adaptive perspectives to emerge (Cavanagh, 2016). Developmental coaches are trained in effective dialogue that can both affirm a person's current perspective and create the tension and paradox to challenge that perspective (Laske, 2007). Nevertheless, while tension and paradox

may create the impetus for growth, it is support that provides the possibility for growth.

The difficulty inherent in accommodative change often necessitates scaffolding. Generating more complex ways of making meaning is challenging, and in the absence of support, people will often fall back on their existing frames of reference. The scaffolding that coaches provide can help to hold people in a place of tension, connected to what is inadequate about their current ways of making meaning, long enough to enable them to build new ways of understanding. In this way, coaches function as an “evolutionary bridge, a context for crossing over” (Kegan, 1994, p. 43).

The stress and anxiety that is an integral part of the growth process underscores the importance of trust and a safe, empathic environment (Cavanagh, 2006; Vincent, 1995). Developmental coaches are capable not only of maintaining an emotionally safe space for doing transformational work, but also of supporting the skill-building and development itself. Through shared dialogue, a developmental coach can help a person to suspend their frame of reference, to be in a position to positively examine their own perspectives, as well as to engage the tension between perspectives, so that those perspectives can be explored and understood (Cavanagh, 2016). It is in this space that the energy and information exists for new, more adaptive perspectives to emerge (Cavanagh, 2016).

Also relevant is the consideration that coaches are not teachers and can use this distinction to step away from being a figure of authority or the expert that adolescents at their developmental level may look to for the “right” answers. This puts the coaches in a special position to engage and support the participants in a way that has the potential to make the experiential learning process more influential. It is the inclusion of a skilled coach whose primary job is to stimulate and scaffold the process of accommodative change that differentiates the THP program from other OAE programs.

Constructive-Developmental Theory

At the heart of many of the above processes and theories is one’s capacity for meaning-making (or perspective-taking). Kegan’s (1982, 1994) constructive-

developmental theory is a theory of the development in the way one constructs meaning of their experiences,² and it provides a useful framework for considering how the participants make meaning of their experiences in a THP program. Kegan uses the concept of *meaning-construction* to describe the active construction of reality that constructivism speaks to (Kegan, 1994). For Kegan (1994), meaning-construction is not about what we think or exclusively about how we think, but rather about the way in which we construct meaning more broadly: “It is about the organising principle we bring to our thinking and our feelings and our relating to others and our relating to parts of ourselves” (p. 29).

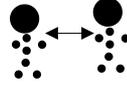
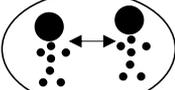
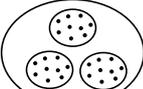
In contrast to Rogers’ focus on separation as the intrinsic process of adaptation and growth, Kegan (1982) considers development toward integration as being equally important to development toward differentiation. Growth, for Kegan, involves both an emergence from embeddedness and a new relating to that which was previously embedded. Piaget’s cycle of assimilation and accommodation is central to the constructive-developmental process, being a cycle that is marked by periods of balance followed by periods of instability and then a new kind of balance. This development reflects a continuous process of negotiation between two conflicting motivations: autonomy and belonging (Bachkirova, 2010). The focus of this process of negotiation and transition is twofold: (a) those elements of our knowing or organising that are so fundamental to us that we are *Subject* to them, and (b) those elements that can be made *Object* (Kegan, 1994). Things that are *Subject* in constructive-developmental theory are invisible to us; they can force us to act but cannot be observed or reflected upon (Kegan, 1994). Things that are *Object* for us are those elements of our knowing that we can take into our perspective, reflect on, and take control of: “We have object; we are subject” (Kegan, 1994, p. 32).

In constructive-developmental theory, there are five stages of qualitatively different ways of constructing meaning, referred to as natural epistemologies, with each stage having a Subject/Object relationship where what is Subject in one stage

² Kegan (1982) acknowledges that his theory originates from the work of Jean Piaget, as well as James Mark Baldwin, John Dewey, and George Herbert Mead.

becomes Object in the next (Kegan, 1994). These stages are outlined in Table 2.1. Development from one stage to the next is gradual and always in the direction of greater complexity. The more a person can take as Object, the more complex that person's perspective becomes because he or she can examine and act on more. However, people are rarely at a distinct stage. Rather, they are generally somewhere between stages and, therefore, are subject often to the conflicting interests of two stages. Rather than refer to these stages hierarchically by number, this thesis uses alternative descriptive terms by which these stages have come to be known: the *Impulsive* way of knowing (stage one), the *Instrumental* way of knowing (stage two), the *Socializing* way of knowing (stage three), the *Self-Authoring* way of knowing (stage four), and the *Self-Transforming* way of knowing (stage five; Kegan, 2000).

Table 2.1
Kegan's Five Stages of Meaning-Making

Constructive-Developmental Stage	Subject	Object	Underlying Structure
Stage One "Impulsive" <i>Fantasy-based</i>	One's impulses, perceptions	One's reflexes	Single point • <i>momentary perception</i>
Stage Two "Instrumental" <i>Rule-based</i>	One's needs, interests, desires	One's impulses, perceptions	Durable categories  <i>concrete, enduring</i>
Stage Three "Socializing" <i>Other-focused</i>	One's mutually reciprocal relationships	One's needs, interests, desires	Across categories  <i>bigger than self</i>
Stage Four "Self-Authoring" <i>Reflective</i>	One's self-determined values, identity, ideology	One's mutually reciprocal relationships	Systemic  <i>complex, multiple role consciousness</i>
Stage Five "Self-Transforming" <i>Interconnecting</i>	The inter-connected-ness of systems	One's self-determined values, identity, ideology	System of systems  <i>paradoxical, trans-complex</i>

Note. Adapted from Kegan, 1994 and Peter W. Pruyne (<http://developmentalobserver.blog.com>).

The Instrumental, Socializing, and Self-Authoring stages are most relevant to the developmental level of participants in a THP program, as adolescents are likely to be somewhere in, or transitioning between, these stages. Instrumental knowers are focused on concrete and specific understandings of themselves and others (Berger, 2003). They construct meaning through the filter of their own needs, wants, and interests. While they can distinguish other's perspectives from their own, they cannot hold another's perspective and their own at the same time. They

are focused on rules and consequences, and they feel supported when others provide specific advice and explicit procedures so that they can accomplish their goals. Students at this stage will adopt the teacher's point of view but will not be able to reflect on it or consider it in relation to their own perspective (McCann, 2005). They view knowledge as a possession, an accumulation of facts and skills, and focus on finding the right answers and the correct way of doing things (Drago-Severson et al., 2001). At this stage, reflection and making abstractions is difficult (Kegan, 1994).

A Socializing knower can make abstract generalisations, self-reflect, and be devoted to something that is greater than their own needs (Berger, 2002). However, at this stage people internalise the perspectives of valued others and subordinate them to their own, so that it becomes impossible for them to develop their own point of view, and they find it difficult to resolve conflicting ideas (Berger & Atkins, 2009). Knowledge is seen as something that comes from authority and experts who hand down truth (Helsing, Drago-Severson, et al., 2001). This relationship to public authority is at odds with self-regulation, which requires a sense of personal authority (Kegan, 1994, p. 275). As the adolescent participants in the THP program may be operating at this or an earlier constructive-developmental level, the impact that their capacity for meaning-making has on their ability to self-regulate is important to note, since self-regulation is one of the key outcomes of the THP program. Consequently, scaffolding will be essential to any improvement in self-regulatory skill.

It is less likely for adolescents to be making meaning fully at the Self-Authoring stage (Baxter Magolda, Creamer, & Meszaros, 2010). However, this stage is important as it is the form of mind that the program structure and those who implement the program may be apt to expect. At this stage, people can recognise, understand, generate, and evaluate various standards, values, and ideas, and can mediate among them using their own self-governing system (Berger & Atkins, 2009). Knowledge at this stage can be understood as constructed (Helsing, Drago-Severson, et al., 2001). Students at this stage can create and explain their own complex ideas and are comfortable holding ideas or opinions that differ from their teachers' (Drago-Severson et al., 2001). They are able to take responsibility for their

own learning and can evaluate experiences by reference to their own self-constructed goals (Drago-Severson et al., 2001).

As indicated in the previous section, OAE provides healthy and challenging environments for exploring self-awareness and broader perspective-taking, and it is these types of environments that can provide the impetus for transformational growth. However, an OAE program developed by adults may be designed with an expectation for a level of meaning-construction that exceeds the capacities of its adolescent participants, resulting in what Kegan (1994) describes as a “mismatch between external epistemological demand and internal epistemological capacity” (p. 41). For example, the requirements in the experiential learning framework for reflective observation and abstract conceptualisation can be difficult for adolescents. Building self-regulation strategies can also be challenging, as it requires participants to take ownership of their thoughts, feelings, and behaviours; to understand themselves as separate from others. These gaps in understanding create a challenging environment for the participants that must be met with ample support. However, if OAE program facilitators make different meaning of these expectations, they may not comprehend the challenges experienced by the participants to offer the necessary support. It has been suggested that developmental learning frameworks can provide important awareness and understanding for educators, including in the field of OAE (Collins, Paisley, Sibthorp, & Gookin, 2012). It is anticipated that skilled coaches with a sound understanding of constructive-developmental theory will be able to recognise and work from a participant’s developmental level and that doing so will provide important benefits for program participants.

Previous Research Findings and Existing Gaps in the Literature Outdoor Adventure Education

Previous research findings. The field of outdoor adventure education encompasses a wide range of experiential programs for a variety of populations with diverse aims. Similarly, the research findings on OAE programs also vary widely, owing to a range of protocols for conducting and measuring program effectiveness (Cason & Gillis, 1994; Hattie et al., 1997). One particular challenge is the outcome

measure of personal and social development, which is an outcome that is especially difficult to measure in any field (Dillon, 2012).

Beyond variability, much of the research that has been conducted in the field of OAE has been found to be methodologically weak, owing to a lack of control groups, lack of randomisation, and lack of longitudinal follow-up (Barrett & Greenaway, 1995; Cason & Gillis, 1994). Hattie et al. (1997) suggested that many of the research papers they considered for their meta-analysis “read more like program advertisements than research” (p. 45). They indicated that those studies that had looked at outcomes, were primarily one-off studies using pre- and post-test comparisons with small samples, and ignoring important variables such as program length, facilitator experience, and program differences. Additional issues include the number of studies that are unpublished and the bias in published studies toward statistically significant findings (Neill, 2008; Sibthorp & Arthur-Banning, 2004).

Despite the empirical issues, there has been research that demonstrates OAE has a small-to-moderate short-term effect on various aspects of the personal and social development of adolescents (Bowen & Neill, 2013; Cason & Gillis, 1994; Ewert, 1987; Ewert, Garvey, Prouty, Panicucci, & Collinson, 2007; Ewert & Yoshino, 2011; Hattie et al., 1997; Hayhurst, Hunter, Kafka, & Boyes, 2015; Hunter et al., 2010; Marsh et al., 1986a, 1986b; Neill, 2008; Neill & Dias, 2001; Paxton & McAvoy, 2000; Propst & Koesler, 1998; Rose, Williams, Olsson, & Allen, 2018; Scarf et al., 2018; Sibthorp, Paisley, Furman, & Gookin, 2008), including disadvantaged adolescents (Norton & Watt, 2014). The outcomes arising from these programs also have been found to lead to improvements in educational outcomes which continue over the long term (Hattie et al., 1997; Marsh et al., 1986a; Norton & Watt, 2014). In particular, research has found OAE programs to be associated with increases in participants’ self-esteem, self-efficacy, self-confidence, self-regulation, and resilience, as well as other important life skills both technical and social (e.g., Davidson, 2001; Dillon, 2012; Ewert & McAvoy, 2000; Hans, 2000; Hattie et al., 1997; Mygind et al., 2019; Propst & Koesler, 1998; Richmond et al., 2018).

The first meta-analysis in the field of OAE was undertaken by Cason and Gillis (1994). They reported on 147 effect sizes from 43 studies conducted on an

adolescent population within the 25 years prior to 1994, finding an average effect size of .31. They further reported a significant positive correlation between program length and effect size. However, there was no distinction found between the types of participants (e.g., normal versus delinquent).

In their meta-analysis, Hattie et al. (1997) estimated 1,728 effect sizes from 96 studies published over roughly the same period as the Cason and Gillis meta-analysis. However, the majority (75%) of participants in these studies were adults or university students. They found an average short-term effect size of .34, as well as a generally positive follow-up effect, indicating that the effects continue to increase over time but noting that there are marked variations among these follow-up effect sizes and that any positive follow-up effect size could be the result of a *sleeping effect* (referring to the delayed impact of an intervention). In fact, the results showed that only some OAE programs were effective and only on some outcomes, with it being suggested that only some components of these programs are influencing the outcomes. As with Cason and Gillis, longer programs had greater effects. Owing to a lack of proper background variables, the reviewed research did not lend itself to determining whether there was any distinction between type of participants. In terms of the outcomes, those with the greatest and most long-lasting effects included independence, confidence, self-efficacy, self-understanding, assertiveness, internal locus of control, and decision-making. Given that these outcomes all relate to a sense of control over the self, Hattie et al. (1997) suggested that among the various outcomes measured, OAE programs appear to be most effective at developing self-regulation.

Adventure therapy programs are closely related to OAE, differing mainly in that the primary aim is to help participants deal with their psychosocial problems (Bowen & Neill, 2013). Nevertheless, adventure therapy programs often incorporate outdoor activities and experiential learning exercises as key components of their program (Bowen & Neill, 2013). They also have additional aims, program elements, and processes similar to OAE. As a consequence, evidence from adventure therapy research can provide useful insight into OAE. Bowen and Neill (2013) recently conducted the most comprehensive and robust meta-analytic review of adventure therapy programs to date. This review considered 2,908 effect sizes from 197 studies

of adventure therapy programs published between 1967 and 2012, comparing the short-term and follow-up effects of those programs with alternative treatment and no treatment groups. The authors found a moderate, positive, and significant pre-post effect size of .47 for the adventure therapy group which was greater than the small, positive, and significant effects for the alternative treatment (ES = .14) and no treatment (ES = .08) groups. The follow-up effect for the adventure therapy group was very small, positive, and non-significant (ES = .03). This positive effect is in addition to the short-term effect and although non-significant, is quite important because there is typically a diminishing of outcomes over time in almost all intervention research. The results, therefore, indicate that the short-term gains of adventure therapy are retained over the longer-term. Participant age was found to be the only significant moderator of the effectiveness of these adventure therapy programs. Other program and participant characteristics were found to explain little variance. Based on this study, Bowen and Neill (2013) have suggested that adventure therapy programs with overall effect sizes between .40 and .60 are within the expected range. However, they note that effect sizes between .30 and .50 are more typical of programs for adolescents, with effect sizes between .50 and .70 being more typical of programs for participants aged 18 years and older.

Existing gaps in the literature. While the research findings on the impact of OAE have been largely positive, only some programs are effective and then only in relation to some outcomes (Hattie et al., 1997; Sheard & Golby, 2006). Moreover, the research is limited and what exists is undermined by criticisms of a lack of methodological rigour in the research (Neill, 2008; Scrutton & Beames, 2015). Furthermore, a number of prior studies have specified a need for longitudinal research to consider program outcomes over time (Ewert & McAvoy, 2000; Sibthorp, 2003; Sibthorp et al., 2008), as well as a greater focus on OAE for youth experiencing various risk factors, including disadvantage (Norton & Watt, 2014). The present research proposes a robust design and rigorous statistical analysis that aim to meet the criticisms of existing research.

In addition to more robust research designs and analyses, a call has also been made for future research to be directed at better understanding which program

variables and theoretical elements are most related to developmental outcomes (Hans, 2000; McKenzie, 2000; Neill, 2008; Sheard & Golby, 2006; Sibthorp & Arthur-Banning, 2004). One particular difficulty in the research relates to the nature of the most established outdoor adventure programs, such as Outward Bound and the National Outdoor Leadership School; participants in these programs generally come together for the specific purpose of their involvement in the program outside of their day-to-day communities. It has been suggested that intact groups (e.g., from the same school or workplace) will have a better opportunity to reinforce the program outcomes (Sibthorp & Jostad, 2014). Moreover, research on extracurricular activities has found that school-based extracurricular activities are more beneficial than activities outside of school (Marsh & Kleitman, 2002). In particular, research suggests that having a sense of belonging and support within the school community can improve student engagement and motivation, increasing opportunities for academic success (Anderman & Freeman, 2004; Sibthorp & Jostad, 2014). THP works with schools to provide their programs through the school to intact school groups of students. Furthermore, THP offers its program groups funding (following a successful pitch) to undertake a community project following the program, providing additional opportunities for participants to reinforce their newly acquired skills and other outcomes, as well as opportunities to further engage with their program group and their community. Consequently, the present investigation will be in a position to consider these aspects of OAE and extracurricular activities more closely.

More research also is needed to better understand how program aspects such as program type, program length, group size, and participant characteristics (e.g., age, gender, personality, experience) impact on program outcomes. Work needs to be done to establish the full value of OAE, not only for the adolescent participants but for their schools and community as well (Dillon, 2012). The facilitators of OAE programs are the program variable that has been given the most attention by researchers (McKenzie, 2000). Their importance to the experiential learning process and transfer of learning has been emphasised in the OAE literature (Ewert & McAvoy, 2000; Gass & Gillis, 1995; Hattie et al., 1997; Luckner & Nadler, 1997; Martin & Leberman, 2005; Martin & Legg, 2002; Sibthorp et al., 2011; Sibthorp,

Paisley, & Gookin, 2007). It has also been suggested that the quality of the processing and coaching that occurs prior to, during, and following any adventure experience is critical to successful outcomes (Kemp, 2006; see also Paisley et al., 2008).

Skilled facilitators can enhance the opportunities for, and quality of, feedback, reflection, and other forms of processing in OAE programs (Hattie et al., 1997; Kemp, 2006; McKenzie, 2000). They have been found to be instrumental in curriculum delivery, but also as role models, mentors, supporters, and sources of inspiration (McKenzie, 2003; Sibthorp et al., 2011). Other valuable facilitator qualities include an attitude that is encouraging and non-judgemental, empathy, presence, good listening skills, effective communication, appropriate expectations, and flexibility (McKenzie, 2000; Vincent, 1995). In general, a humanistic approach that respects each participant's capacity for development is said to be essential for OAE facilitators (McAvoy, Mitten, Stringer, Steckart, & Sproles, 1996). However, more information is required, particularly in relation to the interaction between facilitator qualities, behaviours, and attitudes and program effectiveness (Kemp, 2006; McKenzie, 2000). Typically, OAE facilitators are not trained in developmental psychological approaches. As indicated above, this training is important for supporting accommodative change. It is suggested, therefore, that in addition to the skilled adventure program facilitators, having experienced coaches with developmental coach training and appropriate supervision will maximise the opportunities for the most effective experiential learning process, which in turn will have a positive impact on the program outcomes.

Finally, it has been argued that multidimensional research approaches are needed for understanding the complex processes and benefits associated with OAE programs, including the relationship among program participants, the various program elements, and the numerous outcome measures (Barrett & Greenaway, 1995; Ewert & McAvoy, 2000; Harper, 2010; Klint, 1990; Martin & Leberman, 2005; McKenzie, 2000; Rowley, 1987). In particular, it has been suggested that OAE providers need to better understand the developmental characteristics of students in order to adapt OAE programs to meet their capabilities (Collins et al., 2012). The

mixed-methods design and application of a constructive-developmental lens in this research seeks to meet these needs.

Developmental Coaching

Previous research findings. The empirical literature on coaching is young, with the majority of studies occurring after 2000. Additionally, the bulk of this research comes from non-experimental studies. Accordingly, the evidence base for coaching psychology remains limited (Grant, 2012a; Grant & Cavanagh, 2007; Grant et al., 2010; Lowman, 2005; Whybrow, 2008). Between 2000 and 2011, there have been 234 outcome studies focused on the effectiveness of coaching: 131 case studies, 77 within-subject studies, and 25 between-subject studies (only 14 of which used a randomised controlled design) (Grant, 2012a).³ The research that does exist has been described as “disjointed and somewhat fragmented” (Grant, 2012a). Notwithstanding the early stage of development in coaching research, there is an emerging body of empirical support for the effectiveness of coaching, particularly with respect to executive coaching and life coaching for adults. In particular, research has begun to demonstrate an association between coaching and the positive development of self-regulation, resilience, hope, and wellbeing, even when these psychological constructs are not the main goals of the intervention (Grant, 2016b).

In the first randomised controlled study of short-term executive coaching by professional external coaches, Grant, Curtayne, and Burton (2009) found that coachees experienced statistically significant increases in goal attainment, resilience, and workplace wellbeing, as well as reduced depression and stress, when compared with the control group. A recent meta-analysis of coaching outcomes and moderators from 18 studies conducted in an organisational context found coaching to have significant positive effects on various individual-level psychological outcomes, including an effect size of .60 on performance/skills, an effect size of .43 on coping, an effect size of .74 on goal-directed self-regulation, and an effect size of

³ For an analysis of the academic coaching literature and empirical studies prior to 2000, see the literature review by Grant (2002), and for a review of the empirical literature between 1980 and 2007, see Grant and Cavanagh (2007). See also Grant et al. (2010).

.46 on wellbeing (Theeboom, Beersma, & van Vianen, 2014). The number of coaching sessions, however, was not found to moderate the effectiveness of the interventions. Jones, Woods, and Guillaume (2016) also conducted a meta-analysis on coaching outcomes and practice moderators from 17 studies of workplace coaching, and found coaching to have positive effects on all outcomes, which they categorised as affective outcomes (with an effect size of .51), skill-based outcomes (with an effect size of .28), and individual-level results outcomes (with an effect size of 1.24). They also found significant moderation of effect size for type of coach (with internal coaches leading to stronger effects) and use of multisource feedback (with use of multisource feedback resulting in smaller positive effects). No moderation of effect size was found for coaching format or duration of coaching. A further meta-analysis on coaching in organisations conducted by Grover and Furnham (2016) found coaching to be effective for improving self-efficacy and goal attainment in individual participants, but also considered there to be an open question on whether self-efficacy is an outcome of coaching or a predictor of coaching effectiveness, or whether there is a reciprocal relationship between the two.

While aimed at goal attainment, the life coaching program used in Grant's (2003) within-subjects study was also associated with positive impacts on the participants' mental health and wellbeing, through observed reductions in depression, stress, and anxiety, and improvements in general life satisfaction. Building on this work, the randomised controlled study by Green, Oades, and Grant (2006) found its group, life coaching program was associated with statistically significant increases, not only in goal striving but also in wellbeing and hope, with change being maintained at a 30-week follow-up on some of the outcome variables.

A study by Green, Grant, and Rynsaardt (2007) is one of the few studies to investigate the impact of life coaching on adolescents. The random waitlist control design involved 56 female, senior high school students. The participants in the intervention group were given 10 coaching sessions with teachers trained in the theories and techniques of coaching psychology. This coaching was associated with significant increases in the participants' levels of cognitive hardiness (a form of resilience) and hope, and significant decreases in their levels of depression, when compared to the waitlist control group.

In relation to primary school boys, a strengths-based coaching program at a private school in Sydney, Australia, was found to be associated with statistically significant increases in engagement and hope (Madden et al., 2011). After taking the Youth Values in Action survey to identify their character strengths, 38 male students in year five, participated in eight group coaching sessions aimed at helping them to identify personally meaningful goals, persist in their goal-striving, and find new ways to use their signature strengths. The program also provided the students with the language of strengths, allowing them to recognise and talk about strengths in themselves and others. In addition to the program's association with increases in aspects of wellbeing, the study found indications of other potential benefits, including increases in motivation and performance, as well as the positive impact that increased wellbeing is likely to have on learning in general.

Only recently have there been studies that have focused specifically on coaching at-risk adolescents. An action research study conducted by Robson-Kelly and van Nieuwerburgh (2016) sought to identify what coaching psychology has to offer adolescents at risk of developing mental health problems. Based on their research, the authors suggest that the coaching experience provides a process, a relationship, and skills that enable adolescents to develop choice and control over their thoughts, feelings, and behaviours, leading to increased confidence. The authors propose that this occurs through growing accountability, awareness, and responsibility. Their coaching program involved individual and group coaching, and both the adolescents and the coaches found the combination of approaches to be the most effective intervention.

Another qualitative study with three adolescent girls from an inner-city London school, sought to analyse their perceived changes in quality of life subsequent to participating in an integrated coaching and positive psychology intervention program (Pritchard & van Nieuwerburgh, 2016). This research uncovered three key changes from participation in the program: an increased ability to control emotions and reactions, increases in the experience of positive emotions and thoughts, and an ability to identify purpose and meaning in life. These changes were considered to be related to increased engagement and accomplishment in school, improved relationships, and higher perceived quality of life.

Existing gaps in the literature. While the research base in coaching psychology is increasing, there is concern that much of the research work lacks the rigour necessary for this field to develop as a scientific enterprise (Grant & Cavanagh, 2007; Lowman, 2005). A good portion of the empirical research is contextual or survey-based, rather than outcome focused (Grant & Cavanagh, 2007; Grant et al., 2010). In respect of the outcome studies that do exist, the lack of randomised controlled studies is said to be a serious shortcoming (Grant, 2012a; Grant et al., 2010; Lowman, 2005). Moreover, the outcome measures that have been used are seen as weak in terms of their validity and reliability (Grant, 2012a; Grant et al., 2010; Lowman, 2005). As an emerging discipline, what is needed is more sophisticated research designs and analysis into the effectiveness of coaching as a methodology for creating and sustaining human change, and the development of coaching-specific theory (Grant, 2012a, 2016b; Grant & Cavanagh, 2007). Whybrow (2008) also has emphasised that the development of the coaching psychology profession requires increased peer-reviewed research in high quality publications focused more specifically on coaching psychology. Further, the study by Green et al. (2006), although promising, is one of the few studies to use a longitudinal design. As coaching is fundamentally about change, Linley (2006) points to the need for longitudinal research designs in order to test the sustainability of outcomes and further inform coaching practice (see also Blackman, Moscardo, & Gray, 2016). Spence et al. (2019) also suggest that longitudinal research is necessary to capture any delayed effects anticipated in transformational change. For all of these reasons, solid research is a priority, using well-validated measures of positive outcomes. The present research design aims to meet these research needs in the field of coaching psychology.

In addition to more rigorous quantitative analysis, there is also a call for continued qualitative analysis, noting that such analysis provides “unique insight into complex phenomenon, such as coaching” (Grover & Furnham, 2016, p. 27). Qualitative analysis may also allow for a greater understanding of the longitudinal impact of coaching. The mixed-methods approach used in this thesis endeavours to satisfy the need for both quantitative and qualitative data in coaching research.

Finally, much of the research into the effectiveness of coaching has focused on executive and leadership coaching, and to a lesser extent on life and health coaching (Passmore & Brown, 2009). There is only limited research on coaching in the educational context, and what exists is more focused on the teachers and administrators in that space (Passmore & Brown, 2009). Where there has been research concerned with helping students to thrive, the work has primarily involved middle-class and private school students (Pritchard & van Nieuwerburgh, 2016). This area of research needs to be deepened and extended to at-risk adolescents (Pritchard & van Nieuwerburgh, 2016). Accordingly, the THP program's focus on disadvantaged adolescents meets a pressing need for coaching research in this area.

Integrating Developmental Coaching with Outdoor Adventure Education: Implications for the Present Investigation

Based on the above, it is apparent that both OAE and coaching have been associated with increases in various positive outcomes that enable individuals (including adolescents) to flourish. Both OAE and coaching have also been subject to similar criticisms from an empirical standpoint.

The critical importance of processing of the experience to OAE outcomes, makes the integration of skilled coaches in OAE a variation worthy of exploration. Because coaches with psychology training have a greater awareness of the different developmental levels from which adolescents may perceive their experiences, and a deep understanding of goal theory and change processes, it is proposed that they can stimulate and support the experiential learning process in a way that increases the outcomes of OAE programs, thereby helping those adolescents to reach their full potential. It has also been argued that OAE methodology, with its positive-developmental and human-centred focus, provides a valuable framework and philosophical foundation for coaching more broadly (Kemp, 2006). Each approach, therefore, has potential application for the other. By combining OAE and developmental coaching, we can further advance the research base of each, while also testing a new approach with the potential to improve both OAE and coaching outcomes. To date no empirical research has been found which has tested the effectiveness of the integration of OAE and developmental coaching. Furthermore,

as a school-based extracurricular program, the THP program provides an opportunity to further assess the benefits of school-based extracurricular programs, as well as to compare the differential effects of these program modes.

In addition, this research uses a random waitlist control design and well-validated measures of a number of important positive outcomes. Longitudinal data will also provide the opportunity to assess the full impact of both OAE and developmental coaching over time. Finally, the use of a mixed-methods approach makes it possible obtain the deepest understanding and most complete picture of the THP program effects. This research, therefore, has the potential to make a significant contribution to the developing empirical base in the areas of OAE, coaching psychology, positive psychology, educational psychology, and adolescent development.

Summary

This chapter has provided an overview of the theories central to the THP program and how those theories are relevant to OAE and coaching. Some of the existing empirical literature on both OAE and coaching psychology was then reviewed, as well as gaps in the existing literature. Questions remain as to which theoretical elements and processes are most effective for OAE. Of particular interest is facilitation of the experiential learning process in outdoor adventure programming. Based on the theory and existing empirical research, it is suggested that integrating developmental coaching into OAE may improve OAE outcomes through more skilled stimulation and support of the experiential learning process. The core elements of the THP program are covered in Chapter Four. The next chapter provides a detailed analysis of the measurement outcomes applied in this research and the theoretical constructs underlying those outcomes.

CHAPTER THREE

OVERVIEW OF QUANTITATIVE RESEARCH OUTCOMES: THEORETICAL CONSTRUCTS AND THEIR MEASUREMENT

Introduction

The primary outcomes of interest in this thesis are abstract concepts that cannot be directly observed. In the social and behavioural sciences, these concepts are characterised as *constructs*. Social and behavioural researchers develop and study constructs for the purpose of better understanding how individuals function in the world. For a construct to be capable of analysis, it is necessary to define the construct in a way that allows for its empirical measurement. Using theory and previous research in connection with a construct, an instrument containing items (also referred to as *indicators* or *manifest variables*) that represent the characteristics of that construct can be developed to measure the construct (also referred to as the *variable, factor, or scale*).⁴

This chapter describes each construct used in the quantitative research for this thesis, including the theoretical underpinnings of each construct and its relationship to the THP program, as well as the instruments used to measure these constructs. A complete list of all measurement instruments and their items, including item cross-references to the original instruments and the Survey, can be found in Appendix A. Further details on the psychometric properties of these measurement instruments is included in Chapter Five. As OAE research suggests that program outcomes should be related to program design and aims (Hattie et al., 1997; Sibthorp et al., 2007), comment is included in this chapter on the relevance of each scale to the THP program design and aims. The quantitative analysis used to

⁴ Where an instrument is multidimensional, the term *scale* is used in this thesis to refer to the dimensions of that instrument as separate constructs or elements of a construct. The terms *measurement instrument, instrument, and measure* also are used interchangeably.

evaluate the effectiveness of the THP program on these outcomes is presented in Chapter Six by reference to the relevance of the scales to the THP program (for details on how scale relevance was assessed, see Chapter Six).

Hope

Theoretical Constructs

Snyder's Hope Theory. Snyder is considered the primary scholar on hope as a psychological construct (Boyatzis, Boyatzis, & Akrivou, 2006). Snyder's hope theory is based on hope as an "overall perception that one's goals can be met" (Snyder et al., 1997, p. 400). This perception is driven by two interrelated, but distinct, components of goal-related cognitive processes: agency and pathways thinking. Pathways thinking refers to the belief in one's abilities to find effective pathways to achieving one's goals. Agency is about the belief in one's ability to commence action and persevere in the pursuit of those goals. These two elements are said to work together to sustain goal engagement (Rose & Sieben, 2018). In this theory, higher levels of both agency and pathways thinking are required for higher levels of hope (Snyder, 2002; Snyder et al., 1997).

Individuals high in agency are self-motivated and are able to persist in the pursuit of their goals even when difficulties arise (Snyder et al., 1997). Those high in pathways thinking are able to quickly and confidently perceive realistic routes, including alternative routes, to achieve their goals (Snyder et al., 1997). Being able to plan goals and take action to achieve them, generates a motivational force that drives goal striving and differentiates hope from other constructs such as optimism (Ciarrochi, Heaven, & Davies, 2007; Snyder, Rand, & Sigmon, 2002). It has been suggested that pathways thinking is a concept not found in any other formulations of optimism (Peterson, 2000), making it a particularly valuable component in this area of study. Importantly, one's level of hope can be developed. Hope interventions across a range of studies have shown an enhancement of hope levels with an average effect size of .39 (Dixon, Keltner, Worrell, & Mello, 2017; cf., Weis & Speridakos, 2011).

Hope is central to theories of wellbeing and effectiveness, and therefore, is posited to have beneficial effects on one's ability to flourish. Several studies have

shown higher amounts of hope to be associated with various valuable outcomes, including higher academic achievement (Ciarrochi et al., 2007; Dixson, 2017; Dixson et al., 2017; Feldman & Kubota, 2015; Snyder et al., 1997; Snyder & Shorey, 2002), enhanced problem-solving abilities (Chang, 1998; Snyder et al., 1991), reduced use of disengagement coping strategies when dealing with stress (Chang, 1998; cf., Snyder, 1995), higher graduation rates (Worrell & Hale, 2001), and the positive promotion of life satisfaction and general wellbeing (Chang, 1998; Snyder, 1995). With specific reference to young people, Snyder, Shorey and Rand argue,

having hope means that students have well-defined goals, a belief in their ability to develop strategies for reaching those goals, and the requisite motivation to use those strategies. ... Believing that they inevitably will succeed, high-hope students are not side-tracked by goal-blocking thoughts of failure. (2006, p. 170)

Accordingly, having a higher level of hope can equip a student for success.

Adolescents who experience adversity may have lower expectations for the future, and this may reduce their motivation, goal setting, and ability to achieve, thereby making them even more vulnerable. While hope is seen to be particularly important for these at-risk students (Dixson et al., 2017; Snyder et al., 2006), very few studies have examined hope in the context of this group (Dixson, 2017; Dixson et al., 2017).

A number of hope enhancement strategies have been used in both clinical and non-clinical settings. Weis and Speridakos (2011) used meta-analysis to assess research on the effectiveness of interventions intended to enhance hopeful cognitions. The meta-analysis included 19 studies that used one of Snyder's measures of hope and seven studies using an alternative measure of hope. Overall, the research found these hope enhancement strategies to have a significant, but small, effect in increasing self-reported hopefulness ($d = .22$) and life satisfaction ($d = .16$) (Weis & Speridakos, 2011). The authors suggested that hope might be an outcome of goal attainment, rather than a determinant of it (Weis & Speridakos, 2011). However, given the small number of studies included in the meta-analysis and substantial differences among those studies (e.g., research design, outcome measure, intervention delivery, participant age, and setting), it is suggested that further research is required.

The development of agency and pathways thinking are an explicit focus of the THP program. During the coaching sessions, coaches outline the basic principles of the theory and provide examples of agency and pathways thinking. The coaches discuss values with participants in order to help them set personally meaningful goals. Coaching sessions are used also to identify multiple pathways toward goal achievement, anticipate obstacles to goal attainment, and work through any setbacks. A focus on strengths is also intended to develop agency. The experiential learning cycle allows for reflection during the goal pursuit process and also provides an opportunity for enhancing hope through post-goal-attainment reflection. All of these strategies are intended to build hope in the participants.

Scheier and Carver's Optimism. Another construct similar to hope, is optimism. While that term has been defined in different ways, the construct considered here is the one developed by Scheier and Carver (1985), which focuses on the generalised expectancy that one will experience positive (versus negative) outcomes in the future. People high in optimism expect positive results when they strive to achieve their goals, and they expect to successfully manage any problems they may encounter along the way (Scheier & Carver, 1985). As with hope theory, Scheier and Carver's theory of optimism is based on optimism as primarily a cognitive process and is intricately linked to the goal striving process (Peterson, 2000). However, it differs from hope theory in that it does not have regard to the means by which future outcomes occur (Carver & Scheier, 2014).

Individual differences in levels of optimism versus pessimism have been found to have implications for one's self-regulated behaviour, including in relation to motivation, goal-striving, coping, and resilience (Carver & Scheier, 2014; Carver, Scheier, & Segerstrom, 2010; Scheier, Carver, & Bridges, 1994). Compared to those who are more pessimistic, optimists have been shown to experience less distress when faced with adversity, to cope more effectively in stressful situations, and to feel confident and persist even when goal pursuit is challenging or progress is slow (Carver et al., 2010). Specifically with regard to adolescents, empirical evidence indicates that optimism can be important in establishing these adaptive behaviours during what can be a difficult transitional phase (Monzani, Steca, & Greco, 2014).

There is some debate in the literature as to whether optimism and pessimism represent two ends of a single spectrum or two distinct dimensions (Alessandri et al., 2010; Carver & Scheier, 2014; Monzani et al., 2014; Rauch, Schweizer, & Moosbrugger, 2007; Vautier, Raufaste, & Cariou, 2003; cf., Herzberg, Glaesmer, & Hoyer, 2006). Carver et al. (2010) argue that optimism should be seen as a bipolar dimension, however, they acknowledge that studies to date have had varying results on this question, so the issue remains open (Carver & Scheier, 2014). For more detail on this issue, refer to Chapter Five.

While optimism is seen as being relatively stable over time, experiences can influence the extent to which one is optimistic or pessimistic (Peterson, 2000). It is, therefore, a quality that appears to be amenable to change (Carver et al., 2010; Gillham, Reivich, & Shatté, 2001). However, it is important to bear in mind that interventions aimed at reducing pessimism may not result necessarily in an improvement in optimism, consistent with the notion that optimism and pessimism may be distinct dimensions (Carver et al., 2010). Further, while one's level of optimism may benefit from interventions aimed specifically at increasing optimism or decreasing pessimism, it is also possible that interventions that focus on related areas, for example, stress management or goal-setting skills, may equally result in improvements in one's optimism (Carver et al., 2010). Nonetheless, questions still remain as to whether such changes will be generalised and long-lasting (Carver & Scheier, 2014). Within the THP program, it is anticipated that the development of goal-setting skills and the experience of the successful completion of a variety of challenges, will lead to increases in optimism. It is also anticipated that there will be a corresponding reduction in pessimism. However, there is no express focus in the THP program on either developing greater optimism or decreasing pessimism.

Self-Efficacy. Another expectancy construct, *self-efficacy*, also overlaps with hope theory, particularly the agency component. Self-efficacy refers to a person's beliefs in his or her personal capabilities to take action to manage future situations, even difficult ones (Bandura, 1997b). It has been described as a judgement of the confidence one has in one's abilities (Pajares & Schunk, 2001). These self-efficacy expectancies, being concerned with the general performance of behaviours, are

broader than hope expectancies, which relate more specifically to the attainment of one's goals (Snyder, 2002). Self-efficacy beliefs are said to be important because people with high self-efficacy have higher aspirations, are more motivated and persistent, and are resilient in the face of setbacks (Bandura, 1997a).

Bandura (1997a, 1997b) describes four sources of self-efficacy beliefs: (a) experiences of success or mastery in challenging tasks; (b) experiences of social modelling of success; (c) social persuasion to believe in oneself, including direct feedback; and (d) building physical strength, which signals personal capability and reduces stress, anxiety, and depression. The sense of achievement derived from OAE experiences has been found to play a key role in increased self-efficacy (Bandura, 1997b; Hattie et al., 1997; Rose et al., 2018). The THP program provides possibilities for experiencing all four of these sources of self-efficacy.

Magaletta and Oliver (1999) examined the relations among hope, optimism, self-efficacy, and wellbeing in a study involving 204 undergraduate psychology students. The results of their investigation found that hope, optimism, and self-efficacy are related, but not identical, constructs (see also Feldman & Kubota, 2015; Vacek, Coyle, & Vera, 2010). In that study, each of pathways, agency, self-efficacy, and optimism showed up as distinct factors in a maximum-likelihood factor analysis, and multiple regression analysis indicated that hope, self-efficacy, and optimism each made a significant and unique contribution to the prediction of wellbeing (Magaletta & Oliver, 1999).

Hope, optimism, and self-efficacy may be particularly important sources of strength for at-risk adolescents, who face challenges beyond the classroom and the usual teenage angst, navigating additional financial obstacles and other home life difficulties. For these reasons, this research measures all three constructs. Hope is measured using the Children's Hope Scale, optimism is measured using the revised Life Orientation Test, and self-efficacy is measured as a component of life effectiveness with the Review of Personal Effectiveness questionnaire (see below for more on measuring life effectiveness skills).

Measuring Hope

Children's Hope Scale (CHS). Snyder et al. (1991) operationalised Snyder's construct of hope as the Hope Scale. This was followed by a separate measurement instrument specifically for children called the Children's Hope Scale (CHS; Snyder et al., 1997). This measure presumes that children are goal-directed and that their cognitive processes in connection with goal pursuit can be understood according to two elements: agency and pathways thinking (Lopez, Ciarlelli, Coffman, Stone, & Wyatt, 2000). Given the strong focus of the THP program on building agency and pathways thinking, both dimensions of this instrument were determined to be highly relevant to the THP program design and aims.

Life Orientation Test, Revised (LOT-R). The Life Orientation Test was designed to measure dispositional optimism through two dimensions: optimism and pessimism. The original Life Orientation Test was developed by Scheier and Carver in 1985, and much of the research on optimism and pessimism since then has made use of this test (Scheier et al., 1994). The Life Orientation Test was revised in 1994 following concern that two of the items measured coping, a mediator of optimism, rather than optimism itself (LOT-R; Scheier et al., 1994). The dimensions of optimism and pessimism were determined to be of moderate and limited relevance, respectively, to the THP program design and aims.

Self-Regulation

Theoretical Construct

Self-regulation has been described as one's capacity for altering responses to bring them into line with standards such as ideals, values, morals, and social expectations, as well as to support goal pursuit (Baumeister, Vohs, & Tice, 2007, p. 351). Self-regulation is associated with positive life outcomes in many spheres from wellbeing to academic achievement (Baumeister et al., 2007; Murray, Rosanbalm, Christopoulos, & Hamoudi, 2015; Tangney, Baumeister, & Boone, 2004), making it a worthy point of focus for helping students to reach their full potential. Such a focus is important particularly during adolescence when higher emotional arousal can impact self-regulatory development (Murray et al., 2015). Adolescents who fail to develop self-regulation skills are more likely to engage in risk-taking and

unhealthy behaviours (Tangney et al., 2004). Moreover, research has demonstrated a positive connection between extracurricular programs and increased self-regulation, particularly among disadvantaged youth (Bandy & Moore, 2011).

There are a wide range of concepts used to describe, or that overlap with, the construct of self-regulation, including self-control, self-discipline, and self-efficacy. For this research, self-regulation is defined as “the act of managing cognition and emotion to enable goal-directed actions, such as organising behaviour, controlling impulses, and solving problems constructively” (Murray et al., 2015, p. 5). This definition was developed specifically to facilitate intervention approaches and is intentionally broad and applied (Murray et al., 2015). It emphasises self-regulation as a purposeful process originating within the individual (Carver & Scheier, 2011). This definition also highlights the important role that self-regulation plays in goal setting and pursuit, a core focus of the THP program.

Self-regulation is considered a multifaceted process involving cognitions, emotions, and behaviour (Moilanen, 2007). The various facets of self-regulation interact and relate in complex ways, making it difficult to separate them empirically when measuring self-regulation (Murray et al., 2015). Furthermore, the ability to self-regulate is impacted by factors internal and external to a person, including that person’s genetics, self-regulatory skillset, motivation (both intrinsic and extrinsic), social support, and environmental context (Murray et al., 2015). Self-regulatory skills are wide-ranging, including self-awareness and self-reflection, planning, monitoring, inhibition of negative emotions, delay of gratification, and flexibility to adapt behaviour and emotions as needed (Moilanen, 2007). Unlike traits that may be more stable over time, self-regulation skills are developed over an extended period from birth through young adulthood and beyond (Murray et al., 2015). Adolescence generally presents a period of marked self-regulatory development under normal circumstances, however, this development can be disrupted or slowed, especially in circumstances of persistent adversity (Hamoudi, Murray, Sorensen, & Fontaine, 2015; Murray et al., 2015). Encouragingly, self-regulation skills can be fostered through instruction and support, and these skills appear to be responsive to interventions like the THP program, which provides supportive learning environments and opportunities for practice (Murray, Rosanbalm, &

Christopoulos, 2016; Murray et al., 2015). Meta-analytic research into OAE has found self-regulation to be a major theme underlying the most significant participant outcomes (Hattie et al., 1997).

Accordingly, it has been suggested that interventions for adolescents should focus more intentionally on self-regulation skills and incorporate adults who can provide support in the learning process (Murray et al., 2016). The THP program has an explicit focus on developing self-regulation skills specific to goal pursuit, with participants learning to set appropriate goals and monitor and evaluate their goal progress through David A. Kolb's (1984, 2014) experiential learning framework. Participants also work on developing an awareness of, and taking ownership for, their thoughts, feelings, and behaviours. Both the coaches and teachers assigned to work with a program cohort, also play an important role in scaffolding this developmental process (see Chapter Four for more information on these roles).

Measuring Self-Regulation: Adolescent Self-Regulatory Inventory

The Adolescent Self-Regulatory Inventory⁵ (ASRI; Moilanen, 2007) was developed to measure the degree to which adolescents are able to use self-regulatory processes in both the short and long term. The short-term self-regulation items are said to be related to impulse, attentional, or emotional control in the "heat of the moment" or immediate context, while the long-term items relate to impulse control or direction of effort over a longer period of time, lasting several weeks, months, or years (Moilanen, 2007). In addition to this temporal context and the facets of self-regulation described above (emotional, attentional, cognitive, and behavioural), self-regulation also can be broken down by reference to its various processes: monitoring, persevering, activating, adapting, and inhibiting (Moilanen, 2007). Those items that aim to tap into goal-related self-regulation were determined to be the most strongly relevant to the THP program design and aims, while items assessing emotional self-regulation were found to be of more moderate relevance and items focused on attentional control were found to have the least relevance to the THP program design and aims.

⁵ A later version of this instrument of the same name comprises 52 items (24 short-term self-regulation items and 28 long-term self-regulation items (Moilanen, 2014).

Resilience

Theoretical Constructs

Resilience. Resilience is considered a critical component within the general context of wellbeing (Cowen, 1994) and is seen to be an element relevant to healthy youth development (Ewert & Yoshino, 2011). Accordingly, it is also an important outcome for consideration. Resilience refers to the capacity one has to adapt to adversity (Masten, 2014). Although self-efficacy is related to resilience, self-efficacy is a broader concept that applies beyond circumstances of adversity. Nevertheless, self-efficacy beliefs are said to promote resilience in such circumstances (Bandura, 1997a). Resilience is important because it is believed to assist individuals to deal more effectively with everyday challenges and to solve problems (Goldstein & Brooks, 2013). In addition, resilience is an outcome variable that has been found to be positively influenced by OAE and coaching interventions (e.g., Beightol, Jeverson, Gray, Carter, & Gass, 2009; Bowen, Neill, & Crisp, 2016; Clough & Strycharczyk, 2015; Ewert & Yoshino, 2011; Grant et al., 2009; Hayhurst et al., 2015; Kelly, 2019; Neill & Dias, 2001; Ungar, Dumond, & McDonald, 2005; Whittington & Aspelmeier, 2018).

Within the behavioural and social sciences, resilience research has aimed to uncover those factors both within an individual and externally occurring in that individual's environment, that have a positive impact on one's resilience (Howard & Johnson, 2000). Some of these factors that have been found to protect against or reduce vulnerability include having supportive and caring relationships, a sense of achievement, competence or success, cognitive and self-regulation skills, motivation, and a sense of purpose and future (Howard & Johnson, 2000; Masten, 2001; Reivich, Gillham, Chaplin, & Seligman, 2013). Much of the research related to childhood resilience has focused on these variables in at-risk children who have experienced extreme adversity, as it was originally believed that resilience processes were only implemented in the context of such adversity (Masten, 2001; Rutter, 2006). More recently, however, resilience research has begun to focus not only on at-risk youth, but adolescents more generally, reflecting on resilience from the

everyday stress and challenge of the high pressure environment in which many people live (Goldstein & Brooks, 2013; Martin & Marsh, 2008, 2009).

As with self-regulation, it has been suggested that resilience ultimately emerges from ordinary developmental processes, indicating that interventions should focus on strategies that nurture the development of, protect, or restore these basic systems (Masten, 2001). In particular, there is said to be a role for challenging experiences in such interventions, provided they offer an opportunity to successfully cope with the challenge or stress afforded by such experiences (Masten, 2001; Rutter, 2006). Outdoor adventure programs that maintain an appropriate balance between challenge and support, providing opportunities to fail and recover, can assist in strengthening resilience (Kelly, 2019). The THP program provides challenges for the individual and the group that present such opportunities in a safe and supported environment and, therefore, should help to build resilience.

In line with a more encompassing approach to resilience, academic resilience (also referred to as *academic buoyancy*⁶) is about a student's ability to sustain school-related achievement motivation and performance despite the presence of challenging events or setbacks that are typical of everyday school life (Buck & Inman, 1998; Martin & Marsh, 2006, 2008, 2009). This type of resilience is said to be relevant to all students as all students are likely to face some type of performance setback or other adversity, challenge, or pressure during their school years (Martin & Marsh, 2006, 2008, 2009). This particular conceptualisation of academic resilience is consistent with the positively-oriented construct of flourishing, and proposes that academic resilience can be enhanced through the development of positive cognitive, affective, and behavioural approaches to academic life (Martin & Marsh, 2008, 2009). Research has identified five motivational predictors of academic resilience, referred to as the 5Cs of academic buoyancy: confidence (self-efficacy), coordination (planning), commitment (persistence or grit), composure

⁶ Martin and Marsh (2008, 2009) use the terms academic resilience and academic buoyancy as two distinct concepts in a hierarchical structure, differentiated in part by the level of adversity. These terms are used here interchangeably but focused on the day-to-day proactive frontline response to more minor academic adversity, rather than the defensive backline response that might be required in situations of more severe or chronic adversity (see also Martin, 2013).

(low anxiety), and control (agency) (Martin, Colmar, Davey, & Marsh, 2010; Martin & Marsh, 2006). It has been further suggested that self-regulation also may have a role to play (Martin, Colmar, et al., 2010; Murray et al., 2015). Participants in the THP program have the opportunity to work with their coaches through the program challenges to develop in all of these areas. Even where there is no specific focus on school-related resilience, it is hypothesised that building these general skills will have a positive effect on academic resilience.

Grit. An attribute related to resilience, is grit. Described as the “perseverance and passion for long-term goals,” gritty individuals persist in the pursuit of challenging goals, maintaining effort and interest in their goals despite setbacks, boredom, or lack of any positive feedback (Duckworth, Peterson, Matthews, & Kelly, 2007, p. 1087). Grit is said to be more than resilience; it also involves deep commitments to which one remains loyal (Perkins-Gough, 2013).

The emphasis of grit theory is on trait-level, long-term stamina (Duckworth et al., 2007; Duckworth & Quinn, 2009). While grit is considered to be a relatively stable trait, Duckworth considers that grit can be developed by changing beliefs (Perkins-Gough, 2013). Of particular relevance to this idea is the concept of having a *growth mindset* versus a *fixed mindset* (Dweck, 2006). A growth mindset is the belief that intelligence is not fixed and can be developed (Dweck, 2006). It is proposed that students who have a growth mindset generally see difficult tasks as a learning opportunity, seek out challenging learning experiences, and persist as a result (Blackwell, Trzesniewski, & Dweck, 2007). Research has shown that targeted interventions can help students to develop a growth mindset (Blackwell et al., 2007; Good, Aronson, & Inzlicht, 2003; Paunesku et al., 2015). A recent study found that having a growth mindset reliably predicted achievement in a national sample of high school students of all socioeconomic levels across Chile (Claro, Paunesku, & Dweck, 2016). In addition, while students from lower-socioeconomic families were less likely to hold a growth mindset when compared to their wealthier peers, those who did hold a growth mindset were buffered to some extent against the detrimental effects of economic disadvantage on achievement (Claro et al., 2016). Duckworth has suggested that one thing that makes someone grittier is having a

growth mindset (Perkins-Gough, 2013). Building on the work of Dweck, grit interventions focus on the idea of *deliberate practice*, which is about engaging in very effortful practice on things you can't yet do (Perkins-Gough, 2013).

Grit theorists argue that individual differences in grit can explain why two people with the same ability in a domain may have different performance outcomes in that domain (Duckworth, Kirby, Tsukayama, Berstein, & Ericsson, 2011; Duckworth et al., 2007; Duckworth & Quinn, 2009). There are others, however, who question the strength of the relation between grit and success (Credé, Tynan, & Harms, 2016). Grit, some argue, may not necessarily be adaptive without general potential to succeed in a domain, as well as the ability to engage in the self-reflection and self-monitoring important to self-regulated learning and performance (Credé et al., 2016). There is also the suggestion that in certain circumstances, too much grit may impede performance, for example, where it reduces opportunities for valuable help-seeking behaviour or results in ineffective persistence of a particular goal.

The goal setting-process in the THP program, together with the development of skills in self-reflection and self-regulation, are relevant to building grit. In addition, learning in the THP program is expressly focused on developing a growth mindset and the benefits of deliberate practice. It is hypothesised that both the challenges in the program and the opportunity to participate in a community project (for more information on this project, refer to Chapter Four) afford the participants opportunities to persevere in the pursuit of their goals in a way that will help to develop adaptive grit.

Measuring Resilience

Academic Resilience Scale (ARS). Martin and Marsh (2006) developed a unidimensional measure of self-reported academic resilience. The Academic Resilience Scale assesses a student's ability to deal with setbacks, pressures, and other everyday challenges, in the school setting (ARS; Martin & Marsh, 2006). Given its express focus on schoolwork and marks, this scale was determined to be of limited relevance to the THP program design and aims.

Life Resilience Scale (LRS). Based on the ARS, a unidimensional measure of self-reported general resilience or buoyancy in life was developed for this research. References in the ARS items to exams, school, schoolwork, and study were replaced with references to life in general. As the THP program focus is on resilience at the level of goal-striving, the breadth of the LRS items to life in general was found to make this scale less relevant to the THP program design and aims.

Short Grit Scale (Grit-S). Duckworth et al. (2007) developed a 12-item self-report measure of grit, which was subsequently revised to an eight-item instrument (Grit-S; Duckworth & Quinn, 2009). Grit is measured through two dimensions: consistency of interest and perseverance of effort. Although the concept of growth mindset may be used with program participants, the focus of these items on passion and persistence were considered to be less directly relevant to the THP program design and aims.

Motivation and Engagement

Theoretical Construct

Motivation broadly refers to a person's drive to pursue something and persist in that pursuit (Pintrich, 2003). Motivation and engagement, in an academic sense, have been described as students' "energy and drive to engage, learn, work effectively, and achieve to their potential at school and the behaviours that follow from the energy and drive" (Martin, 2005, p. 180). Research has demonstrated a positive association between motivation- and engagement-related constructs and a number of important outcomes, including educational achievement, educational aspirations, class participation, and school enjoyment (see Martin, 2007; Schunk, 2014). Students' motivation and engagement can impact their success in school. Importantly, research also indicates that students' levels of motivation and engagement decline after middle school (Martin, 2001, 2002, 2003, 2007), so these areas provide a worthy point of focus for students in their early high school years.

The theoretical field of motivation and engagement science is extensive, and there are numerous measures of student motivation. However, most of these instruments reflect a view of motivation that is grounded in a single theoretical perspective (Martin, 2001, 2003). Martin (2007) suggested several concepts that he

considered critical to students' motivation, including self-efficacy, perceived control of outcomes, orientation to tasks (mastery or performance), self-regulatory capacity, and outcome orientation (success, failure-avoidance, or failure-acceptance). Drawing together a number of theoretical perspectives on motivation and engagement, Martin (2001, 2002) developed the *Motivation and Engagement Wheel* (formerly the *Student Motivation Wheel*), which comprises constructs central to these theories with a focus both on enhanced motivation and reduced motivation. Martin (2001, 2002, 2003, 2007) proposed that these areas provide a variety of focal points for interventions aimed at increasing students' motivation and engagement. For example, teaching students effective goal-setting strategies can lead to goal attainment, which in turn can enhance one's self-efficacy (Martin, 2007). Moreover, goal theory provides a pathway for creating a mastery orientation, as well as improving a range of self-regulatory skills, and improvements in these areas can have a flow-on effect for persistence (Martin, 2007). Furthermore, teaching effective goal pursuit strategies that focus on mastery over performance, connect effort to outcomes, and make room for mistakes, can help to address some of the impeding and maladaptive thoughts and behaviours that undermine students' motivation and engagement (Martin, 2007). Such interventions are aligned with the philosophy and programming underlying the THP program. Research has demonstrated that interventions with high school students focusing on these areas result in significant gains in motivation in both the short term and at follow up (Martin, 2005). Additionally, co-curricular or extracurricular programs can enhance academic engagement and achievement, so these interventions need not be restricted to the classroom (Martin, 2005; Valentine, Cooper, Bettencourt, & Dubois, 2002).

Measuring Motivation and Engagement: Motivation and Engagement Scale, Short (MES-S)

Based on his Motivation and Engagement Wheel, Martin (2007, 2009) developed a 44-item, multidimensional instrument for measuring high school students' (12-18 years) academic motivation and engagement, called the Motivation and Engagement Scale–High School (MES-HS, formerly called the Student

Motivation and Engagement Scale). This measure consists of 11 scales, with four items for each scale. These scales are further grouped into higher-order structures reflecting various dimensions of motivation and engagement: three scales assess adaptive cognitive dimensions (*Booster Thoughts*), three assess adaptive behavioural dimensions (*Booster Behaviours*), three assess impeding or maladaptive cognitive dimensions (*Mufflers*), and two assess maladaptive behavioural dimensions (*Guzzlers*). *Booster Thoughts* include scales that measure self-belief (or self-efficacy), learning focus (or mastery orientation), and valuing school. *Booster Behaviours* include scales that assess persistence, planning, and task management. *Mufflers* consist of items that are negatively worded to assess anxiety, failure avoidance, and uncertain control. *Guzzlers*, also comprised of negatively-worded items, measure self-sabotage (or self-handicapping) and disengagement. The items for each scale are aggregated to determine a score for each scale, and the 11 individual scores can then be converted to four global scores that reflect the average of *Booster Thoughts*, *Booster Behaviours*, *Mufflers*, and *Guzzlers*.

With a large number of scales being incorporated into the primary measurement instrument for this research, it was decided to use only a single item representative of each of the 11 scales from the MES-HS. It was hypothesised that these 11 items would form four scales reflecting the four higher-order factors of *Booster Thoughts*, *Booster Behaviours*, *Mufflers*, and *Guzzlers*. This revised measure is referred to in this thesis as the Motivation and Engagement Scale–Short (MES-S). As these items are directed expressly at schoolwork, homework, and tests they were only considered to be of moderate (*Booster Behaviours*, *Booster Thoughts*) and limited (*Guzzlers*, *Mufflers*) relevance to the THP program design and aims.

Wellbeing

Theoretical Construct

The importance of wellbeing to people globally and in all aspects of life means that there is a wide range of interest in this construct across disciplines and throughout both the private and public sector. This has led to a proliferation of different theories, operationalisations, and constructs of wellbeing. Most of these

conceptualisations of wellbeing are influenced by one of two contrasting philosophical perspectives: *hedonia* and *eudaimonia*. Hedonia focuses on the subjective experience of happiness and feeling satisfied with life, while eudaimonia focuses on psychological functioning and living well. In contrast to philosophy, where hedonia and eudaimonia are seen as competing theories, psychologists have begun to consider the potential compatibility of these concepts (Huta & Waterman, 2014). Scholars of this approach argue that the pursuit of both hedonia and eudaimonia provides the greatest and most well-rounded wellbeing (Huppert & So, 2009; Huta, 2013; Huta & Ryan, 2010; Kashdan, Biswas-Diener, & King, 2008; Seligman, 2002, 2011).

There is a substantial amount of research that has found hedonia and eudaimonia to be two constructs that are highly related, but also distinct components of overall wellbeing (Delle Fave, 2009; Keyes & Annas, 2009; Keyes, Shmotkin, & Ryff, 2002; Linley, Maltby, Wood, Osborne, & Hurling, 2009; Vitterso & Soholt, 2011; Waterman, 1993; Waterman, Schwartz, & Conti, 2008). Research also has been carried out to better understand what ordinary people refer to when they speak of wellbeing and happiness (Delle Fave, Brdar, Freire, Vella-Brodrick, & Wissing, 2011). This research found that people refer to both hedonic and eudaimonic aspects of wellbeing, suggesting that wellbeing is a multifaceted concept.

Additionally, there have been a number of studies (across a variety of people and nations) which have considered both hedonia's and eudaimonia's respective contributions to wellbeing or life satisfaction (Chan, 2009; Huta, 2013; Huta & Ryan, 2010; Keyes & Annas, 2009; Park, Peterson, & Ruch, 2009; Peterson, Park, & Seligman, 2005; Schueller & Seligman, 2010; Steger, Kashdan, & Oishi, 2008; Vella-Brodrick et al., 2009). In each of these studies, people who were high on both hedonia and eudaimonia had the highest wellbeing, life satisfaction, and positive affect. In other research, Schueller and Seligman (2010) noted that this effect was more than an additive combination of the impact each pathway generated on its own. Therefore, pursuing both hedonia and eudaimonia is believed to lead to the greatest and most diverse wellbeing (Henderson, Knight, & Richardson, 2013; Huta & Ryan, 2010).

Waterman (1993) has found eudaimonic activities to be strongly associated with developing one's best potential, investing effort, setting clear goals, feeling assertive, interested, and challenged, and having high concentration. Hedonic activities were associated with feeling relaxed, excited, content and happy, losing track of time, and forgetting one's problems. Vitterso and Soholt (2011) have suggested that eudaimonic experiences are most important for pursuing complex goals and challenging activities, while hedonic emotions are more important for the preservation of stability and rewarding need fulfillment.

Wellbeing has major implications for health and social outcomes and psychological functioning (Stewart-Brown et al., 2009), and adolescence is an important developmental stage that creates a foundation for wellbeing in later life (Lawrence et al., 2015). Accordingly, the promotion of positive mental wellbeing among adolescents is an important overall aim for the THP program and this research.

Measuring Wellbeing

Satisfaction with Life Scale (SWLS). The Satisfaction with Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985) is a well-respected instrument intended to measure the cognitive-evaluative component of subjective wellbeing, as opposed to the affective component which represents emotional wellbeing (Pavot & Diener, 1993). It aims to assess satisfaction with one's life as a whole, rather than in any particular life domain (Pavot & Diener, 1993). The measurement of one's life satisfaction by this instrument, is proposed to be a process in which a person evaluates his or her life using the person's own criteria or standards (Pavot & Diener, 1993, 2009). The breadth of this scale's focus on one's life as a whole was determined to make it less relevant to the THP program design and aims.

Warwick-Edinburgh Mental Well-Being Scale (WEMWBS). The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS; Tennant et al., 2007) aims to capture a wide conception of wellbeing through the measurement of emotional, cognitive-evaluative, and psychological aspects of one's level of mental wellbeing (Tennant et al., 2007). There is also a shorter version of the scale called the Short Warwick-Edinburg Mental Well-Being Scale (SWEMWBS; Stewart-Brown

et al., 2009). Mental wellbeing is said to be one aspect of overall wellbeing, which also includes physical and social aspects of wellbeing (Putz, O'Hara, Taggart, & Stewart-Brown, 2012). The construct of mental wellbeing includes positive affect and life satisfaction, as well as positive psychological functioning, satisfying relationships, and self-realisation/acceptance (Putz et al., 2012; Tennant et al., 2007). However, the SWEMWBS is less broad than the WEMWBS, focusing more on psychological functioning and other aspects of eudaimonic wellbeing, than on emotional or hedonic wellbeing (Stewart-Brown et al., 2009). As wellbeing is an overall aim of the THP program, rather than a direct focus of the program design, this scale was determined to have a moderate relevance to the THP program design and aims.

Self-Concept

Theoretical Construct

Another construct related to self-efficacy is self-concept. Self-concept has been described as an individual's perceptions of the self, which are developed through interactions with the environment in the context of a feedback loop, where perceptions influence behaviour and behaviour in turn influences these perceptions (Shavelson, Hubner, & Stanton, 1976). These self-perceptions are broader than self-efficacy, including feelings of self-confidence, self-worth, self-acceptance, competence, and ability (Marsh, Martin, Yeung, & Craven, 2017). While self-efficacy beliefs are suggested to be future-oriented, self-concept beliefs are said to be largely based on past accomplishments and circumstances (Marsh et al., 2019). It also has been argued that, unlike self-efficacy, self-concept perceptions include both a descriptive component and an evaluative component for which frames of reference or standards of comparison are relevant (Marsh et al., 2019).

While self-concept has been perceived in the past as a broad, unidimensional construct (e.g., Coopersmith, 1967; Marx & Winne, 1978), there is now a preponderance of research that supports a construction of self-concept as multifaceted and hierarchical (Byrne, 1984; Marsh, 1990c; Marsh, Ellis, Parada, Richards, & Heubeck, 2005; Marsh & Gouvernet, 1989; Marsh & Hattie, 1996; Shavelson et al., 1976). General self-concept is broadly divided into two

components: academic self-concept and non-academic self-concept (Shavelson et al., 1976). Academic self-concept is differentiated by subject area, and non-academic self-concept is broken down among various types of physical, emotional, and social self-concepts (Shavelson et al., 1976). At the base of the hierarchy are perceptions of personal behaviour in specific situations, with assessments of the self in broader domains (e.g., social, physical, academic) in the middle, and a general, global concept of self at the top (O'Mara, Marsh, Craven, & Debus, 2006; Shavelson et al., 1976). While Shavelson (1976) proposed that general self-concept was more stable than the specific components of self-concept, research has not borne this out (Marsh, 1990c, 1993; Marsh & Craven, 1997; Marsh, Craven, & Martin, 2006; Marsh & Hattie, 1996; Marsh et al., 1986a, 1986b). Furthermore, within this multidimensional, hierarchical model of self-concept, it is proposed that global self-concept (often referred to as *self-esteem*) is better represented as a component of the multifaceted self-concept structure, rather than a higher-order factor based on specific components of self-concept (Marsh, 2005; Marsh et al., 2006).

Improving self-concept has long been considered a valuable educational outcome, both in its own right and in connection with the contribution that it makes in facilitating other desirable outcomes (Marsh, 2005; Marsh & Yeung, 1997; O'Mara, Marsh, et al., 2006; Shavelson et al., 1976). Research has shown, for example, that self-concept is an essential component for academic success (Bloom, 1976; Valentine et al., 2002; Valentine, Dubois, & Cooper, 2004). A person's perception of their academic ability has wide-ranging implications, including for school performance (Marsh, 1990a, 2005; Marsh & Yeung, 1997), motivation for academic tasks (Bandura, 1986; Byrne, 1984; McInerney, Roche, McInerney, & Marsh, 1997), and subsequent coursework selection (Marsh & Yeung, 1997). In addition to the impact that academic self-concept has on academic achievement, there is said to be a reciprocal effect pursuant to which academic success also brings about increases in academic self-concept (Marsh, 1990a, 2005; Marsh & Martin, 2011; Marsh & Yeung, 1997; cf., Marsh & Yeung, 1998). In a similar way, there also may be a reciprocal effect between non-academic areas of self-concept and one's experiences of achievement in those areas (Marsh, 2005). Beyond its impact on achievement, self-concept has been seen as a measure of personal growth

(Henderson, 2009). Importantly, interventions aimed at improving self-concept have been found to improve self-concept in adolescents (O'Mara, Marsh, et al., 2006).

A meta-analysis of 152 self-concept interventions in school settings confirmed the multidimensional perspective of self-concept and found that targeting specific self-concept domains led to better results in those domains (O'Mara, Green, & Marsh, 2006; see also O'Mara, Marsh, et al., 2006). While there were no specific features of the administration of these interventions that were found to be significant moderators of effect size, interventions in secondary schools were slightly more effective than other settings, face-to-face interventions yielded the highest mean effect size, and standardised interventions also resulted in higher effect sizes (O'Mara, Green, et al., 2006). Outside of the school environment, outdoor adventure programs have been found to be very successful in enhancing multiple dimensions of self-concept (Bowen & Neill, 2013; Capurso & Borsci, 2013; Cason & Gillis, 1994; Hattie et al., 1997; Marsh et al., 1986a, 1986b). However, there is an open question on the role of self-concept in the existing research on outdoor adventure education, and whether it is more appropriately conceived as an outcome variable or as a mediating variable through observed behaviour (Richards, Ellis, & Neill, 2002). Through the outdoor adventure component, the THP program directly targets self-concept in relation to one's general self-concept or self-esteem. Moreover, some of the other self-concept domains may be a focus of the individual or group coaching work depending on the student's or group's needs, including in connection with relationship, emotion, and school/academic domains. Accordingly, a focus on both specific and global aspects of self-concept can be valuable.

Measuring Self-Concept: Self-Description Questionnaire II-Short (SDQII-S)

In order to measure self-concept outcomes across different domains, a multidimensional measure is required. Marsh (1988, 1992b, 1992c) developed a series of instruments (called the Self-Description Questionnaire) to measure the multidimensional model of self-concept proposed by Shavelson et al. (1976): the Self-Description Questionnaire I for pre-adolescent primary school students, the Self-Description Questionnaire II for adolescent high school students, and the Self-

Description Questionnaire III for late adolescents and young adults. A short form of the Self-Description Questionnaire II (SDQII-S; Marsh, 1992b) was designed to measure all of the 11 self-concept domains included in the original instrument (Marsh, Ellis, et al., 2005). These aspects of self-concept fall under three broad areas: non-academic self-concept (seven scales), academic self-concept (three scales), and overall general self-esteem (one scale). The general self-esteem scale was determined to be of high relevance to the THP program design and aims, while self-concept of same sex relationships (given the single sex nature of the groups) and emotional stability were considered to be moderately relevant, with the other scales considered to be of less direct relevance.

Life Effectiveness

Theoretical Construct

Life effectiveness has been described as having competency in the “behaviours, cognitions, and emotions which give cross-situational advantage for ‘surviving and thriving’” in any given situation (Neill, 2008, p. 48). More specifically, life effectiveness reflects on how “actively” and “ably” people handle their goals and challenges (Neill, 2008, p. 49). The construct of life effectiveness has been operationalised through a number of dimensions said to represent generic, practical life skills that are both learnable and enhanceable (Neill, 2008; Neill, Marsh, & Richards, 2003; Richards et al., 2002). Therefore, all of these dimensions aptly find themselves a focus of interventions, including through OAE programs. Moreover, the multifaceted nature of this construct is said to allow for a better understanding of the potential outcomes of intervention programs that aim to assist development in multiple areas (Lane, 2008). Consequently, the various dimensions of life effectiveness are outcome variables that have been used in a number of research studies on outdoor education and other experiential education programs (e.g., Bowen & Neill, 2015; Johnson, 2012; Lane, 2008; Neill, 2008; Sibthorp & Arthur-Banning, 2004; Stenger, 2001).

As conceptualised, life effectiveness is said to be related to, but distinct from, other constructs of interest, such as self-concept, self-regulation, and resilience (Neill, 2008). One dimension of life effectiveness is self-efficacy, and this construct

was described in detail above (see under “Hope”, “Theoretical Constructs”). A related construct is self-confidence, which seems to have a less consistent presence in the academic literature (Oney & Oksuzoglu-Guven, 2015; cf., Stankov et al., 2013). Self-confidence has been operationalised variously as self-esteem, self-efficacy, self-concept, optimism, and positive self-beliefs (Marsh et al., 2017). Moreover, while increased self-confidence is one of the most widely recognised aims of OAE, it is much less often measured as an outcome variable (Neill, 2008). The self-confidence measure that forms part of life effectiveness follows a general understanding of self-confidence as reflecting the positive self-beliefs one has about being able to achieve success in the future (as opposed to more domain- or task-specific conceptions). Self-confidence beliefs are said to derive through the effect of past experience but can be updated by new experience (Oney & Oksuzoglu-Guven, 2015). While self-confidence focuses on perceived outcomes, self-efficacy focuses on the perceived competencies necessary to achieve those outcomes (Bandura, 1986; Oney & Oksuzoglu-Guven, 2015). When compared to general self-esteem, although both concepts include an evaluative component, self-confidence is argued to be a narrower concept (Oney & Oksuzoglu-Guven, 2015).

Another related self-perception concept is *locus of control*. Locus of control as a theory was developed out of Rotter’s (1954, 1966) social learning theory. This theory relates to how a person attributes their outcomes, whether positive or negative (Rotter, 1966). Those with an internal locus of control attribute success and failure mainly to factors within their own control. Although a person might believe that they have control over an outcome, this is distinct from self-efficacy beliefs, which focus on the perceived potential inherent within the person to achieve something.

One’s locus of control is a matter of degree and runs along a continuum, with high external locus of control at one end of the continuum and high internal locus of control at the other end. People with a high internal locus of control perceive reinforcement, reward, or other outcomes of their behaviour as following from, or contingent upon, their own behaviour or personal attributes (Rotter, 1966). They see their future as being in their own hands and believe that they are responsible for their successes and failures. People with a high external locus of

control perceive such reinforcement, reward, or other outcomes as being controlled by forces outside of themselves and independent of their own actions (Rotter, 1966). They believe that they have little or no control over their world, with things like chance, luck, fate, or powerful others playing a greater role (Rotter, 1966).

Locus of control is said to be an important factor in determining a person's level of motivation and engagement at school (Ekstrom, Goertz, Pollack, & Rock, 1986). Adolescence is a stressful and challenging period of life in which decisions and their implementation often are perceived as threatening life events that are beyond one's control (Okun, 1984). Such external control expectancies have been associated with a lower sense of wellbeing, increased instances of depression and anxiety, poor coping strategies, lower school achievement, and increased juvenile delinquency (Twenge, Zhang, & Im, 2004). Research has found these effects to be even more pronounced in circumstances of economic disadvantage (Tesiny, Lefkowitz, & Gordon, 1980).

Locus of control is one of the most widely studied individual differences in psychology (Twenge et al., 2004), and it is a common outcome variable in OAE research (e.g., Cason & Gillis, 1994; Hans, 2000; Hattie et al., 1997). Outdoor adventure programs can instil in the participants a sense of empowerment (Walsh & Golins, 1976), and such a feeling of empowerment may have a direct effect on one's locus of control. A number of meta-analyses have demonstrated a significant improvement in internal locus of control through outdoor adventure education programming (Hans, 2000; Hattie et al., 1997). Promoting internal locus of control through these programs can buffer against the maladaptive effects of external control expectancies during adolescence.

Some of the other dimensions of life effectiveness are concerned with self-perceptions of the extent to which one is actually able to perform particular life skills, including social and organisational skills (Neill, 2008). The THP program provides opportunities for teamwork and leadership. There is also an explicit focus on using supported challenge to develop participants' self-confidence, self-efficacy, and perspective-taking capacity, as well as important coping strategies. Therefore, life effectiveness skills are a relevant area of attention.

Measuring Life Effectiveness: Review of Personal Effectiveness with Locus of Control (ROPELOC)

The Review of Personal Effectiveness with Locus of Control (ROPELOC; Richards et al., 2002) was designed to measure self-perceptions of a number of psychological and behavioural dimensions of the construct of life effectiveness. The instrument was developed over 17 years of research into the effects of a wide variety of experience-based, personal development programs, and is based on the Life Effectiveness Questionnaire (LEQ; Neill et al., 2003). It is intended that the items that make up the instrument have a grounding in self-concept, while being expressed and interpreted in terms of behaviours (Richards et al., 2002). It is also envisaged that in addition to its measurement qualities, this instrument can facilitate the processes of self-examination, goal-setting, and feedback that are an inherent part of the experiential learning process (Neill, 2008; Neill et al., 2003).

The dimensions of life effectiveness measured by the ROPELOC are grouped into a number of categories: personal abilities and beliefs (4 scales), social abilities (3 scales), organisational skills (3 scales), energy (1 scale), and overall life effectiveness (1 scale) (Richards et al., 2002). There are also two scales that measure locus of control. Although not behaviour- or skill-based, the locus of control items are said to provide a strong indication of likely behaviour (Neill, Richards, & Badenoch, 1997).

The LEQ and ROPELOC have been used in several studies that have considered the outcomes of outdoor adventure education and experiential learning programs (Culhane, 2004; Imholt, 2009; Johnson, 2012; Lane, 2008; Merrell, 2009; Powers, 2004). While many of the studies using the ROPELOC did not demonstrate a significant effect in life effectiveness using a composite score, several studies have found significant change for some of the scales, including quality seeking (Culhane, 2004), cooperative teamwork, coping with change, and external locus of control (Merrell, 2009). The scales measuring self-confidence, self-efficacy, open thinking, cooperative teamwork, and active involvement were determined to be highly relevant to the THP program design and aims, with stress management, social effectiveness, leadership ability, coping with change, and internal locus of control

considered to be moderately relevant, and the other scales considered less directly relevant.

Summary

Constructs provide a pathway between theory and empirical research (Bergman, 2010). Consequently, it is an essential primary step in the research process to develop an understanding both of the theoretical basis of the constructs of interest, as well as the instruments designed for their empirical measurement. This chapter has outlined the theoretical foundations for the constructs used in the research for this thesis and provided a summary of the instruments selected to empirically measure those constructs, as well as their relevance to the THP program design and aims. These constructs and related instruments underpin the methodology used to address the overarching research aims, hypotheses, and questions for this thesis. The broad methodology and procedures for the research are discussed in detail in the next chapter.

CHAPTER FOUR

GENERAL METHODOLOGY AND PROCEDURES

Introduction

In order to address the aims of this thesis, three studies have been undertaken: Study 1 examines the psychometric properties of the measurement instruments used in Study 2; Study 2 evaluates the effectiveness of The Helmsman Project (THP) program, using a randomised controlled trial (RCT) with waitlist controls and an extended baseline design; and Study 3 assesses semi-structured interviews to obtain qualitatively rich data on how the participants made meaning of their experiences in a THP program.

The purpose of this chapter is to describe the overarching methodology used in the research and relevant to the three studies. The chapter begins by presenting information on THP and its program, followed by an outline of how the participants were recruited. Next, the research design and broad procedures are presented. These details relate to the outcome measures used, data collection, handling of data, and general data analysis. Methodology and procedures more specific to each study are presented in the chapter on that study. Finally, a brief orientation to the statistical procedures used to analyse the data, is presented.

The Helmsman Project

THP is a not-for-profit, public, benevolent institution registered with the Australian Charities and Not-for-Profit Commission. THP provides a program for socioeconomically disadvantaged adolescents that aims to foster personal and social development by cultivating a range of skills and qualities, including hope, resilience, and self-regulation. The ultimate goals behind the program are to improve participants' educational engagement and wellbeing, thereby assisting participants to flourish and reach their full potential. It is intended that the THP program provides a challenging environment in which learning can take place.

THP Program

The THP program is designed to provide experiential learning opportunities, both at the macro level (across the entire program) and at the micro level (within each session). David A. Kolb's (1984, 2014) experiential learning cycle provides a learning framework for the program. The experiential learning process includes both a making sense, and transformation, of experience. More specifically, effective learning through experience requires not only being immersed in a concrete experience, but also being able to observe and reflect on that experience, and then being able to analyse those reflections into abstract concepts and generalisations (David A. Kolb, 1984; 2014). This learning is then applied to a new experience, beginning the next loop of this iterative process. It is the reflection that transforms an experience into learning.

Combined outdoor adventure education/coaching programs. The primary program provided by THP offers a combination of outdoor adventure experiences and developmental coaching (Adventure Program). Developmental coaching is about more than helping a person to set goals and change behaviour; it is aimed at helping the person to develop a more complex understanding of themselves, others, and the systems in which they live (Grant et al., 2010). Developing broader perspective-taking capacities enables participants to perceive different and bigger possibilities than before and improves their ability to meet life's challenges more effectively. Developmental coaching can play a key role in the experiential learning process.

The outdoor adventure experiences in the Adventure Program take place through one of three different modalities: a small-sized group, sailing-based adventure on a yacht known as the *Arctos* with Flying Fish Sailing School; a large-sized group, sailing-based adventure on a tall ship known as the *James Craig* with Sydney Heritage Fleet; or a medium-sized group, land-based adventure with Outward Bound Australia (referred to in this thesis simply as Outward Bound). As the adventure experiences are multimodal, the intervention takes place partially through a number of different forms. Owing to calls for research to provide greater specificity regarding outdoor adventure program components (Hans, 2000), an

overview of those aspects common to each Adventure Program, as well as a broad outline of each separate modality, is included in Appendix B.

Irrespective of the outdoor adventure mode, each THP program runs for 13 weeks, consistent with research that has found longer, more substantial programs lead to better outcomes (Cason & Gillis, 1994; Hattie et al., 1997; Neill, 2007, 2008). Program group size is 8-10 participants, in keeping with the theory that 10 represents the ideal group size for OAE programs; not so small that the group lacks diversity and not so large that individuals lack the necessary attention (Walsh & Golins, 1976; cf., Neill, 2008). However, the size of the group for the adventure experience differs for each mode, with some adventure components combining two or more program groups (see Appendix B). Each program includes a series of individual and group coaching sessions held at the participants' schools, following research that has found the combination of individual and group coaching to be the most effective approach (Robson-Kelly & van Nieuwerburgh, 2016). For the Adventure Program, coaching also takes place during the adventure experiences by way of facilitated reflections and brief coaching huddles after a challenge or break in activity, as well as opportunistic on-the-fly coaching, as needed. Consequently, the coaches assume some of the traditional role of facilitators in OAE programming. Coaches also facilitate reflection after the adventure experiences on the basis that post adventure processing has been found to be important to learning transfer (Duerden, Witt, & Taniguchi, 2012; Leberman & Martin, 2004).

Coaching only programs. For research purposes, THP also has provided a program that is based solely on developmental coaching and skill development, without the outdoor adventure experiences (Coaching Only Program). The primary purpose of initiating this program was to examine the incremental benefits that the outdoor adventure experiences provide for the THP program. Understanding these benefits is important due to the additional costs and logistics involved in implementing the outdoor adventure components. Whereas the outdoor adventure experiences form the focus of the Adventure Program, a personal project lies at the heart of the Coaching Only Program. Additionally, the outdoor adventure experiences are replaced with two workshops to develop the skills and confidence

necessary for preparing and presenting the personal project, as well as for communicating and working with others in anticipation of the group pursuing a community project (see below). A week-by-week outline comparing the coaching sessions in the Adventure Program with the Coaching Only Program, is set out in Appendix C.

Community project. It has been suggested that OAE programs could have more explicit integration with schools and communities (Sibthorp, 2010). In addition to their program, THP provides program group participants with the opportunity to work together to design a project that will make a positive difference in their school community or the wider community in which they live (Community Project). At the end of the program, participants are able to pitch their idea (and related budget) to a THP panel for a chance to be awarded up to AUD 1,000 in funding for their Community Project. The Community Project (and its pitch) are designed to allow participants to apply the learning and skills they've developed during the program, and to continue to build independence and life skills. This project also provides participants with an opportunity to further develop compassion towards others and form connections within their community. Participants have the following two school terms during which to finish planning and implement the Community Project. Coaches are not involved in the Community Project after the pitch, however, the teachers assigned to work with a program cohort (see below for more details) are there to support the participants, as needed. Participation in a Community Project is encouraged but optional. It should also be noted that the option to participate in a Community Project was introduced as an element of the THP program about halfway through the research. Consequently, not every program participant involved in the research was afforded this opportunity.

Coaches

The majority of coaches who facilitate a THP program have a master's degree in coaching psychology or other professional coaching credential. Most of these coaches are also volunteers. Beginning in 2017, available funding allowed THP to

employ one coach in order to bring a higher level of coaching quality and consistency to its programs.

All coaches must have a Working With Children Check through the Office of the Children's Guardian, established under the *Children and Young Persons (Care and Protection) Act 1998* (NSW). Additionally, all of the coaches participate in a training program developed by Dr. Michael Cavanagh, Deputy Director of the Coaching Psychology Unit in the School of Psychology at The University of Sydney. The training sets out the theoretical background to the THP program, and it gives the coaches an opportunity to practise the approaches underlying the program. A coaching manual provides structure and guidance for the coaching sessions, while maintaining a level of flexibility that is fundamental to effective coaching. Coaches are also expected to participate in a number of one-on-one and group supervision sessions during the course of the program.

Cohort Teachers

Each school that participates in a THP program must assign a teacher to support each group of participants from that school (Cohort Teacher). The Cohort Teacher is responsible for coordinating the meetings between the coaches and the cohort of participants, attending the outdoor adventure experiences with the cohort, and assisting the cohort in any Community Project the cohort decides to pursue following participation in the THP program. A guide is provided to the Cohort Teachers to assist them in coordinating the THP program.

Participants

To provide its program, THP has partnered with several schools in Western Sydney, an area with high rates of disadvantage. THP initially ascertained schools of interest through the *My School* database (<https://www.myschool.edu.au>), focusing on schools with an *Index of Community Socio-Educational Advantage* (ICSEA) rating less than the average. THP then contacted these schools of interest to discuss the THP program. If a school had interest in participating in the program, the school principal completed an expression of interest form. THP then determined whether to proceed with that school. In making this assessment, THP focused not only on the proportion of a school's students detrimentally affected by disadvantage, but

also on the level of school support for the THP program. Strong school support and an integration of the program into the participants' school experience, is intended to increase participants' identification and engagement with their school's ethos and values, which has been shown to give rise to enduring positive outcomes for disadvantaged youth (Marsh & Kleitman, 2002). A copy of the recruitment protocol is included in Appendix D.

The target group from these partner schools were adolescents in year nine, who had potential but were at risk of not fulfilling that potential for a variety of reasons, including socioeconomic disadvantage and poor family environments (see Dietrich et al., 2012; Heckman, 2006; Parker et al., 2016). Self-reported demographic information on the participants who participated in this research are included in Appendix E.

Research Design

Consistent with the aims of this thesis, both quantitative and qualitative methodologies were used in the research. Combining methodologies within the context of a single study or research project is often referred to as a *mixed-methods* approach (Gorard, 2010). The literature suggests a variety of reasons why a researcher might use mixed methods in their research, from enhancement of the credibility of research findings, to increased understanding of those findings, to the cross-development of the methods themselves (Hesse-Biber, 2010). Gorard (2010) suggests that mixed methods are really "just a description of how most people would go about researching any topic that they really wanted to find out about" (p. 13). It has been suggested that qualitative interviews are a natural extension of the verbal processing of experience inherent in experiential education (Darl G. Kolb, 1991). Specifically in the field of OAE, it has been suggested that researchers apply qualitative approaches in addition to the more traditional research methods in order to understand the full nature of the OAE experiences and how they influence human perceptions and behaviour (Barrett & Greenaway, 1995; Ewert & McAvoy, 2000; Harper, 2010; Klint, 1990; Martin & Leberman, 2005; McKenzie, 2000; Rowley, 1987). In the context of this research, it is anticipated that a mixed-methods approach will allow for a richer and more complete understanding of the

THP program, the participants, and their relationships with the outcomes of interest.

Quantitative Design

The quantitative design in this research incorporates an RCT with a waitlist control group that allows for an extended baseline design. The use of a waitlist control group overcomes critical ethical issues in RCT design, particularly in educational settings. This design includes both an experimental comparison between the intervention and control groups, as well as an alternative within-subjects comparison of the control group when they participate in the program, providing a test of the replicability of the experimental results. The comparisons are derived from a Likert-style, self-report survey (Survey), which the participants are asked to complete multiple times. Chapter Three provides a detailed overview of the measurement instruments that comprise the Survey and their underlying theoretical constructs, while Chapter Five analyses the psychometric properties of those instruments. Background on the Survey design and data collection procedures is provided below. A diagram of the overall quantitative research design is included in Chapter Six (see Figure 6.1).

Qualitative Design

The qualitative design uses semi-structured interviews that seek to unearth the ways in which participants make meaning of their experiences in a THP program. In qualitative methodology, interviews are said to be ideal when the phenomenon of interest is not something that can be directly observed (Patton, 2002, 2015). In the context of this research, interviews provide a way of gaining insight into how the participants have made meaning of their experiences in a THP program, and how their capacities for meaning making may have developed since their participation. They also provide a different perspective on the impact of the program on our outcomes of interest. Further detail on the interview process is provided below.

Using an interpretative phenomenological analysis and a constructive-developmental lens, the qualitative design includes an evaluation of the interviews to assess participants' capacity for meaning making, as well as any growth in the

complexity of their meaning making. The qualitative design also includes an analysis of the interview data to examine in detail individual participant experience of a THP program, in order to develop a deeper understanding of those experiences and the relationships among the participants, the various program elements, and the outcome measures.

Survey Design and Format

Demographic Information

For each participant, the first time the Survey was administered to that participant, the Survey included nine introductory questions for the purpose of gathering background information on the participant (see Appendix F). This demographic information included the participant's school, grade, month and year of birth, outdoor adventure experience, gender, birth country, parents' birth country, primary language spoken at home, parents'/guardians' highest level of education, and access to certain amenities (room of own, quiet place to study, computer, high speed internet, dictionary, dishwasher).

Survey Items

The Survey originally comprised 214 items from the following 12 measurement instruments: Children's Hope Scale (CHS), Life Orientation Test revised (LOT-R), Adolescent Self-Regulatory Inventory (ASRI), Short Grit Scale (GRIT-S), Academic Resilience Scale (ARS), Life Resilience Scale (LRS), Motivation and Engagement Scale (MES), Satisfaction with Life Scale (SWLS), Warwick-Edinburgh Mental Well-Being Scale (WEMWBS), Self-Description Questionnaire II-Short (SDQII-S), Review of Personal Effectiveness with Locus of Control (ROPELOC), and GSOEP Big Five Inventory (BFI-S). Each measurement instrument comprises one or more scales,⁷ and the original Survey included 45 scales overall.

Subsequently, when the Survey was converted from paper to online format, the 15 items from the GSOEP Big Five Inventory (BFI-S) were removed because there was a desire to reduce the length of the Survey and this instrument was no longer considered relevant to the research. Table 4.1 provides a summary of the 11

⁷ As mentioned in Chapter Three, for multidimensional instruments the term *scale* is used in this thesis to refer to the dimensions of that instrument.

remaining instruments and their scales, including sample items. A complete list of all items, including item cross-references to the original instruments and the Survey, can be found in Appendix A. Further details on these scales and their theoretical and practical connections with the THP program are provided in Chapter Three. More specific information on these scales in connection with their psychometric properties, is provided in Chapter Five.

Table 4.1
Summary of Survey Scales and Sample Items

Scales	Sample Items
Children's Hope Scale (2 scales)	
Agency (3 items)	I think I am doing pretty well
Pathways Thinking (3 items)	When I have a problem, I can come up with lots of
Life Orientation Test-Revised (2 scales)	
Optimism (3 items)	I'm always optimistic about my future
Pessimism (3 items) ^a	I hardly ever expect things to go my way
Adolescent Self-Regulatory Inventory (2 scales)	
Long-Term Self-Regulation (16 items) ^a	I can stay focused on my work even when it's dull
Short-Term Self-Regulation (20 items) ^a	I can start a new task, even if I'm already tired
Academic Resilience Scale (1 scale)	
Academic Resilience (6 items)	I'm good at dealing with setbacks at school (e.g. bad
Life Resilience Scale (1 scale)	
Life Resilience (6 items)	I'm good at dealing with setbacks (e.g. negative feedback on what I do, disappointing outcomes)
Short Grit Scale (2 scales)	
Consistency of Interest (4 items) ^a	I often set a goal but later chose to pursue a different
Perseverance of Effort (4 items)	I finish whatever I begin
Motivation and Engagement Scale-Short (4 scales)	
Booster Thoughts (3 items)	I believe I can do a good job in my schoolwork
Booster Behaviours (3 items)	I plan out how I will do my schoolwork and study
Mufflers (3 items) ^a	I don't think I have much control over how well I do in
Guzzlers (2 items) ^a	I often feel like giving up in my schoolwork
Satisfaction with Life Scale (1 scale)	
Satisfaction With Life (5 items)	I am satisfied with life
Warwick-Edinburgh Mental Well-Being Scale (1 scale)	
Wellbeing (14 items)	I've been feeling good about myself
Self-Description Questionnaire II-S (11 scales)	
Non-Academic Self-Concept (7 scales)	
Physical Abilities Self-Concept (4 items) ^a	I am good at things like sports, gym and dance
Physical Appearance Self-Concept (4 items)	I am good looking
Opp-Sex Relationships Self-Concept (4 items) ^a	I have lots of friends of the opposite sex
Same-Sex Relationships Self-Concept (5 items)	I make friends easily with members of my own sex
Parent Relationships Self-Concept (4 items) ^a	I get along well with my parents
Hon/Trustworthiness Self-Concept (6 items) ^a	I am honest
Emotional Stability (5 items) ^a	I get upset easily
Academic Self-Concept (3 scales)	
Math Self-Concept (4 items) ^a	I get good marks in mathematics
Verbal Self-Concept (5 items) ^a	I get good marks in English
School Self-Concept (4 items) ^a	I am good at most school subjects
Global Self-Concept (1 scale)	
General Self-Esteem (6 items) ^a	Most things I do, I do well

Note. Opp = Opposite; Hon = Honesty.

(continues)

^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

Table 4.1 (continued)
Summary of Survey Scales and Sample Items

Scales	Sample Items
Review of Personal Effectiveness with Locus of Control (14 scales)	
Personal Abilities/Beliefs (4 scales)	
Self-Confidence (3 items)	I am confident in my ability to be successful
Self-Efficacy (3 items)	No matter what happens I can handle it
Stress Management (3 items)	I am calm in stressful situations
Open Thinking (3 items)	I am open to new thoughts and ideas
Social Skills (3 scales)	
Social Effectiveness (3 items)	I am confident and effective in social situations
Cooperative Teamwork (3 items)	I cooperate well when working in a team
Leadership Ability (3 items)	I can be a good leader
Organisational Skills (3 scales)	
Time Efficiency (3 items)	I plan and use my time efficiently
Quality Seeking (3 items)	I try to get the best possible results when I do things
Coping with Change (3 items)	I can cope well when things change
Energy (1 scale)	
Active Involvement (3 items)	I like being active and energetic
Overall Effectiveness (1 scale)	
Overall Effectiveness (3 items)	Overall, in my life I am a very effective person
Locus of Control (LOC) (2 scales)	
Internal Locus of Control (3 items)	If I succeed in life it will be because of my efforts
External Locus of Control (3 items) ^a	My life is mostly controlled by external things

^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

Survey Format

The initial paper version of the Survey (Original Survey) included numbers for each response option, however, the numbering was reversed from what would ordinarily be expected (e.g., 1 for *Strongly agree* and 5 for *Strongly disagree*). Subsequently, the Original Survey was converted into an updated paper version to enable the data to be scanned using Remark Classic OMR[®] software (Remark Survey). The Survey was then moved to an online format using Qualtrics (2017) software, a web-based survey tool (Qualtrics Survey). Numbers were not included for response options in the Remark and Qualtrics Surveys (only words). A copy of the Qualtrics Survey is included in Appendix G.

Research Procedures

Ethics and Participant Consent

Before approaching potential participants for this research, a National Ethics Application Form was submitted to the University of Western Sydney (UWS) Human Research Ethics Committee, and ethics approval was granted on 28 March, 2013. After the research team moved from UWS to the Australian Catholic University (ACU), the ethics approval was transferred to ACU on 2 May, 2014 for the original approval period to 31 December, 2016 (Ethics Approval: 2014-137Q). Subsequently, several modifications and an extension to Ethics Approval: 2014-137Q were sought and obtained. In 2017, the original ethics application, transfer, modifications, and extensions were consolidated and the ACU Human Research Ethics Committee approved this consolidated application on 3 November, 2017 for the period to 31 December, 2018 (Ethics Approval: 2017-252HE). A further modification and extension application was approved on 21 June, 2018, granting an extension to Ethics Approval: 2017-252HE to 30 June, 2020 (see Appendix H).

For research undertaken in New South Wales (NSW) government schools, approval is also required from the Department of Education in NSW. Application for approval was made through the State Education Research Applications Process (SERAP). An initial approval was obtained on 22 July, 2013 (SERAP 2013134), and subsequent extensions were sought in order to maintain the approval through 30 June, 2020. In addition, all persons engaged with children in the context of this research obtained a Working with Children Check through the Office of the Children's Guardian, established under the *Children and Young Persons (Care and Protection) Act 1998* (NSW).

Prior to selection for a program, interested students completed an application form, including a form of parental consent. A copy of the current THP application and consent documentation is included in Appendix I. For those students participating in a research interview, students also signed an additional consent at the time of the interview. A copy of this consent form is included in Appendix J.

Survey Administration and Data Collection Procedures

Participants selected for a THP program or into the waitlist control group were asked to complete the Survey by responding to each item to indicate the extent to which they *strongly agreed*, *agreed*, were *neutral*, *disagreed* or *strongly disagreed* with the statement in that item. Program participants completed the Survey just prior to entry into a THP program (T1), at completion of the program (T2), and approximately three months after program completion (T3). Participants in the waitlist control group completed the Survey together with their corresponding intervention group at T1, T2, and T3 (extended baseline), as well as when they completed a THP program (T4) and approximately three months after program completion (T5).

In general, the project officer from the Institute for Positive Psychology and Education and associated with the research project, administered the Survey to each group of participants together at their school and during school hours. On most occasions, the items in the Survey were read aloud to the participants.

Administration of the Survey had to be coordinated with the schools and at times, this was difficult, particularly when the timing for administration was at the end of a school term. There are occasions when the administration of the Survey had to be deferred until the following term. A table of the program dates and Survey administration dates are included in Appendix K. In addition, there were occasions where one or more participants were absent on a Survey administration date. If it was deemed appropriate in the circumstances, the Cohort Teacher for the particular cohort, was asked to administer the Survey to the absent participant at a later date. In the case of the paper versions of the Survey, the Cohort Teacher provided the late Survey to the project officer, and in the case of the Qualtrics Survey, the Cohort Teacher arranged for the participant to complete the Survey online.

The data output from various versions of the Survey were dealt with in the following ways: data from the Original Surveys were entered by hand into an excel spreadsheet; data from the Remark Surveys were scanned using Remark software and the data saved to an excel spreadsheet, however, scanning was not precise and required subsequent review with adjustments made by hand; and data from the

Qualtrics Surveys were downloaded directly from the secure Australian Catholic University Qualtrics website to an excel spreadsheet. Some excess data needed to be deleted from this download, and some clean-up was required to the column headings, but no response data required amendment. All Survey data was de-identified to remove reference to all participants' names.

Interview Procedures

The schools participating in the THP program invited past program participants to participate in an interview about their experience in the program. Thirteen students from five schools volunteered to participate. The interviews were conducted at the participants' schools. Each interview lasted approximately 90 minutes, including a 30-minute introduction to the interview process. The semi-structured interviews were based on Robert Kegan's *subject-object interview* technique (SOI; Lahey, Souvaine, Kegan, Goodman, & Felix, 2011). Specific details of these 13 participants and the format of the interviews are set out in Chapter Seven.

Quantitative Data Analysis

Software

All quantitative data cleaning and statistical analyses undertaken for this thesis were performed in R version 3.4.1 (R Core Team, 2017) or Mplus version 8 (Muthén & Muthén, 1998-2017). R was used through R Studio (RStudio Team, 2018), an integrated development environment for R. Mplus was used through Mplus Editor and Mplus Diagrammer (Muthén & Muthén, 1998-2017). Specific statistical packages used within R in connection with this thesis are detailed in Appendix L.

Treatment of Missing Data

Missing data is a common issue for most researchers, particularly in the case of longitudinal field-based studies such as the present investigation. As is common in a study involving students and a repeated measures design, many participants had missing data for an entire assessment wave, due primarily to absence from school, change of school, or withdrawal from the program. There also were some challenges coordinating with the schools for timely data collection, and in a couple of instances, this resulted in loss of a wave of data from an entire cohort. Table 4.2

shows the percentage of participants missing data for all items at an assessment wave (T1 -T5). While there were participants in attendance for administration of a Survey who did not provide complete responses to that Survey, all partially completed Surveys were at least 95% complete and considered sufficient for inclusion in this research. For specific details of missing data for Study 2 refer to Table 6.3 in Chapter Six.

Table 4.2

Percentage of Participants Missing All Data for a Wave: T1-T5

Wave	Number of Participants	% Missing All Data
T1	362	11.89
T2	362	19.34
T3	362	19.06
T4	176	43.18
T5	176	63.64

Note. T1 = pre-test (intervention)/extended baseline test (control); T2 = immediate post-test (intervention)/extended baseline test (control); T3 = three months post-test (intervention)/extended baseline test (control); T4 = immediate post-test (control); T5 = three months post-test (control). For details on missing data specific to intervention and control groups relevant to Study 2, see Table 6.3 in Chapter Six.

There are two primary types of missing data: missing data that are missing by design and missing data that are unplanned. The data that is missing for this research project is unplanned. With respect to this missing data, there are three types of mechanisms that can account for the missingness: *missing completely at random* (MCAR); *missing at random* (MAR); or *missing not at random* (MNAR) (Rubin, 1976).

Common approaches to dealing with missing data include *listwise deletion* and *pairwise deletion*. When listwise deletion is implemented, cases with missing data on any variable used in the analysis are removed from the analysis. Pairwise deletion is carried out in connection with analysis that incorporates input vectors (such as means, variances and covariances), and in this context only cases that have data for each input variable (e.g., means, variances) or each pair of input variables (e.g., covariances) are used (Brown, 2015). Both of these approaches, however, are only really appropriate when missing data are MCAR, a condition which is not often

met. Even then, listwise deletion may significantly reduce the available data, which can result in the inflation of standard errors, thereby negatively impacting statistical power and the precision of parameter estimates (Allison, 2002; Brown, 2015; Enders, 2010). Although the method of pairwise deletion preserves more data, it can result in biased standard errors and test statistics and create other issues for model estimation (Allison, 2002; Brown, 2015; Enders, 2010).

Currently, *multiple imputation* and *full information maximum likelihood* (FIML)⁸ are the favoured methods for handling missing data because they can produce unbiased parameter estimates (Allison, 2002; Allison, 2003; Brown, 2015; Enders, 2010). Both approaches are appropriate with missing data that are either MCAR or MAR. In most cases, FIML is regarded to be the best and most appropriate way to handle missing data in the types of statistical analyses conducted in this research (Allison, 2003; Brown, 2015), including where there is a large amount of missing data (Enders, 2010). Accordingly, FIML is the approach that has been used in this thesis, where possible. Specific details on the missing data for each study have been included in the methodology section of the chapter for that study.

Treatment of Response Bias and Negatively-Worded Items

Research has found that survey respondents often respond to survey items independent of the content of those items, a phenomenon variously referred to as *response bias*, *response style* or *method effect* (Marsh, 1996). Such response biases may contaminate the substantive interpretation of a construct under consideration (Alessandri et al., 2010). While specific scales (variously referred to as *lie scales*, *social desirability scales*, or *impression management scales*) have been developed with the aim of detecting certain response biases (e.g., social desirability bias), the validity of these scales as an effective measure of response style has been the subject of criticism (see, e.g., Helmes, Holden, & Ziegler, 2015; MacCann, Ziegler, & Roberts, 2012; McCrae & Costa, 1983; Sackett, 2012; Uziel, 2010). Moreover, in analyses focused on changes in responses by the same person over time, social desirability bias should be less of an issue, assuming there is a similar social

⁸ FIML is also referred to as *maximum likelihood* and *direct maximum likelihood* in the literature (Enders, 2010).

desirability response bias in each wave of data. For these reasons, as well as the desire to avoid any additional increase in the length of the Survey, lie scales were not used in this research.

Other types of response bias (e.g., acquiescence bias) are sometimes managed by including in survey instruments a mixture of positively- and negatively-worded items designed to measure the same construct (Marsh, 1996). Wording items in opposite linguistic direction may also serve to keep respondents more focused on the content of the items overall. This can be particularly important in contexts where respondents are not highly motivated to provide considered, honest responses (Barnette, 2000). However, research has also found that respondents may not read negatively-worded items as carefully or process them in the same way as positively-worded items (e.g., Schriesheim & Hill, 1981; Weems, Onwuegbuzie, & Lustig, 2003). Therefore, including negatively-worded items may impact a scale's validity (Barnette, 2000; Marsh, 1996). This practice can be especially problematic with pre-adolescents, as well as adults with lower education levels (Marsh, 1986a; Melnick & Gable, 1990).

Of the 11 measurement instruments included in the Survey, six of those measures include negatively-worded items. Accordingly, the issues raised above will need to be borne in mind when analysing the data. For the statistical analyses in this thesis, all negatively-worded items were reverse-scored before such analyses were undertaken. In reverse-scoring these items, there is an assumption that agreeing with a positively-worded item and disagreeing with its negatively-worded counterpart is the same. However, this assumption is not always accurate. Whether this presents an issue will be assessed on a scale-by-scale basis. Alternative measures will be taken where deemed necessary.

Structural Equation Models (SEM)

Most of the statistical analyses performed in connection with the research for this thesis, use *structural equation models* (SEM). The specific types of SEMs used in this thesis include confirmatory factor analysis (CFA) and exploratory structural equation modelling (ESEM), which are described in more detail in

Chapter Five, as well as multiple regression models and multiple-group models, which are described in more detail in Chapter Six.

Maximum likelihood (ML) is one of the most widely used estimators for fitting a SEM, and it is the default estimator in Mplus. In cases where the data may not be multivariate normally distributed, a robust version of ML (MLR; Muthén & Muthén, 1998-2017) exists, which corrects standard errors and chi-square test statistics to enhance the robustness of ML from non-normality in the data. Likert-style data with only five response categories are at risk of violating the assumption of multivariate normality (Lubke & Muthen, 2004). Accordingly, MLR is the method of estimation used in the statistical analyses undertaken for this thesis.

Calculating Factor Scores

After an appropriate measurement model has been established through the factor analytic procedures in Study 1 (see Chapter Five “Methodology and Procedures”), factor scores will be derived in Mplus for each latent variable (corresponding to the relevant scale from the Survey) using the regression method. A participant’s factor score for a latent variable is intended to represent the score that would have been observed for the participant were it possible to directly measure the latent variable (Brown, 2015). These factor scores will be used in this Study 2 as proxies for the latent variables. In computing the factor scores, each indicator is weighted for its relative contribution to the factor (Rowe & Rowe, 1999).

Statistical Significance

Statistical significance evaluates the probability of statistical results of sample data analysis, given the sample size and assuming that the sample came from a population in which the null hypothesis is exactly true (Thompson, 2003a). While null hypothesis significance testing (NHST) has been the dominant data analytic approach in the behavioural sciences, it has also been the subject of considerable criticism (Cumming, Fidler, Kalinowski, & Lai, 2012; McCloskey, 2010). In particular, NHST is impacted by sample size and does not directly evaluate effect size or replicability of results (Thompson, 2003a). The Publication Manual of the APA (2010) now emphasises that NHST should be only a starting point, and that estimates of confidence intervals and effect sizes should also be

reported in order to present the most complete understanding of the results. Therefore, while statistical significance is reported in this thesis, the aim is to present a broader perspective of the results where applicable information is available.

Effect Size

Effect size quantifies the extent to which research outcomes diverge from the null hypothesis. This information allows for a better assessment of the practical significance of the studied effect or relationship (American Psychological Association, 2010). In general, an effect size describes the magnitude of a relationship or an effect between two or more variables, in a standardised way. There are a number of different methods for establishing an effect size, and the cut-offs for interpreting the meaning of an effect size statistic is different for each method. In this research, the extent of an item's relationship with its latent factor will rely on the r-squared metric; the size of shared variance between two or more variables will use Pearson's r ; and the size of difference between two or more groups will be determined by the unstandardised regression coefficient using standardised variables, being essentially a standardised effect size (ES). Guidelines for categorising effect sizes for these metrics are set out in Table 4.3. However, as with the other cut-offs suggested in this thesis, these guidelines will not be applied rigidly. Effect size estimates will be presented together with the standard error and an indication of significance, and the practical significance of an effect will be considered with regard to the context of the study and the particular effect (see Ferguson, 2009; Glass, McGaw, & Smith, 1981).

Table 4.3
Guidelines for Categorising Effect Sizes

Effect Size Metric	Size of Effect		
	Small	Medium	Large
R-squared	.02	.13	.26
Pearson's r	.10	.30	.50
ES	.10	.30	.50

Note. The rationale for these benchmarks can be found in Cohen (1988, 1992), rounded to two decimal places for R-squared.

Qualitative Data Analysis

All interviews were recorded with the permission of the relevant participants, and then transcribed manually using HyperTRANSCRIBE software (version 1.6). The transcriptions were then uploaded into NVivo for Mac (version 12). The interview data was scored for constructive-developmental level by the author and an additional scorer, both of whom are certified to score a Subject Object Interview. This data then was assessed using an interpretive phenomenological approach to identify rich narratives and key themes that would further extend the findings of the other components of this thesis.

Summary

The research on which this thesis is based, incorporates a mixed-methods research design. It has been suggested that a mixed-methods approach allows for a more complete understanding of the impact of the THP program on the participants and the outcomes of interest, than a quantitative or qualitative design could achieve on its own. This chapter has provided a broad overview of both the quantitative and qualitative methodologies used in the research. Following this chapter is a separate chapter for each of the three studies, with each chapter presenting a detailed overview and analysis of the specific research aims, hypotheses and questions, methodologies, results, and brief discussion for the relevant study.

CHAPTER FIVE

STUDY 1: EVALUATING THE PSYCHOMETRIC PROPERTIES OF THE MEASUREMENT INSTRUMENTS AND THEIR SCALES

Introduction

The outcomes of interest in this thesis include hope, resilience, self-regulation, motivation, self-concept, wellbeing, and various life effectiveness skills. A Likert-style, self-report survey (Survey) administered to the participants multiple times, attempts to indirectly measure these outcomes. This measurement process forms a critical link between the research hypotheses and questions underlying this thesis and the data collected to represent the outcomes of interest. In order to be in a position to test our research hypotheses and questions with the data, it is essential that these measurements are accurate.

Although the Survey is comprised primarily of previously validated instruments,⁹ it is good practice to reassess an instrument's psychometric properties in connection with each use, because the accuracy of a measurement depends as much on the data collected (including the nature of the participants and the measurement protocols) as on the instrument itself (Marsh & Hau, 2007; Pedhazur & Schmelkin, 1991; Thompson, 1994). Beyond good practice, it is also essential to re-evaluate the psychometric properties of an instrument where its use differs in any way from the instrument's original form. In relation to this research, the Survey is comprised of items from multiple measurement instruments interspersed among each other rather than being included as intact measures. Moreover, as a result of consolidating the measures into a single instrument, the response wording and response range for each of the measures may be different to

⁹ The Life Resilience Scale was created from the Academic Resilience Scale and has not been validated separately from the Academic Resilience Scale. As mentioned in Chapter Three, the terms *instrument*, *measure*, and *measurement instrument* are used interchangeably in this thesis.

its original instrument. Finally, while some of the measures were developed specifically with an adolescent population in mind, others were developed, at least initially, for an adult population. As a consequence, it is important to assess whether each of the measures included in the Survey takes on the same properties when combined into a single instrument in this way, and when used with this particular population. Accordingly, the first study in this thesis is devoted to a psychometric analysis of the measures included in the Survey and the data derived from the administration of the Survey. The psychometric properties of the relevant measurement instruments will be assessed by considering: (a) the reliability of each scale;¹⁰ (b) the structural validity of each instrument and its a priori hypothesised factors; (c) the stability of the structural model for each instrument; and (d) the validity of the construct or constructs each instrument purports to measure.

Reliability refers to the ability of an instrument to measure a construct consistently. In practice, no measure can be perfectly reliable because all data will include some *random error* (Thompson, 2003b). Random error refers to error that is unpredictable and not reproducible, as opposed to *systematic error*, which is more consistent. Reliability testing provides one way of evaluating the impact of random error on an instrument's validity (Kline, 2005). An approach to estimating reliability, referred to as *internal consistency*, focuses on the extent to which there is consistency in the inter-relatedness (or correlations) among the measurement items that purport to measure a construct (Pedhazur & Schmelkin, 1991; Streiner, 2003; Tavakol & Dennick, 2011).

It is also important to confirm the accuracy of the latent structure of a measurement instrument by evaluating the pattern of item-factor relationships through *factor analysis* (Brown, 2015). Beyond this structural evaluation, there is also the consideration of the stability of the factor model both over time and across populations. *Measurement invariance* assesses the extent to which instruments

¹⁰ The terms *scale*, *latent variable*, *factor*, and *latent factor* are used interchangeably throughout this chapter. As mentioned in Chapter Three, for multidimensional instruments the term *scale* is used in this thesis to refer to the dimensions of that instrument. For ease of reference, scale names will be capitalized.

measured in different circumstances maintain the same psychometric properties (Meade, Johnson, & Braddy, 2008).

In order for a measurement instrument to be valid, it is also necessary to establish that it measures what it is intended to measure (referred to as *construct validity*). Construct validity is evaluated through two elements: *convergent validity*, which assesses whether concepts that should be related theoretically are related in reality, and *discriminant validity*, which assesses whether concepts that should not be related theoretically are not related in reality (Campbell & Fiske, 1959).

Particularly in the context of this thesis where multiple scales with overlapping theoretical underpinnings are at the heart of the research, it is important to assess the convergent and discriminant validity of these different scales and the constructs they purport to measure.

The main purpose of investigating the psychometric properties of these instruments and their scales, is to enable us to make legitimate inferences about the participants from their Survey data (e.g., Hubley & Zumbo, 2013). Evaluating the psychometric properties of the instruments in a statistically rigorous manner helps to ensure that the comparisons made in this thesis represent true differences in the constructs of interest. Having given an overview of the purpose of this study, the research aims and hypotheses for this study are outlined next. Following the research aims and hypotheses, the instruments used to measure the constructs of interest in this thesis are discussed. Particular attention is given to outlining the existing research establishing their psychometric properties. The methodology and procedures specific to this study will then be addressed, followed by the results of the analyses and a brief discussion of those results. A more detailed discussion of the complete results for this thesis are included in Chapter Eight.

Research Aims and Hypotheses

Research Aims

The principal aim of this study is to evaluate the participants' Survey data in order to assess the reliability and validity of the measurement instruments and their scales, so as to support the interpretations of the data in the further analyses undertaken in Study 2. This investigation into the psychometric properties of these measures will cover the following research aims:

1. *internal consistency reliability*: to assess the internal consistency reliability of the Survey data for each scale in the measurement instruments;
2. *factor structure*: to assess the a priori hypothesised factor structure of each measurement instrument using the Survey data;
3. *invariance*: to assess the stability of the factor structure for each measurement instrument; and
4. *construct validity*: to assess the convergent validity and discriminant validity of each scale with the Survey data.

Statement of Research Hypotheses

The research hypotheses in this study are based on existing theory and research in connection with each of the relevant measurement instruments and their scales, as well as psychometric theory more generally. Owing to the large number of scales in this study, the research hypotheses are referred to generically, but apply to each of the instruments and their underlying scales. The numbers for the research hypotheses begin with this study number and then are numbered sequentially. Detailed information on the analyses used to assess the research hypotheses is presented in the "Methodology and Procedures" section below.

Research Hypothesis 1.1: Internal consistency reliability. The Survey data for each scale will be reliable as demonstrated by acceptable tests of internal consistency.

Research Hypothesis 1.2: Factor analysis. The a priori factor structure of each instrument will be supported by the Survey data as demonstrated by a factor model with acceptable model fit.

Research Hypothesis 1.3: Invariance analysis. The factor structure for each instrument will be consistent over time, as demonstrated by tests of configural, metric, and scalar invariance from the pre-test (T1) to the immediate post-test (T2).

Research Hypothesis 1.4: Construct validity.

Research Hypothesis 1.4.1: Convergent validity. Correlations of the factor scores for each scale derived from the factor analysis will support the convergent validity of the responses for those scales.

Research Hypothesis 1.4.2: Discriminant validity. Correlations of the factor scores for each scale derived from the factor analysis will support the discriminant validity of the responses for those scales.

Summary

Before proceeding to test the key research hypotheses and questions underlying this thesis, it is essential to establish the accuracy of the instruments used to measure the outcomes of interest, as well as the integrity of the data collected from the participants. Accordingly, the overarching research aims and hypotheses for this study were developed to investigate the psychometric properties of each of the Survey instruments and their scales using the Survey data. The next section provides an overview of each measurement instrument with a particular focus on the existing research in connection with its reliability and validity, in order to establish the rationale for the research hypotheses for this study outlined above.

Measurement Instruments

The Survey incorporates items from 11 instruments with a total of 41 scales. Information on these measurement instruments in relation to their theoretical background and practical connections and relevance to the THP program has been provided in Chapter Three. A brief outline of each instrument follows, highlighting existing research associated with the psychometric properties of the instrument. These details help to establish the rationale for the research hypotheses and provide information relevant to the analysis for this study.

Children's Hope Scale (CHS)

The Children's Hope Scale (CHS; Snyder et al., 1997) is comprised of two scales: Agency and Pathways Thinking. The CHS is one of the most widely used dispositional measures of youth hopefulness, and while it was originally designed for youth aged 8-16 years, subsequent validation studies indicate it to be appropriate for youth up to age 19 (Valle, Huebner, & Suldo, 2004).

Research suggests that this instrument as a whole evidences acceptable internal consistency, test-retest reliability, as well as convergent, discriminant, and incremental validity (Moon & Snyder, 2000; Snyder et al., 1997). Regarding reliability, CHS scores have been found to have internal consistency reliability with a median Cronbach's alpha of .77 among a series of samples, and test-retest correlations over a week and a month of .73 and .71, respectively (Snyder et al., 1997). In a more recent study by Dixson (2017), CHS scores for each scale individually, as well as the total score, were found to be internally consistent in all three samples, with all alpha coefficients equal to or greater than .70. Supporting convergent validity, scores on the CHS were found to correlate positively with scores on various measures of self-concept, self-esteem, optimism, self-efficacy, and wellbeing (Edwards, Rand, Lopez, & Snyder, 2007) and not to demonstrate meaningful associations with hopelessness or intelligence (Snyder et al., 1997). Although the CHS was developed as a dispositional measure and the test-retest reliability indicates stability in CHS scores, it has been suggested that hope is learned (Snyder, 2002) and that the CHS is able to detect change in hope levels over time (Dew-Reeves, Michele Athay, & Kelley, 2012).

In relation to structure, the instrument authors and others have found a two-factor model (Agency and Pathways Thinking) to be the best fit, although the instrument authors intend the item scores to be added together to form an overall hope score (Ciarrochi et al., 2007; Snyder et al., 1997; Valle et al., 2004). However, there are other scholars who do not support this structure, suggesting that the two factors correlate too highly and there are too few items for each factor, to warrant a two-factor model (Dew-Reeves et al., 2012). A recent study examined the factor structure of the CHS in three groups of high school students spanning the range of

academic achievement from academically gifted, to general education, to academically at-risk (Dixson, 2017). In a comparison between the one- and two-factor models, the two-factor model was found to fit the data significantly better in the academically gifted and general education groups (Dixson, 2017). While this was not the case in the academically at-risk group, there were only minimal differences in the fit indices between the two models, and within the two-factor structure, the two scales shared 30% or less of the variance across the three groups (Dixson, 2017). Based on these results, together with the underlying theory and previous research, the two-factor model was accepted as the preferred model for that study (Dixson, 2017). However, Dixson (2017) has suggested that more research on CHS scores in academically at-risk samples is needed. Further, there is no research that could be found that has tested a two-factor ESEM with CHS data.

Past studies indicate that scores on the CHS are invariant across socioeconomic status and age (Snyder et al., 1997; Valle et al., 2004). A more recent study has found CHS scores to be invariant across gender (Dixson, 2017), however, no research has been found that assesses measurement invariance of the CHS items in longitudinal data.

The CHS instrument consists of six, positively-worded items with responses ranging from 1 (None of the time) to 6 (All of the time) on a six-point Likert scale. Three of the items make up the Pathways Thinking scale (e.g., “When I have a problem, I can come up with lots of ways to solve it”) and three of the items make up the Agency scale (e.g., “I think I am doing pretty well”). While all six items were included in the Survey in identical form, the response wording and range differed from the original instrument. Despite these differences, as previous research suggests that the CHS has strong psychometric properties, it is hypothesised that the CHS data will be reliable and valid, and the data will support either a one-factor or two-factor structure. Although there is a lack of longitudinal research in connection with this instrument, given the existing psychometric research, it is hypothesised that the CHS data will also evidence measurement invariance over time.

Life Orientation Test, Revised (LOT-R)

The revised Life Orientation Test (LOT-R; Scheier et al., 1994) is comprised of two scales: Optimism and Pessimism. The LOT-R is one of the most widely used measures of dispositional optimism (see also original Life Orientation Test; Scheier & Carver, 1985), and while it was not developed expressly for adolescents, it has been used with adolescents in a number of studies (e.g., Creed, Patton, & Bartrum, 2002; Monzani et al., 2014; Vassar & Bradley, 2010).

The initial validation study for the LOT-R found the measure to have internal consistency reliability, reporting an alpha coefficient of .78 (Scheier et al., 1994). However, research studies conducted since the initial study have reported a wide variety of reliability coefficients in connection with both the LOT and LOT-R (Vassar & Bradley, 2010). Moreover, a recent meta-analysis of coefficient alphas across studies using the LOT and LOT-R, found adolescent samples to have lower reliability coefficients than other populations (Vassar & Bradley, 2010). While the mean alpha coefficient across non-adolescent populations from this meta-analysis was .74, the mean alpha coefficient for adolescents was much lower at .61, with two studies yielding extremely low coefficients below .5 (Vassar & Bradley, 2010).

The structure of the LOT-R has been the subject of considerable debate. The instrument authors have considered optimism and pessimism to represent opposite ends of a single dimensional scale, and the scoring instructions suggest using a single composite score (Scheier et al., 1994). However, the weight of empirical evidence considering LOT-R data suggests a two-factor structure with low to moderate correlation between the factors (Appaneal, 2012; Carver & Scheier, 2014; Glaesmer et al., 2012; Herzberg et al., 2006; Hinz et al., 2017; Vautier et al., 2003). This two-dimensional structure seems to become more pronounced for older respondents (Herzberg et al., 2006), although research with a sample of 504 adolescents also supported a two-factor model (Creed et al., 2002).

Some argue that the inconsistency between the structure of the operationalised construct and its theoretical conceptualisation brings into question the construct validity of the LOT-R (Monzani et al., 2014). However, there is also the question of what role method variance plays in the factor structure, with the

measure containing both positively- and negatively-worded items (Carver & Scheier, 2014). Vautier et al. (2003) suggest that the empirical support for a two-dimensional structure can be reconciled with the single dimensional optimism construct if consideration is given to the tendency for self-report respondents to present themselves in a positive manner (also referred to as “faking positive”). In that study, the researchers found that the best-fitting model was a bi-factor model, with the first factor (predicting all of the items) measuring optimism and the second factor (predicting the three positively-worded items) measuring an artifact of response style (see also Alessandri et al., 2010; Rauch et al., 2007). They also included the filler items in their model and further argued that the correlation of the positively-framed filler items with the other positive items corroborated their response-style hypothesis (Vautier et al., 2003). A recent study by Monzani et al. (2014) with adolescent subjects, also found this bi-factor model to be the best-fitting model when compared to a congeneric model and a two-factor correlated model (cf., Herzberg et al., 2006). However, it does not appear that this model has been compared with an ESEM for the LOT-R data.

Equivalence of the LOT-R items has been assessed across both age and gender, and the two-dimensional structure of the LOT-R has been found to be gender and age invariant (Glaesmer et al., 2012; Herzberg et al., 2006; Hinz et al., 2017). However, no research has been found which assesses the measurement invariance of the LOT-R items with longitudinal data.

The LOT-R instrument has a total of 10 items: three positively-worded items, three negatively-worded items and four filler items. The 5-point Likert-style response scale ranges from 0 (Strongly disagree) to 4 (Strongly agree). The three positively-worded items make up the Optimism scale (e.g., “In uncertain times, I usually expect the best”), and the three negatively-worded items make up the Pessimism scale (e.g., “If something can go wrong for me, it will”). All 10 items for the LOT-R were included in the Survey, and the response wording for the original instrument is almost identical to that used in the Survey (although the response range is different). Accordingly, given the existing research results in connection with the psychometric properties of the LOT-R, it is anticipated that the LOT-R data will be reliable and valid and will support either a one-factor, two-factor or bi-

factor model. Although there is a lack of longitudinal research in connection with this instrument, given the existing psychometric research, it is hypothesised that the LOT-R data will also evidence measurement invariance over time.

Adolescent Self-Regulatory Inventory (ASRI)

The Adolescent Self-Regulatory Inventory¹¹ (ASRI; Moilanen, 2007) is broken down into two scales: Short-Term Self-Regulation and Long-Term Self-Regulation. Given its recent development, only two psychometric studies of this instrument have been found, with one of those studies using a form of the ASRI translated into Portuguese (Dias, del Castillo, & Moilanen, 2014; Moilanen, 2007). The initial validation study on the ASRI was based on 169 cases and covered only a single point in time (Moilanen, 2007). Accordingly, test-retest reliability could not be evaluated.¹² Following initial reliability analysis in that study, only 27 items were retained of the original 36 items (13 items in the Short-Term Self-Regulation scale and 14 items in the Long-Term Self-Regulation scale). These revised scales demonstrated an alpha coefficient of .70 for Short-Term Self-Regulation and .82 for Long-Term Self-Regulation (Moilanen, 2007). It should be noted that of the negatively-worded items, seven of these are in the Short-Term Self-Regulation scale (more than 50% of that scale) and only two are in the Long-Term Self-Regulation scale (less than 15% of that scale). There is a risk that the unbalanced placement of these negatively-worded items could create an unequal method effect in the scales.

In relation to factor structure, this initial study revealed a slightly better fit for a two-factor structure over a single-factor structure, however, there were issues with some of the loadings (both poor loadings and cross-loadings) and the two factors were strongly correlated ($r = .83$; Moilanen, 2007). The instrument's author suggested that revision or omission of some of the items, and inclusion of additional items, may improve these results. Similar issues were experienced in the

¹¹ A later version of this instrument of the same name comprises 52 items (Moilanen, 2014).

¹² Moilanen (2014) states that the ASRI has displayed high test-retest reliability in a further study, however, she cites her own unpublished manuscript as the source, and that paper could not be obtained for review.

Portuguese study, although not consistently with the same items (Dias et al., 2014). No research on the invariance of the ASRI items has been found.

The ASRI uses a 5-point Likert-style response scale which ranges from 1 (Not at all true for me) to 3 (Neither true nor untrue for me) to 5 (Really true for me). All 36 items were included in the Survey, however, the Survey response wording differed from the original instrument (although the range was the same). The instrument author currently recommends using only 29 items in scoring the ASRI (Moilanen, 2011). Of these 29 items, 15 represent Short-Term Self-Regulation (e.g., “I can start a new task even if I’m already tired”) and 14 represent Long-Term Self-Regulation (e.g., “I can stay focused on my work even when it’s dull”), with an unbalanced mix of positively- and negatively-worded items (eight in the Short-Term factor and two in the Long-Term factor). Notwithstanding the limited psychometric research on this instrument and somewhat problematic results, it is hypothesised that scores for the 29 suggested items will be reliable and valid and will support either a one-factor or two-factor structure. Although there is a lack of invariance research in connection with this instrument, based on the limited research available, it is hypothesised that the ASRI data will also evidence measurement invariance over time.

Short Grit Scale (Grit-S)

The Short Grit Scale (Grit-S; Duckworth et al., 2007; Duckworth & Quinn, 2009) is comprised of two scales: Consistency of Interest and Perseverance of Effort. Consistency of Interest refers to the tendency to stick with goals and interests over time, and Perseverance of Effort refers to the tendency to work hard and maintain effort even in the face of setbacks.

While research has shown the scales to have differential associations with predicted outcomes, the total scale score has been found a better predictor of success than either scale on its own (Duckworth et al., 2007; Duckworth & Quinn, 2009). For this reason, Duckworth and Quinn (2009) describe grit as a compound trait. Consequently, the Grit-S is typically operationalised as a higher-order factor structure with the two scales acting as the first-order factors. An average total score is calculated to determine one’s level of grit. Créde et al. (2016), however, argue that

such a model is problematic for two reasons: first, it cannot be identified at the higher-order level without additional constraints, and second, it cannot be distinguished from a model with two correlated factors. In their meta-analytic study, these researchers found Perseverance of Effort to have a much stronger relation with all academic performance criteria than Consistency of Interest (Credé et al., 2016). They considered the correlations between the two factors to determine whether a high-order factor was plausible, but they found wide variation in the strength of this relation, putting that aspect of the structure in doubt (Credé et al., 2016; Midkiff, Langer, Demetriou, & Panter, 2017). Similar to the LOT-R, there is also the question of what role method variance plays in the factor structure, with Consistency of Interest consisting only of negatively-worded items and Perseverance of Effort consisting only of positively-worded items.

In their validation studies, Duckworth and Quinn (2009) provide evidence for the internal consistency of the Grit-S as a single measure ($\alpha = .73-.83$), as well as its test-retest stability (Duckworth & Quinn, 2009). While the Consistency of Interest scale also demonstrated good internal consistency ($\alpha = .73-.79$), the results for the Perseverance of Effort scale were more variable ($\alpha = .60-.78$) (Duckworth & Quinn, 2009). This study also evidenced a better fit for a two-factor second order structure, when compared with a one-factor model, although the two-factor model did not demonstrate acceptable fit statistics across all samples (Duckworth & Quinn, 2009). The Grit-S has also been found to have invariance across gender (Duckworth & Quinn, 2009), however, a Polish version of the Grit-S was found to have only partial scalar invariance.¹³ While the Grit-S has been found to be strongly associated with Conscientiousness (one of the Big Five factors), there has been evidence of predictive validity beyond Conscientiousness (Duckworth & Quinn, 2009). However, in a recent meta-analytic study, Créde et al. (2016) found that only the Perseverance of Effort factor provided incremental predictive validity over Conscientiousness (Credé et al., 2016; see also MacCann & Roberts, 2010).

¹³ More detailed information on the different types of invariance and their implications is included below in the “Methodology and Procedures” section of this chapter under the heading “Psychometric Analysis”, sub-heading “Invariance testing.”

The Grit-S has eight items, four negatively-worded items for Consistency of Interest (e.g., “I often set a goal but later choose to pursue a different one”) and four positively-worded items for Perseverance of Effort (e.g., “I finish whatever I begin”). Respondents indicate how much the description in the items are like them, using a 5-point Likert-style scale ranging from 1 (Not at all like me) to 5 (Very much like me). All eight items were included in the Survey, and although the original instrument response wording is not identical to the wording in the Survey, the scale range is the same. Notwithstanding the limited psychometric research on this instrument and somewhat mixed results, it is hypothesised that the data for the Grit-S items will be reliable and valid and will support either a one-factor, two-factor, or higher-order model. Although there is a lack of longitudinal research in connection with this instrument, based on the limited research available, it is hypothesised that the Grit-S data will also evidence measurement invariance over time.

Motivation and Engagement Scale-Short (MES-S)

The Motivation and Engagement Scale–Short (MES-S) is comprised of a single item representative of each of the 11 scales from the Motivation and Engagement Scale–High School (MES-HS; Martin, 2007; Martin, 2009) is hypothesised to comprise four scales: Booster Thoughts, Booster Behaviours, Mufflers, and Guzzlers. There is no current research that has assessed the psychometric properties of these items and their scales.

There is broad research to substantiate the validity and reliability of the MES-HS as a measure of academic motivation and engagement, as well as its invariance across gender, age, and school level (Green, Martin, & Marsh, 2007; Marsh, Liem, Martin, Morin, & Nagengast, 2011; Martin, 2007, 2009; Martin & Hau, 2010; Martin, Malmberg, & Liem, 2010). Analysis of data collected from 21,579 high school students from 58 schools supported both the first-order and higher-order factor structures of the 11 scales, as well as the reliability of those scales, with a mean alpha coefficient of .79 (Martin, 2009; Martin, Malmberg, et al., 2010). The research also indicates that much of the variance in motivation and engagement is

at the level of the individual, suggesting a need to focus interventions at the student level (Martin, Malmberg, et al., 2010).

Of particular interest in the research are the correlations among the scales, as this data provides an indication of how the items used to form the revised scales of the MES-S may perform. Existing research found the three dimensions of Booster Thoughts to be strongly and positively correlated with each other, as were the three dimensions of Booster Behaviours (Martin, 2009). However, the three dimensions of Mufflers and two dimensions of Guzzlers had a more moderate relationship. All correlations indicated lower levels of shared variance between the first-order factor groupings than within those factor groupings. At the second-order level, Booster Thoughts and Booster Behaviours correlated strongly (positively) with each other, and both correlated strongly (negatively) with Guzzlers but correlated only slightly (negatively) with Mufflers. Guzzlers and Mufflers correlated at a moderate-to-strong level (positively).

Also of interest is the empirical evidence on the nature of the relationships between the different dimensions of motivation and engagement and important educational outcomes, such as educational aspirations, class participation, and school enjoyment. Martin (2007) found the dimensions of Booster Thoughts and Booster Behaviours to be strongly positively associated with such constructs, and the Guzzlers dimension to be moderately-to-strongly negatively associated. However, the Mufflers dimension was either only weakly negatively, or not at all, associated with those constructs, with the uncertain control scale having the strongest association of the three Muffler scales (Martin, 2007; see also, Martin, 2009).

The MES-S has three positively-worded items for each of Booster Behaviours and Booster Thoughts, three negatively-worded items for Mufflers, and two negatively-worded items for Guzzlers. Each item in the MES-HS is rated on a 7-point Likert-style response scale ranging from 1 (Strongly disagree) to 7 (Strongly agree). The 11 individual scale scores in the MES-HS can be converted to four global scores that reflect the average of Booster Thoughts (e.g., "I believe I can do a good job in my schoolwork"), Booster Behaviours (e.g., "I plan out how I will do my schoolwork and study"), Mufflers (e.g., "I don't think I have much control over how

well I do in my schoolwork”), and Guzzlers (e.g., “I often feel like giving up in my schoolwork”). Given that the MES-S is a modified version of the MES-HS based on this higher-order factor structure, it will be important to test its psychometric properties with the data. However, given the broad existing research on the MES-HS, it is hypothesised that the data for the MES-S items will be reliable and valid and will support a four-factor model, with such model to be invariant over time. However, in the event that the initial analysis is not satisfactory, it may be necessary to go back to an exploratory form of factor analysis.

Academic Resilience Scale (ARS)

The Academic Resilience Scale (ARS) is a unidimensional measure of academic resilience for adolescents (Martin & Marsh, 2006). The initial validation study for the ARS demonstrated high internal consistency ($\alpha = .89$) and excellent fit statistics for a one-factor CFA (CFI = .98, TLI = .96; Martin & Marsh, 2006). Although this study also found the ARS items to be invariant across gender (Martin & Marsh, 2006), no research was found that assessed longitudinal invariance of the measure.

The ARS consists of six positively-worded items (e.g., “I’m good at dealing with setbacks at school (e.g., bad mark, negative feedback on my work)”), with a 7-point Likert-style response scale ranging from 1 (Strongly disagree) to 7 (Strongly agree). All six items were included in the Survey, and the response wording is the same but with fewer response options in the Survey. Therefore, consistent with the original instrument, it is hypothesised that the ARS data will be reliable and valid, and will support the a priori single factor structure. Although there is a lack of longitudinal research in connection with this instrument, based on existing research, it is hypothesised that the ARS scores will also evidence measurement invariance over time.

Life Resilience Scale (LRS)

The Life Resilience Scale (LRS) was created from the ARS for the purpose of this research. Accordingly, there is no existing psychometric research on this measure. The LRS consists of six positively-worded items (e.g., “I’m good at dealing with setbacks (e.g., negative feedback on what I do, disappointing outcomes)”)

with a 7-point Likert-style response scale ranging from 1 (Strongly disagree) to 7 (Strongly agree). All six items were included in the Survey, and the response wording is the same but with fewer response options in the Survey. Although there is no existing psychometric research on this measure, given its similarity to the ARS, it is hypothesised that the LRS data will be reliable and valid, and will support the a priori single factor structure, with such structure to be invariant over time.

Satisfaction with Life Scale (SWLS)

The Satisfaction with Life Scale (SWLS; Diener et al., 1985), a unidimensional measure of subjective wellbeing, has been used in hundreds of studies since its inception (Pavot & Diener, 2008). The measure has demonstrated sound psychometric properties, reflecting high internal consistency with alpha coefficients ranging from .79 to .89 (Pavot & Diener, 1993; see also Vassar, 2008). However, a negative relationship has been found between the SWLS reliability and youth populations (Vassar, 2008). This finding is despite the measure having been developed for use with a wide range of ages and groups (Pavot & Diener, 2008). Accordingly, it will be important to consider reliability of this measure in the context of our sample, which consists of youth, and more specifically, disadvantaged youth.

The SWLS has also exhibited good test-retest correlations, although these correlations show a decline in stability over longer periods (Pavot & Diener, 1993). While the SWLS demonstrates some temporal stability, it has also been shown to be sufficiently sensitive to detect change in life satisfaction over the duration of an intervention (Pavot & Diener, 1993). These results were reaffirmed in Pavot and Diener's (2008) more recent meta-analytic study.

Factor analysis has evidenced a single factor solution (Diener et al., 1985), and this model has been replicated across a variety of populations, languages, and cultural contexts (Pavot & Diener, 1993, 2008). While the fifth item ("If I could live my life over, I would change almost nothing") has shown lower factor loadings and item-total correlations than the other items, it still has been found to be highly correlated with those items (Pavot & Diener, 2008). Pavot and Diener (2008) have suggested that the fifth item could be excluded in circumstances where a researcher

is particularly interested in respondents' satisfaction with their current life, as the fifth item seems to orient respondents to a summary evaluation over years.

In addition, scores on the SWLS correlate moderately to highly with other measures of subjective wellbeing and do not correlate or correlate negatively with clinical measures of distress, impulsivity, and negative affect, providing support for construct validity (Diener et al., 1985; Pavot & Diener, 1993). No evaluation of invariance was found in the literature, suggesting that this is an important area for further examination (Vassar, 2008).

The SWLS consists of a single factor made up of five positively-worded items representing general satisfaction with life (e.g., "I am satisfied with life"). Respondents use a 7-point Likert-style scale to respond, ranging from 1 (Strongly disagree) to 7 (Strongly agree) with 4 being a neutral response (Neither agree nor disagree). All five items have been included in the Survey, and the response wording is the same but with fewer response options in the Survey. Given the existing psychometric research on this instrument, it is hypothesised that the SWLS data will be reliable and valid, and will support the a priori single factor structure. Further evaluation will be made of the instrument's reliability and factor structure within a disadvantaged adolescent population, as well as an assessment of invariance. However, based on the strong psychometric properties of this measure, it is hypothesised that the SWLS data will also evidence measurement invariance over time.

Warwick-Edinburgh Mental Well-Being Scale (WEMWBS)

The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS; Tennant et al., 2007) is a unidimensional measure intended to capture a wide conception of wellbeing. The WEMWBS was developed in part to assess the effect of programs designed to influence mental wellbeing (Stewart-Brown et al., 2009), and has been found to be sensitive to a range of interventions across a number of different populations (Maheswaran, Weich, Powell, & Stewart-Brown, 2012). Research has demonstrated strong internal consistency for the scale, with Cronbach's alpha often well-exceeding the .7 cut-off, but also suggesting some item redundancy (Tennant et al., 2007). Confirmatory factor analysis has also supported the single factor

hypothesised structure of the WEMWBS (Tennant et al., 2007). The scale has been found to correlate with other wellbeing scales, including the SWLS ($r = .73$, $p < .01$; Tennant et al., 2007).

The internal construct validity of the WEMWBS has also been tested within a Rasch Measurement Model, which is said to provide a more robust assessment of the internal construct validity of ordinal scales (Stewart-Brown et al., 2009). In this study, a number of the items were found not to be a good fit to the model, and many of these items also performed differently depending on the gender of the respondent (Stewart-Brown et al., 2009). Based on these results, a shorter 7-item measure was developed called the Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS), which is said to satisfy all criteria for measurement demanded by the Rasch Measurement Model, including strict unidimensionality (Stewart-Brown et al., 2009; see also Bartram, Sinclair, & Baldwin, 2013). However, the SWEMWBS is less broad than the WEMWBS, focusing more on psychological functioning and other aspects of eudaimonic wellbeing, than emotional or hedonic wellbeing (Stewart-Brown et al., 2009). Accordingly, it has been suggested that researchers continue to gather data on the full 14-item instrument (Stewart-Brown et al., 2009).

These measures have also been assessed with adolescent populations. Research with a sample of 1647 adolescents in Scotland and England, aged 13 to 16, reported excellent fit statistics (Goodness of Fit Index [GFI] = 1.000, RMSEA = .003) and good internal consistency ($\alpha = .87$) for the WEMWBS (Clarke et al., 2011; see also Lloyd & Devine, 2012). However, follow-up single-sex focus group interviews with 80 of the participants in that study, revealed that there may be difficulty for adolescents in understanding some of the items (e.g., “energy to spare”) and potential for misinterpretation, particularly those items that were less concrete and required more self-reflection (e.g., “feeling useful”, “thinking clearly”, “interested in new things”). Participants also indicated that some of the relationship items could be understood as relating to a sexual or romantic relationship (e.g., “interested in other people”, “feeling loved”, “feeling close to other people”), and therefore had the potential to lead to misinterpretation or embarrassment (Clarke et al., 2011).

Both measures have been evaluated also in the Australian context (Hunter, Houghton, & Wood, 2015). In that study, which involved 829 Australian adolescents (aged 13 to 16), Hunter et al. (2015) found less strong, but acceptable, fit for the WEMWBS (CFI = .93, TLI = .92, RMSEA = .080 [90% CI: .074, .087]). They found close to excellent fit for the SWEMWBS (CFI = .98, TLI = .97, RMSEA = .066 [90% CI: .050, .083]). Internal reliability of the scores for the SWEMWBS was very good with a Cronbach's alpha of .87 (no alpha was reported for the full WEMWBS). Having considered the modification indices, the authors determined it was appropriate to correlate the uniqueness of item six ("I've been dealing with problems well") and item seven ("I've been thinking clearly"), resulting in an even better fitting model. Although the authors indicated that there appeared to be conceptual justification for this modification, the reasoning is unclear. This SWEMWBS model was then evaluated for invariance across gender and age, and while scalar invariance was found across age groups, only metric invariance was found across gender (Hunter et al., 2015).

A recent study conducted by McKay and Andretta (2017) examined the psychometric properties of both the WEMWBS and the SWEMWBS in more than 9,000 high school students in Scotland and Northern Ireland. This research found the scores in both samples and for both versions of the instrument to have high internal consistency (WEMWBS average $\alpha = .89$ and SWEMWBS average $\alpha = .79$), and the inclusion of other health-related measures indicated sound construct validity for both measures (McKay & Andretta, 2017). In addition, confirmatory factor analysis demonstrated that the one-factor model fit the data well for both samples, with the best fit statistics for the SWEMWBS in the Scotland sample (CFI = .991, TLI = .985, RMSEA = .052; McKay & Andretta, 2017). Gender invariance was also found for both instruments, with the exception of scalar invariance for the SWEMWBS, however, that conclusion was based on a much stricter test than is being used in this research (McKay & Andretta, 2017).¹⁴

¹⁴ For details on the criteria for invariance being used in this study, refer to the "Methodology and Procedures" section in this chapter under the heading "Psychometric Analysis," sub-heading "Establishing goodness-of-fit."

The WEMWBS asks respondents to rate their experience of each of the 14 positively-worded items over the previous two weeks, using a 5-point Likert-style response scale ranging from 1 (None of the time) to 5 (All of the time). Items cover various aspects of mental wellbeing, such as autonomy (“I’ve been able to make up my own mind on things”), positive relationships (“I’ve been feeling close to other people”), positive functioning (“I’ve been thinking clearly”), and positive affect (“I’ve been feeling cheerful”). All 14 items have been included in the Survey, although with different response wording but the same response range. Moreover, there is no request to consider these items in the context of the previous two weeks. However, given the strength of the considerable research indicating that both the WEMWBS and SWEMWBS are psychometrically sound instruments, it is hypothesised that the WEMWBS scores will be reliable and valid, and they will support the a priori single factor structure. Further evaluation will be made of the scale’s reliability and factor structure within a disadvantaged adolescent population, as well as an assessment of measurement invariance of the scale structure. However, based on the strong psychometric properties of this measure, it is hypothesised that the WEMWBS data will also evidence measurement invariance over time.

Self-Description Questionnaire II—Short (SDQII-S)

The Self-Description Questionnaire II—Short (SDQII-S; Ellis, Marsh, & Richards, 2002) is designed to measure 11 self-concept scales, falling under three broad areas: non-academic self-concept, academic self-concept, and global self-concept. Factor analyses of Self-Description Questionnaire data compiled from diverse populations have consistently supported the multidimensionality of self-concept and the distinct facets proposed by the Self-Description Questionnaire suite of instruments (for reviews see Marsh, 1990c, 2005; Marsh & Craven, 1997). Beyond factor analysis, these instruments have been the subject of extensive psychometric evaluation and, having been found reliable, valid, and psychometrically sound across a variety of populations, are said to be the most well-validated measures of multidimensional self-concept in use (Byrne, 1996; Gilman, Laughlin, & Huebner, 1999; Leung, Marsh, Yeung, & Abduljabbar, 2015; Marsh, 1990c).

Within the academic self-concept scales, the Math and Verbal Self-Concept scales have been found to be nearly uncorrelated with each other but highly correlated with the School Self-Concept scale (Marsh, Byrne, & Shavelson, 1988). The non-academic self-concept scales have been found to be somewhat correlated with each other and more highly correlated with each other than with the academic self-concept scales (Marsh & Craven, 1997). This lack of correlation between the Math and Verbal Self-Concept scales is said to result from the way in which academic self-concept develops, referred to as the *internal/external frame of reference model* of self-concept (I/E model; Marsh, 1986b, 1990c, 2005). According to the I/E model, academic self-concept in a subject is based on two frames of reference: an external reference of comparison between one's self-perceived performance in a subject relative to other students and relative to other external standards of actual achievement; and an internal reference of comparison between one's performance in one subject with their performances in other school subjects (Marsh, 2005). The I/E model represents a modification to the *Shavelson model* of self-concept, known as the *Marsh/Shavelson revision* (Marsh, 1990c).

The SDQII-S, in particular, has been found to have strong internal consistency, with reliability estimates consistently at an acceptable level for each scale (Bodkin-Andrews, Ha, Craven, & Yeung, 2010; Ellis et al., 2002; Marsh, Ellis, et al., 2005). Factor analysis across a number of studies has also evidenced good fit for the instrument's a priori 11-factor structure (Bodkin-Andrews et al., 2010; Ellis et al., 2002; Marsh, Ellis, et al., 2005). In a sample of 1,725 indigenous and non-indigenous students, a CFA of the SDQII-S evidenced strong fit statistics: CFI = .97, TLI = .96, and RMSEA = .048 (Bodkin-Andrews et al., 2010). That study also demonstrated invariance of factor loadings and intercepts for the factor structure of the SDQII-S across indigenous and non-indigenous male and female students (Bodkin-Andrews et al., 2010). Multitrait-multimethod analysis conducted by Marsh, Ellis, et al. (2005) has also provided good support for the short-term stability of the latent factors over time, as well as the convergent and discriminant validity of responses to the SDQII-S.

The SDQII-S has 51 items, 32 of which are positively worded and 19 of which are negatively worded. Sample items are set out in Table 5.1. Respondents use a 6-

point Likert-style response scale to indicate for each item from 1 to 6 how false (not like me at all) or how true (this statement describes me well) each statement is as a description of them. All 51 items were included in the Survey, although the Survey response scale is different to the range and wording used in the original SDQII-S instrument. However, based on the extent of the psychometric research and consistency of the findings, it is hypothesised that the SDQII-S data will be reliable and valid, and will support the a priori 11-factor structure, as well as measurement invariance of the scale structure over time.

Table 5.1
SDQII-S Sample Items

Scales	Sample Items
<i>Non-Academic SC (7 scales)</i>	
Physical Abilities SC (4 items) ^a	I am good at things like sports, gym and dance
Physical Appearance SC (4 items)	I am good looking
Opposite-Sex Relationships SC (4 items) ^a	I have lots of friends of the opposite sex
Same-Sex Relationships SC (5 items) ^a	I make friends easily with members of my own sex
Parent Relationships SC (4 items) ^a	I get along well with my parents
Honesty/Trustworthiness SC (6 items) ^a	I am honest
Emotional Stability (5 items) ^a	I get upset easily
<i>Academic SC (3 scales)</i>	
Math SC (4 items) ^a	I get good marks in mathematics
Verbal SC (5 items) ^a	I get good marks in English
School SC (4 items) ^a	I am good at most school subjects
<i>Global SC (1 scale)</i>	
General Self-Esteem (6 items) ^a	Most things I do, I do well

Note. SDQII-S = Self-Description Questionnaire II-Short.

^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

Review of Personal Effectiveness with Locus of Control (ROPELOC)

The Review of Personal Effectiveness with Locus of Control scale (ROPELOC; Richards et al., 2002) was developed to measure important psychological and behavioural dimensions of life effectiveness. It consists of 14 scales that can be broadly grouped into six categories: personal abilities and beliefs, social skills,

organisational skills, energy, overall effectiveness, and locus of control. The ROPELOC (and its predecessors) were designed to be sensitive to the effects of personal development programs based in experiential learning, and they have been used extensively in research involving youth and experiential learning programs, including those with an emphasis on outdoor adventure (e.g., Culhane, 2004; Ellis, Marsh, & Craven, 2009; Imholt, 2009; Johnson, 2012; Luo, 2011; Merrell, 2009; Neill et al., 2003; Richards et al., 2002).

The ROPELOC scales have been found to have good internal reliabilities. The first trial of 1,250 high school students demonstrated Cronbach alphas for the 14 scales of between .79 and .93, with an average internal reliability of $\alpha = .85$ (Richards et al., 2002). Similar internal reliabilities were found in the second trial sample ($n = 1,450$). The factor structure of the ROPELOC has been supported by both exploratory and confirmatory factor analysis, with factor loadings ranging between .65 and .90, and a TLI for each trial of .925 and .940, respectively (Richards et al., 2002). In addition, factor correlations among the scales (average $r = .43$) evidence good discrimination between the closely-related dimensions of life effectiveness (Richards et al., 2002). The factor structure and reliabilities of the LEQ have been found to be consistent over age and gender (Neill et al., 2003), however, no research considering the measurement invariance of the ROPELOC over time was found.

The ROPELOC has three items for each of the 14 scales, as well as three additional control items. All items are positively worded except for the external locus of control scale where the items are all negatively worded. Sample items are set out in Table 5.2. The items are rated on an 8-point Likert-style response scale, ranging from 1 (False, not like me) to 8 (True, like me). All 45 items have been included in the Survey, although the Survey response scale is different to the range and wording used in the original ROPELOC instrument. However, given the strength of the existing psychometric findings on the ROPELOC and related instruments, it is hypothesised that the ROPELOC data will be reliable and valid, and will support the a priori 14-factor structure, as well as measurement invariance over time.

Table 5.2
ROPELOC Sample Items

Scales	Sample Items
<i>Personal Abilities/Beliefs (4 scales)</i>	
Self-Confidence (3 items)	I am confident in my ability to be successful
Self-Efficacy (3 items)	No matter what happens I can handle it
Stress Management (3 items)	I am calm in stressful situations
Open Thinking (3 items)	I am open to new thoughts and ideas
<i>Social Skills (3 scales)</i>	
Social Effectiveness (3 items)	I am confident and effective in social situations
Cooperative Teamwork (3 items)	I cooperate well when working in a team
Leadership Ability (3 items)	I can be a good leader
<i>Organisational Skills (3 scales)</i>	
Time Efficiency (3 items)	I plan and use my time efficiently
Quality Seeking (3 items)	I try to get the best possible results when I do things
Coping with Change (3 items)	I can cope well when things change
<i>Energy (1 scale)</i>	
Active Involvement (3 items)	I like being active and energetic
<i>Overall Effectiveness (1 scale)</i>	
Overall Effectiveness (3 items)	Overall, in my life I am a very effective person
<i>Locus of Control (LOC) (2 scales)</i>	
Internal Locus of Control (3 items)	If I succeed in life it will be because of my efforts
External Locus of Control (3 items)	My life is mostly controlled by external things

Note. ROPELOC = Review of Personal Effectiveness with Locus of Control.

Summary

The Survey scales are at the heart of the investigations in this study. The constructs underlying this thesis and the instruments used to measure those constructs, were outlined in Chapter Three. This section reintroduced those measurement instruments and highlighted the existing research in connection with their reliability and validity to establish the rationale for the research hypotheses for this study. Before presenting the results, the next section outlines the methodology and procedures used in this study to conduct the psychometric analysis of the measurement instruments included in the Survey.

Methodology and Procedures

Participants

The sample for this study consisted of 362 students in Year 9, 51.7% of which were female. Each participant had the opportunity to take part in a THP program, either as a member of the intervention group or through the waitlist control group. The THP program and the different modes through which the program is offered, are described in detail in Chapter Four. Participants were from 11 high schools all located in Western Sydney, an area with high rates of socioeconomic disadvantage. Self-reported demographic information on these participants is included in Appendix E.

Survey

Participants in the intervention group completed the Survey just prior to entry into the THP program (T1), at completion of the program (T2), and approximately three months after program completion (T3). Participants in the waitlist control group completed the Survey together with their corresponding intervention group at T1, T2, and T3 (extended baseline), as well as when they completed a THP program (T4) and approximately three months after program completion (T5). The measurement instruments included in the Survey and the existing research on their psychometric properties have been described above.

Data Analysis

Treatment of missing data. For an overview of the missing data and its general treatment in this research, refer to Chapter Four. Most of the analyses in this study used long form data. Missing data presents less of an issue with long form data as each assessment wave for a participant is considered separately. Accordingly, if a participant had missing data for an entire assessment wave, that assessment wave for that participant was not included in the long form data set. As indicated in Chapter Four, missing data in a partially completed Survey was handled primarily by using *full information maximum likelihood* (FIML).¹⁵ In most

¹⁵ FIML is also referred to as *maximum likelihood* and *direct maximum likelihood* in the literature (Enders, 2010).

cases, FIML is regarded to be the best and most appropriate way to handle missing data in CFA and SEM (Allison, 2003; Brown, 2015). For internal consistency analysis conducted in R, missing data was handled by listwise deletion.

Treatment of negatively-worded items. Of the 41 scales included in the Survey, 15 of those scales include negatively-worded items. Potential issues in connection with negatively-worded scale items have been discussed in Chapter Four. For the statistical analyses in this study, all negatively-worded items were reverse-scored before such analyses were undertaken, so that a high value indicates the same type of response on a negatively-worded item as it does on a positively-worded item. Reverse scoring is essential for assessing reliability in a scale that includes both positively- and negatively-worded items.

Psychometric Analysis

Reliability.

Coefficient alpha. One method of estimating internal consistency that is widely used by researchers is known as Cronbach's (1951) *coefficient alpha* or α (Dunn, Baguley, & Brunsden, 2014; Hogan, Benjamin, & Brezinski, 2003; Pedhazur & Schmelkin, 1991). Although coefficient alpha is a popular statistic for confirming reliability of a scale, it does have limitations. To begin, the formula for coefficient alpha incorporates the number of items in a scale, so as the number of items increases, the value of coefficient alpha increases (Pedhazur & Schmelkin, 1991; Streiner, 2003). The formula also depends on total score variance, which will differ depending on the sample. The more heterogeneous the sample, the larger the total variance and the higher the reliability (Streiner, 2003). In addition, the grounding of coefficient alpha in the essentially *tau equivalent* measurement model may have an impact on its value. Tau equivalence requires that items that measure the same latent factor or construct do so using the same scale and with identical precision and error; that is, with equality of factor loadings (Brown, 2015; Dunn et al., 2014; Raykov, 2004). However, this condition is often not met, particularly where the scale is arbitrary, as is the case with Likert-style scales (Raykov, 2001; Raykov, 2004). Where this condition is not met, coefficient alpha may underestimate scale reliability, depending on whether the measurement model contains correlated

uniquenesses or not, as well as other underlying measurement parameters (Dunn et al., 2014; Graham, 2006; Raykov, 1997; Raykov, 2001; Raykov, 2004; Zimmerman, 1972). As a result, coefficient alpha provides only a lower-bound estimate of the true reliability of a scale when measures are congeneric.

Coefficient omega. *Coefficient omega* or ω , first suggested by McDonald (1999), does not assume tau equivalence. Based on a factor analytic model, coefficient omega uses the item factor loading and uniqueness to estimate reliability. Thus, when estimating reliability with coefficient omega, both the error variances and the factor loadings can vary by item. A number of researchers have shown coefficient omega to be generally a more appropriate index of internal consistency, both in relation to coefficient alpha and when compared to other alternatives (Graham, 2006; Zinbarg, Revelle, Yovel, & Li, 2005; Zinbarg, Revelle, & Yovel, 2007). Therefore, coefficient omega was relied upon in this study as a test of internal consistency, although coefficient alpha is also reported in Appendix M for good measure.

Threshold estimates. Although there is no universally agreed minimum threshold for reliability reporting, values of .70 are generally considered “adequate”, with values around .80 being “very good” and values around .90 being “excellent” (Kline, 2016, p. 92). However, it has been acknowledged that with psychological constructs such as those the subject of this research, the values of reliability estimates can fall below .70 (Field, Miles, & Field, 2012). This is because scale items for these constructs may be less homogenous, thereby reducing reliability estimates (Cohen, Swerdlik, & Sturman, 2013). For the purpose of this study, values of coefficient omega at or above .70 were considered acceptable. However, owing to the limitations of the reliability estimates outlined above and consistent with calls for these “rules of thumb” to act only as a preliminary assessment (Marsh & Hau, 2007; Marsh, Hau, & Wen, 2004), any final determination as to a scale’s reliability included a subjective analysis of the data, taking into account the particulars of the research. In particular, where a scale’s reliability coefficients were below the .70 threshold, consideration also was given to the inter-item correlations. It has been proposed that inter-item correlations between .20 and .40 are appropriate as they

suggest that the items have sufficient content overlap, but are not so homogenous as to make them redundant (Cohen et al., 2013). Therefore, analyses may have proceeded with scales demonstrating a reliability coefficient below .70, where there was satisfactory assessment of their inter-item correlations and factor analytic properties. It is also relevant to note that the analyses in Study 2 are based on factor scores derived from the latent variables (as opposed to manifest variable scale scores) which account for measurement error, thereby reducing attenuation of parameter estimates due to issues of reliability.

Finally, reliability estimates traditionally have been reported as point estimates. However, it has been argued by many researchers that reporting a single reliability coefficient is not sufficient as it does not reflect the level of reliability of the estimate itself (Dunn et al., 2014; Fan & Thompson, 2001; Iacobucci & Duhachek, 2003; Kelley & Pornprasertmanit, 2016; Raykov & Shrout, 2002). The reporting of confidence intervals is considered to be a benchmark for rigorous statistical reporting in psychology and other disciplines (American Psychological Association, 2010; Association for Psychological Science, 22 October, 2018). Confidence intervals indicate the range of values likely to include the true effect. Therefore, in addition to reporting the point estimates for coefficient omega, the 95% confidence intervals also are reported. All reliability analyses used long form data and were conducted in R using the function `ci.reliability` from package MBESS (version 4.4.3) with maximum likelihood parameter estimation robust to non-normality (MLR). Missing data was handled for these analyses with listwise deletion.

Factor analysis. *Structural equation modelling* (SEM) is a tool used for assessing modelled interrelations among a set of items. In SEM, sample size has an impact on the ability to accurately assess model fit and the precision of the parameter estimates (Brown, 2015). However, the question of an appropriate sample size is not one that is easily answered, as it depends on a number of characteristics, including study design, various qualities of the indicators, estimator type, missing data, and model complexity (Brown, 2015).

As a result of the longitudinal nature of the data collected for this thesis and in order to increase power in the statistical analysis, long form data was used for the factor analysis. The analysis was performed in Mplus (Muthén & Muthén, 1998-2017) using the participant as the cluster variable together with the Mplus complex design option, in order to appropriately adjust standard errors to take into account the non-independence of observations that is a consequence of using long form data. Maximum likelihood parameter estimation was used with standard errors and a chi-square test statistic that are robust to non-normality and non-independence of observations (MLR; Muthén & Muthén, 1998-2017).

Confirmatory factor analysis. CFA is a type of SEM that is concerned specifically with evaluating the relationships between observed items and latent constructs in a measurement model (Brown, 2015). In research that uses previously validated instruments (as is mostly the case in this thesis), researchers typically use CFA to confirm the a priori hypothesised factor structure of those instruments.

A CFA was conducted separately for each measurement instrument used in this thesis. The CFA models in this study scale the latent variables by fixing the variance of each latent factor to 1. All items for a latent variable were allowed to freely load on that factor, but they were constrained to have a zero loading on all other factors in the model. In addition, while the factors were permitted to correlate, item uniquenesses (also referred to as measurement errors) were uncorrelated, except as expressly specified in the results section.

Although there are many advantages to using CFAs, these models often do not reflect an acceptable model fit (Marsh, Lüdtke, et al., 2010). As indicated, a simple structure for a CFA model is generally specified in which items load on only a single factor, with non-target loadings constrained to be zero (Brown, 2015), sometimes referred to as independent cluster models of confirmatory factor analysis (Marsh, Morin, Parker, & Kaur, 2014a). CFA provides for a more parsimonious measurement model, however, items in a multidimensional measure may have small cross-loadings, often prompted by item content (e.g., method effects) or theoretically supported correlations among the scales (Asparouhov & Muthén, 2009). Marsh, Lüdtke, et al. (2010) also argue that particularly in the case of

psychological measurement, items often have multiple determinants, with non-zero cross-loadings a natural consequence of those interrelationships (see also Marsh et al., 2014a; Morin, Arens, & Marsh, 2016). In such a situation, fixing the cross-loadings to zero may result in a model that is more parsimonious than is suitable for the data (Asparouhov & Muthén, 2009; Asparouhov, Muthén, & Morin, 2015). It may also overestimate the factor correlations in the CFA, which in turn may lead to other biased estimates in analyses based on the CFA measurement model (Asparouhov & Muthén, 2009; Marsh, Lüdtke, et al., 2010; Marsh et al., 2009; Marsh, Nagengast, & Morin, 2013).

In this thesis, where the a priori model did not fit the data well, a *model-generating* approach was adopted (Byrne, 2012). In that situation, the model may be re-specified where such re-specification is based on empirical, conceptual, or practical considerations, and supported by the applicable theory (Brown, 2015). One alternative model is a bifactor model, which has been used in prior research particularly in connection with multidimensional measures (Reise, 2012). A bifactor model specifies a single general factor that accounts for common variance among all of the items in a measure, as well as one or more group factors that reflect additional common variance among specified clusters of items (Reise, 2012). In a confirmatory bifactor model, each item may load on the general factor and only one group factor, with all other loadings fixed to zero (Hublely & Zumbo, 2013). The group and general factors are specified to be orthogonal, and where the general factor is the substantive factor, the group factors are also specified to be orthogonal.¹⁶ Another alternative approach to CFA that was considered is ESEM. This approach is described in more detail below.

Exploratory structural equation modelling. It is suggested that exploratory factor analytic approaches are preferable to utilising modification indices for model refinement (Asparouhov & Muthén, 2009; Asparouhov et al., 2015; Brown, 2015; Marsh et al., 2009). While basic exploratory factor analysis

¹⁶ Another alternative is the second-order or hierarchical model, which is a direct alternative to a CFA with correlated factors. In this model the first-order latent factors load onto a second-order latent factor and the first order factors are specified to be orthogonal. No such models were used in this study.

(EFA) avoids some of the pitfalls of CFA, it has its own limitations. For example, as traditionally applied, EFA does not provide fit indices, does not allow for measurement invariance to be tested, and does not allow the inclusion of additional variables to validate the empirical factors (Marsh et al., 2014a). A more recent approach being used in latent construct measurement modelling is ESEM, a method that integrates EFA, CFA and SEM (Asparouhov & Muthén, 2009; Marsh, Lüdtke, et al., 2010; Marsh et al., 2009). Unlike CFA which restricts cross-loadings to zero, ESEM allows all factor loadings to be estimated, excluding any loadings that require constraint for model identification (see Asparouhov & Muthén, 2009; Marsh et al., 2009). In particular, with ESEM it is possible to specify target loadings for the primary items predicted by a latent factor using a target rotation, while still allowing non-target items to be estimated close to zero (see Asparouhov & Muthén, 2009; Marsh et al., 2009). A target rotation lies between the mechanical approach to EFA rotation and the strict approach to CFA model specification (Asparouhov & Muthén, 2009). This type of rotation is consistent with ESEM as being more confirmatory than exploratory in nature, and is said to be particularly appropriate where there is a clearly defined a priori factor structure (Marsh et al., 2014a).

ESEM can overcome many of the important limitations of CFA, while at the same time avoiding those limitations inherent in traditional EFA. This modelling methodology often results in better fit, as well as latent factors that are more accurately estimated and less correlated (Asparouhov & Muthén, 2009; Asparouhov et al., 2015). Importantly, valuable information (e.g., SEM-style parameter estimates, standard errors, and goodness-of-fit statistics) and statistical advances typically associated with CFA and SEM modelling, are accessible through ESEM (Marsh, Lüdtke, et al., 2010; Marsh et al., 2009). However, ESEMs with a large number of items lack parsimony, and this lack of parsimony can make it difficult to attain a solution as a result of computational issues.

Accordingly, for multidimensional measures, an ESEM using target rotation was conducted and compared to the CFA. When assessing an ESEM against a CFA, the ESEM was preferred if the factors were appropriately identified, the fit indices for the ESEM were meaningfully better than for the CFA, and the factor correlations were meaningfully smaller for the ESEM than for the CFA (Marsh et al., 2014a). An

ESEM with substantively better fit indices indicates that the estimated factor correlations for the CFA solution are likely to be substantially biased (Marsh et al., 2014a). Where the fit indices and factor correlations were nearly the same, the CFA was preferred on the basis of parsimony (Marsh et al., 2014a). Information on fit indices is provided below in the section headed “Establishing goodness-of-fit.”

Invariance testing. After establishing an acceptable fitting measurement model for each instrument, the latent factors were evaluated for invariance. Longitudinal invariance ensures that a measure is consistently assessed over time (Vandenberg & Lance, 2000). Such testing can also meet concerns about potential response-shift bias (Oort, 2005). Testing the measures used in this study for longitudinal invariance allows for the longitudinal comparisons to be made in Study 2. Although not directly relevant to the research questions in this thesis, the latent factors were also evaluated for invariance across gender. For completeness, these results are presented in Appendix N as a supplemental research hypothesis.

Invariance testing involves the step-by-step comparison of a number of models in which successive aspects of the factor structure are systematically constrained across time or groups, and then comparing these competing models (Cheung & Rensvold, 2002). The first model is completely unconstrained, testing only the equality of factor structure, referred to as *configural invariance* or *weak factor invariance* (Horn & McArdle, 1992). This model tests whether the measures reflect the same underlying structure, and it serves as the baseline model for subsequent tests of invariance. The second model constrains the factor loadings to be equal but allows the other parameters to remain freely estimated, referred to as *metric invariance* or *strong invariance* (Horn & McArdle, 1992). This model tests whether the measures have the same meaning and structure for different groups of respondents (e.g., males and females) or at different times. The third model constrains the factor loadings and the item intercepts, while allowing the other parameters to remain freely estimated, referred to as *scalar invariance* or *strict invariance* (Horn & McArdle, 1992). Although there are additional models in which further parameters may be constrained, these are the three models most commonly

investigated in order to establish invariance (Meade et al., 2008). Accordingly, these are the models that have been used in this study.

Longitudinal invariance was tested using the best-fitting model from the factor analysis described above. While invariance testing mostly has been discussed in connection with CFA, it is also available with ESEM (Asparouhov & Muthén, 2009). For each measurement instrument, the best-fitting factor structure was first modelled using the wide form data for T1 and T2 simultaneously for the relevant scales completely unconstrained, other than correlated uniquenesses for the same item at each timepoint¹⁷ (the configural model). This model was then compared with a similar model with constrained factor loadings (the metric model), and then with a similar model with constrained factor loadings and constrained item uniquenesses (the scalar model). These analyses were performed in Mplus with MLR estimation.

Establishing goodness-of-fit. There are various goodness-of-fit statistics that provide information about how well a model's parameter estimates reproduce the variances and covariances in the sample data (i.e., how well the model fits the data). While the chi-square test was the first fit statistic to be developed, it has a number of disadvantages. To begin, it is based on very stringent standards of perfect fit, so minor misspecifications can lead to rejection of good or reasonable fitting models (Brown, 2015; Marsh, Morin, Parker, & Kaur, 2014b). In addition, the chi-square statistic is sensitive to sample size, as well as to small deviations from multivariate normality, which can also lead to an inappropriate result (Brown, 2015; Cheung & Rensvold, 2002; Marsh & Balla, 1994; Marsh, Balla, & McDonald, 1988; Marsh et al., 2014b).

¹⁷ When the same item is used on multiple occasions, a correlation between the unique components of each item on those two occasions (beyond the correlations explained by the factors) is likely to exist. The failure to include correlated uniquenesses for these items may bias the parameter estimates and inflate the test-retest correlations among the matching latent factors (Marsh & Hau, 1996; Marsh et al., 2011). Therefore, correlated uniquenesses have been included between matching items at T1 and T2 in order to account for these unique associations.

Fit indices for factor analysis. Issues with the chi-square test statistic have led to the development of various fit indices to supplement the chi-square test (Cheung & Rensvold, 2002). These alternative fit indices tend to fall into one of two categories: absolute fit indices (which assess how well an a priori model reproduces sample data) and incremental fit indices (which measure the proportionate improvement in fit by comparing a target model with a more restricted, nested baseline model, also referred to as comparative fit indices) (Hu & Bentler, 1999).

Consistent with recent approaches in applied SEM research, the CFAs and ESEMs for this study were evaluated on the basis of the following fit indices: the *comparative fit index* (CFI; Bentler, 1990) and the *Tucker-Lewis index* (TLI; Tucker & Lewis, 1973), both incremental fit indices, and the *root mean square error of approximation* (RMSEA; Steiger & Lind, 1980), an absolute fit index (Brown, 2015; Cheung & Rensvold, 2002; Hu & Bentler, 1999; Marsh & Balla, 1994; Marsh, Balla, et al., 1988; Marsh, Hau, & Grayson, 2005; Marsh et al., 2004; Marsh et al., 2011; Meade et al., 2008). Of these fit indices, the TLI and RMSEA penalise for model complexity that does not meaningfully improve fit, while the CFI contains no such penalty. Therefore, it is important to recognise that the inclusion of additional parameters in a model may result in the CFI indicating an improved fit that is unwarranted.

The values of the TLI and CFI vary along a continuum from zero to one,¹⁸ and values at or greater than .90 and .95 are generally taken to reflect “acceptable” and “excellent” fits, respectively, to the data (Bentler, 1990; Marsh & Balla, 1994; Marsh et al., 2004; Marsh et al., 2014b). For the RMSEA, zero indicates a perfect fit, and while the upper range of the RMSEA is unlimited, it is said to be rare to see a value above one (Brown, 2015). In terms of assessing fit, values of the RMSEA at or less than .08 and .06 support an “acceptable” and “good” fit to the data, respectively (Browne & Cudeck, 1992; Marsh et al., 2014b). The 90% confidence interval for the RMSEA is also reported. While the chi square test is not used as an indicator of

¹⁸ Technically, this refers to the population values so that actual observed values of the TLI can fall outside the range of 0 to 1, however, it is generally interpreted in a similar manner to the CFI (Brown, 2015).

model fit for the reasons already mentioned, the test statistic is nevertheless reported for good measure in the appendices (See Appendix O and Appendix P).

Additional parameters of evaluation. These indices provide only one aspect of evaluating model adequacy (Marsh & Hau, 2007; Marsh et al., 2004). In addition, and consistent with the suggestions of Marsh et al. (2014b), detailed consideration also was given to factor loadings, factor correlations, and any other parameter estimates that were relevant in the particular circumstances. In relation to factor loadings, there are various cut-off criteria that researchers use. Particularly in the context of exploratory factor analysis, the most common indicator for a salient loading has been an absolute value of at least .30 for the standardised coefficient (Osborne, 2014). In this study, the .30 cut-off was considered minimally acceptable, and for ESEM, factor loadings for each item on its target factor should be higher than cross-loadings on the other factors. In addition, excessively high factor correlations (e.g., greater than .80) were found to detract from discriminant validity (see the section headed “Validity” below). While the amount of variance in an item’s score that is attributable to its factor is also an important parameter to consider and report (Jackson, Gillaspay, & Purc-Stephenson, 2009), this statistic can be derived directly from the factor loadings and, therefore, is reported in detail. Ultimately, some level of subjectivity and judgement play a role in selecting the most appropriate model for further analysis (Browne & Cudeck, 1992). Where the fit statistics for a measure’s a priori factor structure were found to be less than acceptable modelled with either a CFA or ESEM, then if there was an alternative structure that has been evidenced in the literature or otherwise could be supported with existing theory, that alternative structure was modelled. Any alternative models were evaluated against the original CFA and ESEM by comparing their fit statistics and parameter estimates.

Fit indices for invariance testing. For invariance testing, the chi-square difference test has been the most common way of comparing alternative models (Cheung & Rensvold, 2002; Meade et al., 2008; Sass, Schmitt, & Marsh, 2014). However, as with the chi-square test of model fit, this test may be limited by sample size, model complexity, non-normal data, and model misspecification (Cheung &

Rensvold, 2002; Meade et al., 2008; Sass et al., 2014). Similar to tests of model fit, many researchers now use alternative fit indices which have less sensitivity to these issues, as the preferred method for comparing alternative models (Meade et al., 2008; Sass et al., 2014). Of all the fit indices, research indicates that the change in CFI performs best (Chen, 2007; Meade et al., 2008). In addition, change in TLI and RMSEA are also seen as important in evaluating invariance because they contain a correction for parsimony (Marsh, Lüdtke, et al., 2010; Marsh et al., 2009). However, as a result of the correction for parsimony, it is also possible for a more constrained model to result in a better fit than a less constrained model. Following Chen (2007) and Sass, Schmitt and Marsh (2014), and assuming the fit of the best-fitting model is acceptable, there was support in this study for the more constrained (and more parsimonious) model if the CFI and TLI decreased by less than .010, and the RMSEA increased by less than .015 (Chen, 2007; Cheung & Rensvold, 2002). Although Meade et al. (2008) have suggested much stricter cut-off recommendations based on the results of a Monte Carlo simulation, it has been suggested that their simulation parameters were too strict for actual models, potentially resulting in overly conservative cut-off values (Little, 2013). For completeness, the Satorra-Bentler scaled chi square difference test (Satorra & Bentler, 2001, 2010) also is reported in Appendix P. If absolute invariance was not established for a measure, partial invariance was considered acceptable provided a majority of the items for each scale were invariant (Byrne, Shavelson, & Muthen, 1989; Steenkamp & Baumgartner, 1998; Vandenberg & Lance, 2000; cf. Steinmetz, 2013). Partial invariance was approached by releasing constraints one item at a time based on the highest modification index (Yoon & Kim, 2014). In addition, because the primary analyses in Study 2 are not focused on mean comparisons, metric invariance was the primary focus of the invariance testing and failure to achieve scalar invariance was not considered essential.

Validity. Campbell and Fiske's (1959) MTMM matrix approach is now widely used as a method for evaluating a scale's convergent and discriminant validity (Marsh, Ellis, et al., 2005; Marsh, Martin, & Jackson, 2010). In its original design, the MTMM matrix incorporated several constructs (called *traits*) by several

methods or approaches (e.g., self-report, direct observation, performance; Campbell & Fiske, 1959). This design has since been extended to allow multiple time points in longitudinal data to act as the multiple methods in an MTMM analysis (Campbell & O'Connell, 1967; Marsh, Martin, et al., 2010). Moreover, where the traits in the MTMM matrix are based on factor analytic models of latent factors, the correlations will be purged of measurement error (Marsh, Martin, et al., 2010).

Approaching an MTMM analysis in this way allows for an evaluation of the temporal stability of the latent factors, as well the convergent and discriminant validity of the responses for the scales (Marsh, Ellis, et al., 2005). The correlations between responses to the same scale measured on different occasions, referred to as *monotrait-heteromethod* (MTHM) correlations or *convergent validities*, represent test-retest stabilities of the latent factors (Marsh, Ellis, et al., 2005). As the two measures being correlated are of the same scale, high correlations support convergent validity. The correlations between responses for different scales measured at the same time (referred to as *heterotrait-monomethod* (HTMM) correlations), and the correlations between responses for different scales measured at different times (referred to as *heterotrait-heteromethod* (HTHM) correlations), should be lower than the convergent validities to support discriminant validity. Moreover, the HTMM correlations should not exceed the highest convergent validity.

In this study, all of the scales formed an MTMM matrix. The MTMM correlations were based on factor scores derived from the preferred model for the latent structure for each instrument following the results of the factor structure analysis. The MTMM matrix was produced in R using the psych package (see Appendix L for more information).

Summary

This section described the methodology and procedures specific to the psychometric analysis carried out in this study. The methodology and procedures outlined in this section, together with the details provided in Chapter Four, demonstrate the rigorous approach that was taken to test the proposed research

hypotheses for this study. The following section presents the results of the analyses undertaken in connection with those research hypotheses.

Results

Assessing the psychometric properties of the measurement instruments through the application of advanced statistical analyses, generated a considerable amount of quantitative data. In order to assist with the consideration of the data, Table 5.3 first provides a visual overview of these results for each instrument and its scales. It indicates for each instrument or scale (as relevant): (a) whether the threshold omega estimate of .70 for internal consistency reliability was achieved; (b) the preferred factor model that satisfied the fit criteria for the instrument's factor structure; (c) whether longitudinal scalar invariance was achieved or if not, the level that was achieved; and (d) whether convergent and discriminant validity were achieved. Following this overview, detailed results are presented in a series of tables organised by measure, except for the unidimensional measures, which are presented together. The tables set out the results for the evaluation of internal consistency reliability,¹⁹ factor structure,²⁰ and invariance.²¹ These within-construct analyses are presented first, followed by the results of the MTMM analysis for all scales and the between-construct analyses. Presenting the results in this way is consistent with the construct validation process, which emphasises a preliminary focus on within-construct considerations before moving to between-construct analyses (see Marsh, 1990b, 1993; Marsh, Ellis, et al., 2005). All of the psychometric analyses are then briefly summarised, with specific comments addressing any important findings.

¹⁹ The tables in this section present the omega coefficient (and its 95% confidence interval) to assess internal consistency for each scale. A table of Cronbach's alpha coefficients for all scales is included in Appendix M.

²⁰ The tables in this section present the CFI, TLI, and RMSEA (and its 90% confidence interval) to assess fit of the factor analytic models. Results for the chi square test statistic, degrees of freedom, and related *p*-values in connection with all factor analyses are presented in Appendix O.

²¹ The tables in this section present the CFI, TLI, and RMSEA to assess measurement invariance over time. Results for the Satorra-Bentler scaled chi square difference tests, and related *p*-values, are presented in Appendix P.

Table 5.3
 Overview of Results for Psychometric Analyses of Measurement Instruments and Scales

Scale	Internal Consistency	Factor Analysis	Invariance	Construct Validity	
				Convergent	Discriminant
Children's Hope Scale		2-fac CFA-M	Pt Scalar		
Agency	x			●	x
Pathways Thinking	●				
Pathways Thinking (revised)	●			●	x
Hope	●				
Life Orientation Test, Revised		2-fac CFA	●		
Optimism	●			●	●
Pessimism ^a	●			●	●
Total Optimism ^a	x				
Adolescent Self-Regulatory Inventory		x			
Long-Term Self-Regulation ^a	●				
Short-Term Self-Regulation ^a	●				
Total Self-Regulation	●				
Adolescent Self-Regulatory Inventory, Revised		3-fac ESEM	●		
Focus	●			●	●
Goal Self-Regulation	●			●	●
Emotion Self-Regulation	x			●	●
Short Grit Scale		2-fac ESEM	●		
Consistency of Interest ^a	x			●	●
Perseverance of Effort	●			●	●
Total Grit ^a	x				
Motivation and Engagement Scale - Short		3-fac ESEM	●		
Booster Behaviours	●			●	●
Booster Thoughts	●			●	●
Mufflers	x				
Guzzlers	●				
Hampering ^a	x			●	●
Life Resilience Scale		1-fac CFA	●		
Life Resilience	●			●	●
Academic Resilience Scale		1-fac CFA	●		
Academic Resilience	●			●	●
Satisfaction with Life Scale		1-fac CFA	Pt Scalar		
Life Satisfaction	●			●	●
Warwick-Edinburgh Mental Well-Being Scale		1-fac CFA	●		
Wellbeing	●			●	●

Note. For Internal Consistency, ● indicates the scale demonstrated an omega estimate > .70 or satisfactory inter-item correlations, and x indicates that it did not achieve either threshold. For Factor Analysis, the preferred model that satisfied the fit criteria is shown, and x indicates that no model was found that evidenced acceptable fit statistics. For Invariance, ● indicates the preferred factor model evidenced scalar invariance (otherwise, the level of invariance achieved is specified). For Convergent Validity, ● indicates that the scale demonstrated a significant monotrait-heteromethod correlation (i.e. convergent validity). For Discriminant Validity, ● indicates that the heterotrait-monomethod correlations for the scale at T₁ (pre-test) and T₂ (immediate post-test) were less than the highest convergent validity for all scales ($r = .80$), and x indicates that at least one such correlation for the scale exceeded .80. A blank indicates that the test was not relevant for that instrument or scale (e.g., internal consistency is applied at the scale level and factor analysis is applied at the item level).
^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

continues

Table 5.3 (continued)
Overview of Results for Psychometric Analyses of Measurement Instruments and Scales

Scale	Internal Consistency	Factor Analysis	Invariance	Construct Validity	
				Convergent	Discriminant
Self-Description Questionnaire II-Short		10-fac ESEM	●		
Physical Abilities SC ^a	●	/ 1-fac CFA		●	●
Physical Appearance SC	●			●	●
Opposite-Sex Relationships SC ^a	●			●	●
Same-Sex Relationships SC ^a	●			●	●
Parent Relationships SC ^a	●			●	●
Honesty-Trustworthiness SC ^a	●			●	●
Emotional Stability SC ^a	●			●	●
Math SC ^a	●			●	●
Verbal SC ^a	●			●	●
School SC ^a	●			●	●
General Self-Esteem ^a	●			●	x
Review of Personal Effectiveness and Locus of Control		14-fac CFA	●		
Self-Confidence	●			●	x
Self-Efficacy	●			●	x
Stress Management	●			●	x
Open Thinking	●			●	x
Social Effectiveness	●			●	●
Cooperative Teamwork	●			●	x
Leadership Ability	●			●	●
Time Efficiency	●			●	●
Quality Seeking	●			●	x
Coping with Change	●			●	x
Active Involvement	●			●	●
Overall Effectiveness	●			●	●
Internal Locus of Control	●			●	x
External Locus of Control ^a	●			●	●

Note. For Internal Consistency, ● indicates the scale demonstrated an omega estimate > .70 or satisfactory inter-item correlations, and x indicates that it did not achieve either threshold. For Factor Analysis, the preferred model that satisfied the fit criteria is shown, and x indicates that no model was found that evidenced acceptable fit statistics. For Invariance, ● indicates the preferred factor model evidenced scalar invariance (otherwise, the level of invariance achieved is specified). For Convergent Validity, ● indicates that the scale demonstrated a significant monotrait-heteromethod correlation (i.e. convergent validity). For Discriminant Validity, ● indicates that the heterotrait-monomethod correlations for the scale at T₁ (pre-test) and T₂ (immediate post-test) were less than the highest convergent validity for all scales ($r = .80$), and x indicates that at least one such correlation for the scale exceeded .80. A blank indicates that the test was not relevant for that instrument or scale (e.g., internal consistency is applied at the scale level and factor analysis^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

Children's Hope Scale (CHS)

Results of Research Hypothesis 1.1: Internal consistency reliability of

CHS. Given the outstanding question of whether the CHS is a one- or two-factor measure, Table 5.4 sets out the omega estimates for each of the Agency and Pathways Thinking scales, as well as the CHS as a whole. As reported in Table 5.4, the reliability estimates for both scales (including their 95% confidence intervals)

are below the proposed threshold level of .70, however, the reliability estimate for the CHS as a whole demonstrates an acceptable omega coefficient of .70, 95% CI [.67, .73].

Table 5.4

Reliability Estimates (Omega Coefficients) for the CHS and its Scales using Long Form Data

CHS Scale	Omega Coefficient	95% CI
Agency (3 items; $n = 1,059$)	.59	.54 / .63
Pathways (3 items; $n = 1,061$)	.61	.56 / .65
Hope (6 items; $n = 1,052$)	.70	.67 / .73

Note. CHS = Children's Hope Scale; CI = confidence interval; n = number of observations; Pathways = Pathways Thinking.

As a result of the low estimates, the inter-item correlations were also considered, and these are set out in Table 5.5. All of the items in the CHS have good inter-item correlations between .24 and .43 (including across scales), with an average inter-item correlation of .29. It is also worth noting that two of the Agency items: Ag1 (“I think I am doing pretty well”) and Ag2 (“I am doing just as well as other kids my age”) have a much higher inter-item correlation than either of those items has with the third Agency item: Ag3 (“I think the things I have done in the past will help me in the future”). The third Agency item has a similar inter-item correlation with the other Agency items as it has with the Pathways Thinking items. Retesting the Pathways Thinking scale including Ag3 resulted in an omega coefficient of .63, 95% CI [.58, .67], which is better than the omega coefficient for the original scale, but still below the .70 threshold.

Table 5.5
Inter-Item Correlations for the CHS Items using Long Form Data

CHS Items	Ag1	Ag2	Ag3	Pth1	Pth2	Pth3
Ag1	1.00					
Ag2	.43	1.00				
Ag3	.24	.27	1.00			
Pth1	.26	.24	.26	1.00		
Pth2	.25	.29	.26	.33	1.00	
Pth3	.25	.26	.26	.31	.37	1.00

Note. CHS = Children's Hope Scale; Ag = Agency; Pth = Pathways Thinking. Inter-item correlations within the same scale are highlighted grey.

Accordingly, Research Hypothesis 1.1 was supported only for the CHS as a whole. Although the reliability estimate was on the low end of the threshold, this lower result may reflect the conceptual heterogeneity of the items rather than low reliability. While Research Hypothesis 1.1 was not supported for the Agency and Pathways Thinking scales, the reliability estimate for the Pathways Thinking scale was above .60 and the inter-item correlations among the Pathways Thinking items were stronger than any of the inter-item correlations between Pathways Thinking items and Agency items. Internal consistency reliability for the Agency scale was the weakest and requires additional analysis. These results are given further consideration in the factor structure analysis detailed below.

Results of Research Hypothesis 1.2: Factor analysis for CHS. Due to the debate in connection with the CHS factor structure, a number of models were tested in the CHS factor analysis. The fit statistics for these models are set out below in Table 5.6.

Table 5.6
Fit Statistics for the Hypothesised CHS Factor Models using Long Form Data

Model	CFI	TLI	RMSEA	90% CI
One-Factor CFA	.937	.896	.062	.045 / .080
Two-Factor CFA	.983	.968	.035	.013 / .056
Two-Factor ESEM	.999	.997	.010	.000 / .048
Two-Factor CFA-M	1.000	.101	.000	.000 / .018

Note. CHS = Children's Hope Scale; CFI = comparative fit index; TLI = Tucker-Lewis fit index; RMSEA = root mean square error of approximation; CI = confidence interval; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model; M = modified. A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value can be found in Appendix O. Number of observations = 1,068.

The first step in the factor analysis, was to use a CFA to evaluate the a priori single factor structure for the CHS. With fit indices just at or below the acceptable thresholds, as indicated in Table 5.6, this model was not a strong fit to the data. The standardised factor loadings were significant, ranging from .47 to .57 (mean loading = .54, median loading = .54).

A two-factor solution, which has been strongly supported in the literature, was also tested. The two-factor CFA exhibited excellent fit statistics. The standardised factor loadings were significant and similar in size to those in the one-factor model, ranging from .47 (Ag3) to .64 (Ag2) (mean loading = .58, median loading = .60; see Table 5.7 for the full set of factor loadings). However, the correlation between the factors was high ($r = .76, p < .01$). The modification indices indicated that model fit would be improved by an additional path between Ag3 and the Pathways Thinking factor, as well as a correlated uniqueness between Ag1 and Ag2. A study using this instrument with Mexican-American youth also found Ag3 to exhibit a stronger association with the Pathways Thinking factor than with the Agency factor (Edwards, Ong, & Lopez, 2007). In their analysis, it was suggested that this item may have been interpreted by their sample as having more to do with routes toward desired goals, than energy or determination to move toward those goals. It is possible that the same is true for our sample, particularly given the focus of the THP program on setting action steps toward goals. In that study, the researchers chose to use a model in which Ag3 was moved to the Pathways

Thinking factor. There was no theoretical justification for a correlated uniqueness between Ag1 and Ag2 other than somewhat similar item wording.

Table 5.7
Standardised Factor Loadings for the CHS Two-Factor CFA and Two-Factor ESEM

CHS Items	Agency Factor Loadings			Pathways Factor Loadings		
	CFA	ESEM	CFA-M	CFA	ESEM	CFA-M
Ag1	.61*	.59	.63*		.02	
Ag2	.64*	.77*	.67*		-.09	
Ag3	.47*	.20*	.17*		.30*	.33*
Pth1		.07		.54*	.47*	.54*
Pth2		.04		.62*	.59*	.61*
Pth3		.01		.59*	.59*	.59*

Note. CHS = Children's Hope Scale; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model; M = modified; Ag = Agency; Pathways or Pth = Pathways Thinking. Target loadings highlighted grey.

Before modifying the two-factor CFA, a two-factor ESEM was modelled, which would allow Ag3 to cross-load onto the Pathways Thinking factor. The two-factor ESEM had excellent fit statistics with the CFI and TLI both close to 1. In addition, the correlation in the ESEM between the Agency and Pathways Thinking factors was lower than in the two-factor CFA ($r = .67, p < .01$), indicating that the restriction in the CFA on cross-loadings likely biased the correlation between the factors. The standardised factor loadings for this model are set out in Table 5.7. The standardised target loadings were wider ranging than in the two-factor CFA, going from .20 (Ag3) to .77 (Ag2) (mean target loading = .55, median target loading = .59). Moreover, the target loading for Ag1 was not significant, and Ag3 had a stronger standardised cross-loading on the Pathways Thinking factor (.32, $p < .05$) than the target loading on its own latent factor, which, though small, was also significant (.20, $p < .05$).

Given the results of the ESEM, a modified two-factor CFA was modelled (to be referred to in the balance of this chapter as the “two-factor CFA-M”). While Edwards et al. (2007) chose to move Ag3 from the Agency factor to the Pathways Thinking factor, doing so leaves the Agency factor with only two items. A minimum of three items is generally recommended for each latent factor in a multi-factor model, in order to avoid issues with model identification (Brown, 2015).

Furthermore, the ESEM demonstrated Ag3 to maintain a significant relationship with the Agency factor despite being allowed to cross-load onto the Pathways Thinking factor. Accordingly, a two-factor CFA was modelled which allowed Ag3 to load onto both the Agency and Pathways Thinking factors. This model had the best fit statistics of all the models, as set out in Table 5.6. The correlation between the factors also was lower than the two-factor CFA ($r = .68, p < .01$), and the standardised factor loadings were all significant, ranging from .17 (Ag3 on Agency) to .67 (Ag2; mean loading = .51, median loading = .59; see Table 5.7 for the full set of standardised factor loadings). With the best fit statistics, acceptable factor loadings, and only a slightly higher factor correlation than the ESEM, the two-factor CFA-M was selected as the preferred model, being more parsimonious than the ESEM. On this basis, Research Hypothesis 1.2 was deemed satisfied for the CHS.

Results of Research Hypothesis 1.3: Invariance analysis for CHS.

Having selected the two-factor CFA-M as the optimal measurement model for the CHS data, the longitudinal invariance of that data was tested using the selected model with the CHS data collected at T1 and T2. The results of these invariance tests are set out in Table 5.8.

Table 5.8
Longitudinal Invariance Models Based on CHS Two-Factor CFA-M at T1 and T2: Change in Fit Statistics

Models	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
Configural (1)	.992		.987		.017	
Metric (2)	.983	.009	.975	.012	.023	.006
Scalar (3a)	.954	.029	.938	.037	.037	.014
Partial Scalar (3b)	.979	.004	.972	.003	.025	.002

Note. CHS = Children's Hope Scale; CFA-M = confirmatory factor analysis, modified; T1 = pre-test; T2 = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). For the Partial Scalar (3b) model, the intercept for Pth3 (being the third item on the Pathways Thinking factor) was free. A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations = 350.

The configural model evidenced excellent fit, demonstrating that the two-factor CFA-M for the CHS data is sustainable at both timepoints. In addition, the changes in fit statistics from the configural model to the metric model were within the acceptable range, suggesting that the CHS items evidence equivalent relationships to their latent factor over time. However, the changes in fit statistics from the metric model to the scalar model did not demonstrate invariance. The modification indices suggested that the intercept of Pth3 was the largest source of model misfit. Accordingly, a partial scalar model was run with the Pth3 intercept free. The partial scalar model had a better fit than the full scalar model, and the changes in fit statistics from the metric model to the partial scalar model were within the acceptable range. These results suggest that for the Agency factor mean differences over time can be taken to reflect all mean differences in the shared variance of the Agency items, and for the Pathways Thinking factor, the same can be said for all but one of the items for that factor. Accordingly, it is argued that any analysis of mean change over time can be attributed to true change in the constructs of Agency and Pathways Thinking, each as measured by the CHS.

Conclusion. Research Hypothesis 1.1 was only supported for the CHS as a whole, with reliability estimates for each scale below the .70 threshold, consistent

with some of the previous research on this measure with adolescent data. However, given the small number of items on each factor, this result is not conclusive. As an alternative assessment, the inter-item correlations were found acceptable, but did reveal a potential issue in connection with the Agency items, with two being more strongly related to each other and the third seeming to be more aligned with the Pathways Thinking scale. The factor analysis was consistent with the internal consistency analysis, revealing the two-factor CFA-M to be the preferred model, with Ag3 loading onto both factors. This model had excellent fit statistics, and although both the two-factor CFA and two-factor ESEM also had excellent fit statistics, the two-factor CFA-M provided the best model as it reflected the relationship between Ag3 and the Pathways Thinking scale, while being more parsimonious than the ESEM. Permitting Ag3 to load onto both factors was considered preferable to moving Ag3 to the Pathways Thinking scale in order to stay as true to the original model as possible and avoid creating any issues for model identification. The two-factor CFA-M was found, on balance, to be invariant over time, with Research Hypothesis 1.3 supported only for a partial scalar model.

Life Orientation Test-Revised (LOT-R)

Results of Research Hypothesis 1.1: Internal consistency reliability of LOT-R. Given the outstanding question of whether the LOT-R is a one- or two-factor measure, Table 5.9 sets out the omega estimates for each of the Optimism and Pessimism scales, as well as the LOT-R scale as a whole. As reported in Table 5.9, the reliability estimates for both scales, as well as the complete scale, were all below the proposed threshold level of .70. The reliability estimate for the Pessimism scale is closest to the threshold with a 95% confidence interval that includes the threshold level. These results are consistent with the lower alpha coefficients found for adolescents in the meta-analysis conducted by Vassar and Bradley (2010).

Table 5.9
Reliability Estimates (Omega Coefficients) for the LOT-R and its Scales using Long Form Data

LOT-R Scale	Omega Coefficient	95% CI
Optimism (3 items; $n = 1,041$)	.54	.49 / .60
Pessimism ^a (3 items; $n = 1,050$)	.66	.62 / .71
LOT (6 items; $n = 1,027$)	.61	.56 / .65

Note. LOT-R = Life Orientation Test, Revised; CI = confidence interval; n = number of observations.

^a The items for this scale are all negatively worded and were reverse-scored prior to analysis.

As a result of the low estimates, the inter-item correlations were also considered, and these are set out in Table 5.10. All of the items within each scale have good inter-item correlations between .23 and .43, with an average inter-item correlation for Optimism of .28 and Pessimism of .39, lending support to Research Hypothesis 1.1 for the LOT-R scales considered separately. However, the inter-item correlations among the Optimism and Pessimism items are all below .20, suggesting that they are not well-related to each other and, therefore, may not be suitable for measuring as a single construct. Further assessment will be made in connection with the factor analysis.

Table 5.10
Inter-Item Correlations for the LOT-R Items using Long Form Data

LOT-R items	Opt1	Opt2	Opt3	Ps1 ^a	Ps2 ^a	Ps3 ^a
Opt1	1.00					
Opt2	.31	1.00				
Opt3	.29	.23	1.00			
Ps1 ^a	.12	.15	.05	1.00		
Ps2 ^a	.09	.17	.10	.39	1.00	
Ps3 ^a	.09	.13	.07	.34	.43	1.00

Note. LOT-R = Life Orientation Test, Revised; Opt = Optimism; Ps = Pessimism. Inter-item correlations within the same scale are highlighted grey.

^a This item is negatively worded and was reverse-scored prior to analysis.

Results of Research Hypothesis 1.2: Factor analysis for LOT-R. Due to the debate in connection with the LOT-R factor structure, a number of models were tested in the LOT-R factor analysis. The fit statistics for these models are set out below in Table 5.11.

Table 5.11
Fit Statistics for the Hypothesised LOT-R Factor Models using Long Form Data

Model	CFI	TLI	RMSEA	90% CI
One-Factor CFA	.690	.483	.114	.097 / .131
Two-Factor CFA	.998	.997	.009	.000 / .038
Bifactor ("faking good")	1.000	1.018	.000	.000 / .026
Two-Factor ESEM	1.000	1.012	.000	.000 / .039

Note. LOT-R = Life Orientation Test, Revised; CFI = comparative fit index, TLI = Tucker-Lewis fit index, RMSEA = root mean square error of approximation; CI = confidence interval; CFA = confirmatory factor analysis; Bifactor ("faking good") = a bifactor model with a general factor for all LOT-R items and a specific factor for the positively-worded items; ESEM = exploratory structural equation model. A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF) and *p*-value can be found in Appendix O. Number of observations = 1,067.

The first step in the factor analysis was to use a CFA to evaluate the a priori single factor structure for the LOT-R. With fit indices well below the thresholds, as indicated in Table 5.11, this model was a poor fit to the data. Next a two-factor CFA was tested. This model resulted in strong fit statistics, demonstrating an excellent fit to the data, with a low correlation between the Optimism and Pessimism factors ($r = .32, p < .01$). The standardised factor loadings for this model, all within an acceptable range, are set out in Table 5.12 (mean loading = .58, median loading = .58).

Table 5.12
Standardised Factor Loadings for the LOT-R Two-Factor CFA and Two-Factor ESEM

LOT-R Items	Opt Factor Loadings		Pess Factor Loadings	
	CFA	ESEM	CFA	ESEM
Opt1	.59*	.58*		-.05
Opt2	.54*	.43*		.09*
Opt3	.45*	.39*		-.02
Ps1 ^a		.03	.56*	.55*
Ps2 ^a		.00	.71*	.71*
Ps3 ^a		-.02	.61*	.61*

Note. LOT-R = Life Orientation Test, Revised; Opt = Optimism; Pess and Ps = Pessimism; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model. Target loadings highlighted grey.

*indicates significant p -value < .05.

^aThis item is negatively worded and was reverse-scored prior to analysis.

A bifactor model was also tested, following the model proposed by Vautier et al. (2003). While this model had superior fit statistics, the factor loadings for the Optimism items on the general factor (being the substantive factor) were not satisfactory, with standardised loadings for those items ranging from .12 (Opt3) to .24 (Opt2), all well below the cut-off of .30. Finally, a two-factor ESEM was modelled. This model had the best fit to the data, but not substantially better than the two-factor CFA. The factor loadings (mean target loading = .55, median target loading = .57) and factor correlation ($r = .31$, $p < .01$) were also similar to the two-factor CFA (for a full set of factor loadings see Table 5.12). Accordingly, the two-factor CFA was accepted as the preferred model on the basis that the fit statistics and parameter estimates for the ESEM did not outweigh the greater parsimony of the CFA.

Results of Research Hypothesis 1.3: Invariance analysis for LOT-R.

Having selected the two-factor CFA as the optimal measurement model for the LOT-R data, the longitudinal invariance of that data was tested using the selected model. The results of these invariance tests are set out in Table 5.13.

Table 5.13

Longitudinal Invariance Models Based on LOT-R Two-Factor CFA at T1 and T2: Change in Fit Statistics

Model	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
Configural (1)	.993		.989		.014	
Metric (2)	1.000	-.007	1.000	-.011	.000	-.014
Scalar (3)	.996	.004	.995	.005	.009	.009

Note. LOT-R = Life Orientation Test, Revised; CFA = confirmatory factor analysis; T1 = pre-test; T2 = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations = 350.

The configural model evidenced excellent fit, demonstrating that the two-factor CFA for the LOT-R data is sustainable at both timepoints. In addition, the changes in fit statistics from the configural model to the metric model, and from the metric model to the scalar model, were each within the acceptable range, suggesting that the LOT-R items evidence equivalent relationships to their latent factor over time and their intercepts are also invariant over time. Accordingly, any analysis of mean change over time can be attributed to true change in the constructs of Optimism and Pessimism, each as measured by the LOT-R.

Conclusion. Research Hypothesis 1.1 was not supported for either the LOT-R as a whole or for its separate scales, based on the reliability estimates which were below the .70 threshold. As an alternative assessment, the inter-item correlations were found acceptable for the Optimism and Pessimism scales but revealed a lack of consistency among the items as a whole. This result does not lend support for a single factor model. Consistent with these results, the factor analysis found the two-factor CFA to be the preferred model with excellent fit statistics, appropriate factor loadings, and an acceptable correlation between the factors, giving support to Research Hypothesis 1.2. The two-factor CFA was found to be invariant over time, supporting Research Hypothesis 1.3.

Adolescent Self-Regulatory Inventory (ASRI)

Results of Research Hypothesis 1.1: Internal consistency reliability of

ASRI. Given the limited research on the ASRI, Table 5.14 sets out the omega estimates for both the Long-Term Self-Regulation (LT-SR) and Short-Term Self-Regulation (ST-SR) scales, as well as the measure as a whole (using only the scored items, see Moilanen, 2011). All scales reflect omega estimates, as well as 95% confidence intervals, above the threshold and near to or exceeding a very good rating.

Table 5.14

Reliability Estimates (Omega Coefficients) for the ASRI and its Scales using Long Form Data

ASRI Scale	Omega Coefficient	95% CI
LT-SR ^a (14 items; $n = 1,015$)	.80	.78 / .82
ST-SR ^a (15 items; $n = 1,015$)	.75	.72 / .77
ASRI (29 items; $n = 977$)	.87	.85 / .88

Note. ASRI = Adolescent Self-Regulatory Inventory; CI = confidence interval; LT-SR = Long-Term Self-Regulation; ST-SR = Short-Term Self-Regulation; n = number of observations.

^a Some of the items for this scale are negatively worded and were reverse-scored prior to analysis.

Results of Research Hypothesis 1.2: Factor analysis for ASRI. The a priori hypothesised model for the ASRI was tested first with a two-factor CFA. Owing to the poor fit of this model (see Table 5.15) and high factor correlation ($r = .91, p < .01$), a two-factor ESEM was modelled. This model also demonstrated a poor fit to the data, as set out in Table 5.15. However, the correlation between factors was substantially reduced ($r = .35, p < .01$).

Table 5.15
Fit Statistics for the Hypothesised ASRI Factor Models using Long Form Data

Model	CFI	TLI	RMSEA	90% CI
One-Factor CFA	.706	.683	.060	.057 / .062
Two-Factor CFA	.712	.689	.059	.056 / .062
Two-Factor ESEM	.855	.831	.044	.041 / .047

Note. ASRI = Adolescent Self-Regulatory Inventory; CFI = comparative fit index; TLI = Tucker-Lewis fit index; RMSEA = root mean square error of approximation; CI = confidence interval; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model. A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value can be found in Appendix O. Number of observations = 1,068.

For both the CFA and ESEM, a number of the standardised factor loadings were close to or below the threshold level (CFA mean loading = .39, median loading = .36; ESEM mean loading = .34, median loading = .41). In addition, there were many significant cross-loadings in the ESEM. Standardised loadings for these models are set out in Table 5.16. Given the excellent omega coefficient for the total scale, a one-factor CFA was also tested. This model had the poorest fit of the three models (see Table 5.15).

Table 5.16
Standardised Factor Loadings for the ASRI Two-Factor CFA and Two-Factor ESEM

ASRI Items	<u>LT Factor Loadings</u>		<u>ST Factor Loadings</u>	
	CFA	ESEM	CFA	ESEM
LT1	.52*	.41*		.01
LT2	.57*	.45*		.35*
LT4 ^a	.34*	.10		.35*
LT5 ^a	.19*	-.05		.41*
LT6	.51*	.48*		.11
LT7	.64*	.44*		.36*
LT8	.35*	.34*		.11
LT9	.59*	.50*		-.04
LT10	.29*	.30*		-.01
LT11	.62*	.52*		.26*
LT12	.38*	.42*		-.12*
LT13	.64*	.58*		-.01
LT14	.64*	.55*		.05
LT16	.48*	.39*		.12*
ST2		.38*	.31*	-.06
ST4 ^a		-.03	.35*	.61*
ST5		.30*	.28*	.02
ST6		.51*	.56*	.03
ST7		.35*	.25*	-.10
ST8		.50*	.58*	.19*
ST9 ^a		.00	.36*	.49*
ST10 ^a		.02	.40*	.53*
ST11 ^a		.07*	.50*	.65*
ST12		.48*	.62*	.25*
ST13 ^a		.01	.41*	.52*
ST14 ^a		-.05	.24*	.39*
ST15 ^a		.00	.28*	.34*
ST16		.46*	.51*	.10*
ST19 ^a		-.06	.26*	.44*

Note. ASRI = Adolescent Self-Regulatory Inventory; LT = Long-Term Self-Regulation; ST = Short-Term Self-Regulation; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model. Target loadings highlighted grey.

*indicates significant p -value < .05.

^aThis item is negatively worded and was reverse-scored prior to analysis.

Given the lack of research on the measure and poor fit of the hypothesised models to the research data, consideration was given to whether a subset of the items might be used to form a new scale or scales more closely connected to the THP program design and aims. In particular, the self-regulation at the heart of the THP program is less focused on the long-term and short-term self-regulation distinctions and more on the cognitive and emotional facets of self-regulatory behaviour. Particular attention in the THP program is placed on improving cognitive self-regulation, especially self-regulatory behaviours relevant to goal pursuit. Accordingly, items were selected a priori that were considered to be predicted by one of three proposed latent constructs: Focus (the ability to self-regulate despite distractions or other difficulties, e.g., “I can stay focused on my work, even when it’s dull”), Goal Self-Regulation (proactive behaviour, problem-solving, and persistence in the pursuit of goals, e.g., “If something isn’t going according to my plans, I can change my actions to try and reach my goal”), and Emotion Self-Regulation (the ability to self-regulate one’s emotions or to self-regulate one’s behaviour in the face of emotions, e.g., “I can calm myself down, especially when I’m excited or all wound up”). Five items were chosen for each latent factor, and the full set of items is set out in Appendix A. Figure 5.1 illustrates the factor structure for the revised ASRI (to be referred to as ASRI-R).

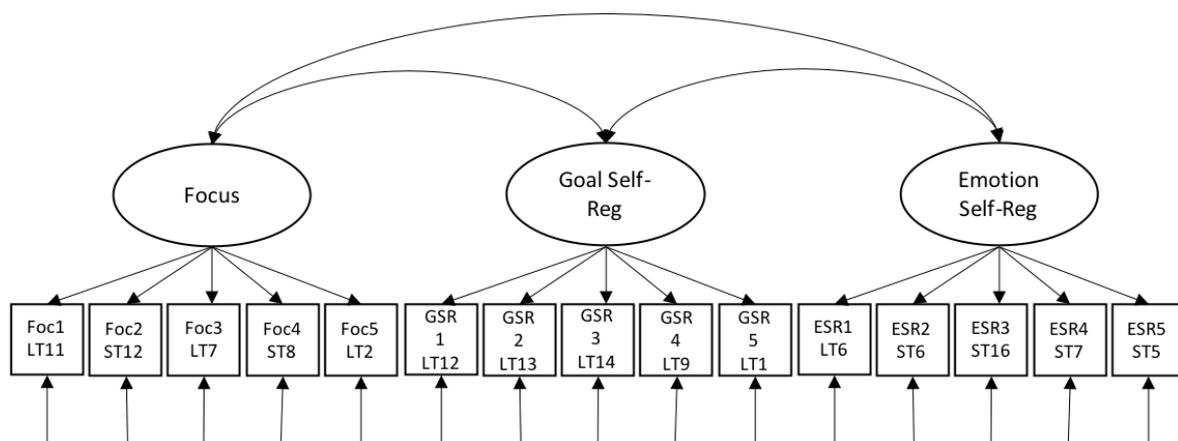


Figure 5.1. Revised ASRI three-factor structure (ASRI-R).

Note. Self-Reg = Self-Regulation; Foc = Focus; LT = Long-Term; GSR = Goal Self-Regulation; ESR = Emotion Self-Regulation; ST = Short-Term. The boxes include item numbers for the new ASRI-R on top and item numbers for the original ASRI underneath.

Both a CFA and an ESEM were modelled, and the results are set out in Table 5.17. The CFA had excellent fit statistics, and the factor loadings were all significant. The standardised loadings ranged from .34 (ESR5: “I can usually act normal around everybody if I am upset with someone” on Emotion Self-Regulation) to .73 (GSR2, “I can find a way to stick with my plans and goals, even when it's tough” on Goal Self-Regulation; mean loading .59, median loading = .62). However, correlations between the factors were quite high ranging from $r = .73, p < .01$ (Emotion Self-Regulation/Goal Self-Regulation) to $r = .76, p < .01$ (Goal Self-Regulation/Focus), with mean $r = .75$ and median $r = .74$.

Table 5.17

Fit Statistics for the Hypothesised ASRI-R Factor Models using Long Form Data

Model	CFI	TLI	RMSEA	90% CI
Three-Factor CFA	.970	.964	.029	.022 / .036
Three-Factor ESEM	.986	.977	.023	.014 / .032

Note. ASRI-R = Adolescent Self-Regulatory Inventory, Revised; CFI = comparative fit index; TLI = Tucker-Lewis fit index; RMSEA = root mean square error of approximation; CI = confidence interval; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model. A complete table including chi square test statistics (χ^2), degrees of freedom (df), scaling correlation factor (SCF), and p -value can be found in Appendix O. Number of observations = 1,068.

The ESEM demonstrated an improved fit to the data, and all three pairs of factor correlations were reduced (all three being the same, $r = .63, p < .01$). All target factor loadings were significant and standardised target loadings were all above the acceptable threshold, ranging from .38 (Foc4 “I can start a new task, even if I’m already tired” on Focus) to .71 (GSR2, as above), with mean loading = .55 and median loading = .56. The full set of factor loadings for both the CFA and the ESEM is set out in Table 5.18. While there were a few items with significant cross-loadings in the ESEM, the target loadings for those items were stronger than any cross-loadings. Accordingly, the ESEM was accepted as the preferred model.

Table 5.18
Standardised Factor Loadings for the ASRI-R Three-Factor CFA and Three-Factor ESEM

ASRI-R Items	Focus Factor Loadings		GSR Factor Loadings		ESR Factor Loadings	
	CFA	ESEM	CFA	ESEM	CFA	ESEM
Foc1	.66*	.62*		.05		.03
Foc2	.65*	.55*		.06		.08
Foc3	.69*	.66*		.04		.01
Foc4	.59*	.38*		.10		.16*
Foc5	.62*	.66*		.01		-.04
GSR1		-.15*	.46*	.58*		.02
GSR2		.04	.73*	.71*		.02
GSR3		.23*	.70*	.55*		-.03
GSR4		.03	.66*	.56*		-.01
GSR5		.12*	.56*	.47*		.12
ESR1		.09		-.14*	.60*	.68*
ESR2		.01		.30*	.64*	.38*
ESR3		.11		-.09	.60*	.61*
ESR4		-.18		.13	.35*	.41*
ESR5		-.04		-.02*	.34*	.40*

Note. ASRI-R = Adolescent Self-Regulatory Inventory, Revised; GSR = Goal Self-Regulation; ESR = Emotion Self-Regulation; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model; Foc = Focus. Target loadings highlighted grey.

*indicates significant p -value $<.05$.

Reliability of the data was retested using these new scales. Omega estimates for Focus and Goal Self-Regulation were acceptable ($\omega = .77$ and $.76$, respectively), while the omega estimate for Emotion Self-Regulation was below the threshold ($\omega = .63$, 95% CI [.59, .67]). An inter-item correlation matrix revealed that item ESR5 had correlations with both ESR3 and ESR4 that were below the .20 threshold but not without relationship (at .17 and .16, respectively). In addition, the inter-item correlations among the other ESR items were acceptable, ranging from .22 to .44. Therefore, notwithstanding the slightly weaker result for the internal consistency of the Emotion Self-Regulation scale, the decision was made to retain the Emotion Self-Regulation scale and each of its items. The inter-item correlation matrix also reflected reasonable inter-item correlations between items from different factors, reinforcing the appropriateness of an ESEM, which would allow the items to cross-

load. Accordingly, it was decided to proceed with this revised three-factor ASRI (ASRI-R) and to model the factor structure as an ESEM.

Results of Research Hypothesis 1.3: Invariance analysis for ASRI-R.

Having selected the three-factor ESEM as the optimal measurement model for the ASRI-R data, the longitudinal invariance of that data was tested using the selected model. The results of these invariance tests are set out in Table 5.19.

Table 5.19

Longitudinal Invariance Models Based on ASRI-R Three-Factor ESEM at T1 and T2: Change in Fit Statistics

Models	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
Configural (1)	.977		.969		.020	
Metric (2)	.976	.001	.971	-.002	.020	.000
Scalar (3)	.975	.001	.971	.000	.020	.000

Note. ASRI-R = Adolescent Self-Regulatory Inventory, Revised; ESEM = exploratory structural equation model; T1 = pre-test; T2 = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). A complete table including chi square test statistics (χ^2), degrees of freedom (df), scaling correlation factor (SCF), and p -value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations = 350.

The configural model evidenced excellent fit, demonstrating that the three-factor ESEM for the ASRI-R data is sustainable at both timepoints. In addition, the changes in fit statistics from the configural to the metric model, and from the metric to the scalar model, were all acceptable, suggesting that the ASRI-R items evidence equivalent relationships to their latent factor over time and their intercepts are invariant over time. Accordingly, any analysis of mean change over time can be attributed to true change in the newly constituted latent constructs measured by the ASRI-R.

Conclusion. While Research Hypothesis 1.1 was accepted for the a priori hypothesised scales of the ASRI, Research Hypothesis 1.2 was not supported for the two-factor structure with either a CFA or ESEM. As a consequence, the ASRI items were reviewed with the THP program in mind, as well as the theoretical classification of self-regulatory processes by reference to cognitions, emotions, and

behaviour (Moilanen, 2007). A revised measure was created (referred to as ASRI-R) using 15 items from the original ASRI that were considered to be predicted by three constructs: Focus (cognitions), Goal Self-Regulation (behaviour), and Emotion Self-Regulation (emotions). Research Hypothesis 1.1 was supported for two of these three ASRI-R scales. Despite some lower inter-item correlations among the items in the Emotion Self-Regulation scale, a decision was made to proceed with all three scales given the factor analysis was acceptable. The ESEM was accepted as the preferred model, and thus Research Hypothesis 1.2 was supported for the ASRI-R. Invariance testing over time was acceptable, with all models displaying satisfactory fit statistics and model comparisons evidencing changes in fit statistics all within acceptable limits. Accordingly, Research Hypothesis 1.3 also was supported for the ASRI-R.

Short Grit Scale (Grit-S)

Results of Research Hypothesis 1.1: Internal consistency reliability of Grit-S. Given the outstanding question of whether the Grit-S is a one- or two-factor measure, Table 5.20 sets out the omega estimates for each of the Consistency of Interest (COI) and Perseverance of Effort (POE) scales, as well as the Grit-S measured as a whole. As reported in Table 5.20, the reliability estimate for the POE scale is only just below the .70 threshold with a 95% confidence interval that includes the threshold level. However, both the COI scale and the Grit measure as a whole are below the threshold of .70. These results stand in contrast with previous findings in which the COI scale and overall Grit measure demonstrated good reliability, with the POE scale estimates being more variable (Duckworth & Quinn, 2009).

Table 5.20
Reliability Estimates (Omega Coefficients) for the Grit-S and its Scales using Long Form Data

Grit-S Scale	Omega Coefficient	95% CI
COI ^a (4 items; <i>n</i> = 1,050)	.60	.55 / .65
POE (4 items; <i>n</i> = 1,044)	.68	.65 / .71
Grit (8 items; <i>n</i> = 1,030)	.61	.56 / .67

Note. Grit-S = Short Grit Scale; CI = confidence interval; COI = Consistency of Interest; POE = Perseverance of Effort; *n* = number of observations.

^a The items for this scale are all negatively worded and were reverse-scored prior to analysis.

As a result of the low omega coefficient estimates, the inter-item correlations were also considered, and these are set out in Table 5.21. All of the items within the POE scale have good inter-item correlations between .31 and .38, with an average inter-item correlation of .35. Except for the correlation ($r = .16$) between COI1 (“I often set a goal but later choose to pursue a different one”) and COI4 (“I have difficulty maintaining my focus on projects that take more than a few months to complete”), the items within the COI scale also have good inter-item correlations ($r = .26-.32$), with an overall average correlation of $r = .27$. However, the inter-item correlations between the COI and POE items are all well below .20, with the exception of COI4 and POE1 (“I finish whatever I begin”), which has a correlation of $r = .23$. These correlations suggest that the two scales are not well related to each other and, therefore, may not be suitable for measuring as a single construct. Further assessment will be made in connection with the factor analysis.

Table 5.21
Inter-Item Correlations for the Grit-S Items using Long Form Data

Grit-S items	COI1 ^a	COI2 ^a	COI3 ^a	COI4 ^a	POE1	POE2	POE3	POE4
COI1 ^a	1.00							
COI2 ^a	.27	1.00						
COI3 ^a	.32	.32	1.00					
COI4 ^a	.16	.30	.26	1.00				
POE1	.14	.08	.14	.23	1.00			
POE2	.08	.04	.06	.09	.31	1.00		
POE3	.04	.04	.07	.12	.35	.38	1.00	
POE4	.06	.06	.08	.08	.36	.31	.38	1.00

Note. Grit-S = Short Grit Scale; COI = Consistency of Interest; POE = Perseverance of Effort. Inter-item correlations within the same scale are highlighted grey.

^aThis item is negatively worded and was reverse-scored prior to analysis.

Results of Research Hypothesis 1.2: Factor analysis for Grit-S. The first model evaluated in the factor analysis for the Grit-S was a one-factor CFA. With fit indices well below the thresholds, as indicated in Table 5.22, this model was a poor fit to the data.

Table 5.22
Fit Statistics for the Hypothesised Grit-S Factor Models using Long Form Data

Model	CFI	TLI	RMSEA	90% CI
One-Factor CFA	.641	.497	.106	.095 / .118
Two-Factor CFA	.964	.947	.035	.021 / .048
Two-Factor ESEM	.984	.965	.028	.008 / .045

Note. Grit-S = Short Grit Scale; CFI = comparative fit index; TLI = Tucker-Lewis fit index; RMSEA = root mean square error of approximation; CI = confidence interval; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model. A complete table including chi square test statistics (χ^2), degrees of freedom (df), scaling correlation factor (SCF), and p -value can be found in Appendix O. Number of observations = 1,068.

Next a two-factor CFA was tested. This model had mostly excellent fit statistics, with the TLI just below excellent. Factor loadings were all significant and acceptable (mean loading = .56, median loading = .57), and the correlation between factors was low ($r = .27$, $p < .01$; see Table 5.23 for the full set of standardised factor coefficients). Although previous research has found a second-order structure to evidence good fit, the low correlation between factors in the two-factor CFA does not support such a model. An ESEM was conducted, and this model demonstrated

the best fit, with all fit statistics in the excellent range (see Table 5.22). The standardised target factor loadings for this model were all significant and above the .30 threshold (mean loading = .56, median loading = .57), with non-target loadings that were non-significant or otherwise minor. These factor coefficients are reproduced in Table 5.23. Accordingly, the two-factor ESEM was selected as the preferred model.

Table 5.23

Standardised Factor Loadings for the Grit-S Two-Factor CFA and Two-Factor ESEM

Grit-S Items	<u>COI Factor Loadings</u>		<u>POE Factor Loadings</u>	
	CFA	ESEM	CFA	ESEM
COI1 ^a	.59*	.48*		.00
COI2 ^a	.47*	.58*		-.05
COI3 ^a	.56*	.59*		-.01
COI4 ^a	.46*	.43*		.10*
POE1		.12*	.59*	.55*
POE2		-.02	.55*	.56*
POE3		-.06*	.64*	.66*
POE4		-.02	.59*	.59*

Note. Grit-S = Short Grit Scale; COI = Consistency of Interest; POE = Perseverance of Effort; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model. Target loadings highlighted grey.

*indicates significant p -value < .05.

^aThis item is negatively worded and was reverse-scored prior to analysis.

Results of Research Hypothesis 1.3: Invariance analysis for Grit-S.

Having selected the two-factor ESEM as the optimal measurement model for the Grit-S data, the longitudinal invariance of that data was tested using the selected model. The results of these invariance tests are set out in Table 5.24.

Table 5.24
Longitudinal Invariance Models Based on Grit-S Two-Factor ESEM at T1 and T2: Change in Fit Statistics

Models	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
Configural (1)	.938		.905		.038	
Metric (2)	.929	.009	.906	-.001	.038	.000
Scalar (3)	.921	.008	.901	.005	.039	.001

Note. Grit-S = Short Grit Scale; ESEM = exploratory structural equation model; T1 = pre-test; T2 = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations = 350.

The configural model had acceptable fit statistics to support the model at both timepoints. In addition, the changes in fit statistics from the configural model to the metric model, and from the metric model to the scalar model, were each within the acceptable range, suggesting that the Grit-S items evidence equivalent relationships to their latent factor over time and their intercepts are also invariant over time. Accordingly, any analysis of mean change over time can be attributed to true change in the constructs of Consistency of Interest and Perseverance of Effort, each as measured by the Grit-S.

Conclusion. With reliability estimates below .70, Research Hypothesis 1.1 was not supported for either of the Grit-S scales or for the measure as a whole. As an alternative assessment, the inter-item correlations were found acceptable, with the exception of one item pair in the COI scale. The factor analysis revealed the two-factor ESEM to be the preferred model, giving support to Research Hypothesis 1.2. This model had good fit statistics, good factor loadings, and an acceptable inter-factor correlation. Testing the longitudinal invariance of the two-factor ESEM, the configural model was found to have acceptable fit statistics, and the changes in fit statistics across invariance models were within the acceptable range, providing support for Research Hypothesis 1.3.

Motivation and Engagement Scale, Short (MES-S)

Results of Research Hypothesis 1.1: Internal consistency reliability of

MES-S. Table 5.25 sets out the omega estimates for each of the MES-S scales. As reported in Table 5.25, Booster Behaviours demonstrates a good reliability estimate, and the reliability estimate for Booster Thoughts is close to the threshold, with the upper end of the 95% confidence interval being just off the threshold. However, the reliability estimates for both the Mufflers and Guzzlers scales are well below the threshold. This result is not surprising given the prior research indicating lower levels of shared variance between the first-order scales in the MES-HS, from which these scales were derived (see Martin, 2009).

Table 5.25

Reliability Estimates (Omega Coefficients) for the MES-S and its Scales using Long Form Data

MES-S Scale	Omega Coefficient	95% CI
Booster Behavs (3 items; $n = 1,048$)	.76	.73 / .79
Booster Thts (3 items; $n = 1,054$)	.65	.60 / .69
Mufflers ^a (3 items; $n = 1,056$)	.41	.31 / .51
Guzzlers ^a (2 items; $n = 1,061$)	.53	.47 / .59

Note. MES-S = Motivation and Engagement Scale-Short; CI = confidence interval; Behavs = Behaviours; Thts = Thoughts; n = number of observations.

^a The items for this scale are all negatively worded and were reverse-scored prior to analysis.

As a result of the low estimates, the inter-item correlations were also considered, and these are set out in Table 5.26. Each of the Booster Thoughts and Guzzlers scales has items with good inter-item correlations all ranging between .36 and .38. The Booster Behaviours scale has items with higher inter-item correlations, ranging between .42 and .58, indicating that these items may be more homogenous. However, the Booster Thoughts and Booster Behaviours items also have moderate-to-high inter-item correlations between the scales, ranging from .31 to .46. Moreover, the Guzzler items have moderate inter-item correlations with both the Booster Thoughts and Booster Behaviours items, ranging from .19 to .35. The Mufflers scale has items with poor inter-item correlations, ranging from .08 to .21 and an average inter-item correlation of .16. It is also worth noting that Mf3 (“I

don't think I have much control over how well I do in my schoolwork") has a higher inter-item correlation with all of the other items on the MES-S than with Mf1 ("I get quite anxious about schoolwork and tests") and higher inter-item correlations with a number of the other items on the MES-S than with Mf2 ("I mainly do my schoolwork to avoid failing or disapproval from parents or the teachers"). These results make the Mufflers scale difficult to support with these data. Further analysis of the measurement properties of the MES-S scales will be undertaken in connection with the factor analysis below.

Table 5.26

Inter-Item Correlations for the MES-S Items using Long Form Data

MES-S Items	BB1	BB2	BB3	BT1	BT2	BT3	Mf1	Mf2	Mf3	Gz1	Gz2
BB1	1.00										
BB2	.58	1.00									
BB3	.42	.47	1.00								
BT1	.31	.34	.37	1.00							
BT2	.34	.37	.39	.37	1.00						
BT3	.46	.34	.38	.38	.38	1.00					
Mf1 ^a	-.04	-.02	.03	.02	-.03	.05	1.00				
Mf2 ^a	-.08	-.07	-.08	-.08	-.08	-.08	.21	1.00			
Mf3 ^a	.15	.17	.21	.30	.13	.19	.08	.18	1.00		
Gz1 ^a	.27	.35	.26	.19	.23	.22	.08	.01	.19	1.00	
Gz2 ^a	.24	.30	.31	.31	.29	.28	.18	.08	.37	.36	1.00

Note. MES-S = Motivation and Engagement Scale-Short; BB = Booster Behaviours; BT = Booster Thoughts; Mf = Mufflers; Gz = Guzzlers. Inter-item correlations within the same scale are highlighted grey.

^aThis item is negatively worded and was reverse-scored prior to analysis.

Results of Research Hypothesis 1.2: Factor analysis for MES-S. The hypothesised model for the MES-S was tested first with a four-factor CFA. This model had poor fit (CFI = .892, TLI = .844, RMSEA = .065, 90% CI [.056, .073]). In addition, two of the three items for the Muffler scale had standardised loadings well below the threshold (at .20 and .21), putting the structural integrity of this scale in doubt. A significant and high factor correlation between the Booster Behaviours and Booster Thoughts scales ($r = .84$) also raised issues of multicollinearity. A one-

set, four-factor ESEM was modelled, however, convergence could not be obtained for this model. Based on these results, it was determined that the research data did not support a four-factor structure for the MES-S. As this measure was a modification of the original measure (having selected one item to represent each first-order factor), exploratory factor analysis (EFA) was used to ascertain if there was an alternative factor structure that better fit the data.

An EFA was run in Mplus on all 11 MES-S items with geomin rotation and MLR estimation. An oblique rotation was used as the motivation and engagement theory suggests the factors should be correlated. The EFA requested extraction of two, three, and four factors, based on a scree plot of the data. Consistent with the ESEM results, a four-factor solution did not converge, but results were obtained for two- and three-factor solutions. Results for these two solutions are set out below in Table 5.27. The three-factor solution had the best fit, with all fit indices in the acceptable range.

Table 5.27

Fit Statistics from EFA for the Hypothesised MES-S Factor Models using Long Form Data

Model	CFI	TLI	RMSEA	90% CI
Two-Factor Solution	.917	.865	.060	.051 / .069
Three-Factor Solution	.972	.938	.041	.030 / .052

Note. EFA = exploratory factor analysis; MES-S = Motivation and Engagement Scale - Short; CFI = comparative fit index; TLI = Tucker-Lewis fit index; RMSEA = root mean square error of approximation; CI = confidence interval. A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value can be found in Appendix O. Number of observations = 1,068.

Factor loadings for the three-factor EFA solution are set out in Table 5.28. The three Booster Behaviours items load onto the first factor. BB3 (“I persist at schoolwork even when it is challenging or difficult”) also cross-loads onto the second factor at the same level. The three Booster Thoughts items load onto the second factor. Two of these items cross-load onto the first factor with the loading for BT3 (“In my schoolwork, I am focused on learning and improving more than competing and being the best”) being slightly higher on the first factor than the second factor. The five negatively-worded items that make up the Mufflers and Guzzlers scales all load onto factor three. One of the Guzzlers items (Gz1: “In my schoolwork I sometimes reduce my chances of doing well (e.g., waste time, not

study, disrupt others, procrastinate)”) also loads onto the first factor more strongly than the third factor. Other cross-loadings are below the .30 threshold.

Table 5.28
Geomin Rotated Loadings for MES-S EFA with Three Factors

MES-S Items	Factor		
	1	2	3
BT1		.79	
BT2	.31	.32	
BT3	.34	.32	
BB1	.71		
BB2	.80		
BB3	.38	.38	
Mf1 ^a			.35
Mf2 ^a			.37
Mf3 ^a		.29	.41
Gz1 ^a	.40		.29
Gz2 ^a	.21	.20	.52

Note. MES-S = Motivation and Engagement Scale-Short; EFA = exploratory factor analysis; BT = Booster Thoughts; BB = Booster Behaviours; Mf = Mufflers; Gz = Guzzlers. Reported factor loadings include only those loadings significant at 5% level.

^aThis item is negatively worded and was reverse-scored prior to analysis.

Following the EFA, while also seeking to retain as much of the original structure as was warranted, the MES-S was reorganised as a three-factor structure, with Booster Behaviours and Booster Thoughts remaining the same and the Mufflers and Guzzlers factors being combined to form a third factor, to be known as “Hampering”. It should be noted, however, that reliability of this new factor was still below the threshold of .70 ($\omega = .54$, 95% CI [.49, .58]).

Owing to the number of cross-loadings in the EFA, it was determined to model the new MES-S three-factor structure as an ESEM.²² This model demonstrated an acceptable fit to the data (CFI = .972, TLI = .938, RMSEA = .041, 90% CI [.030, .052], and all target factor loadings were significant with standardised loadings ranging from .32 to .79 (mean loading = .50, median loading = .40). Correlations between the factors ranged from $r = .37$ to $r = .54$ (mean $r = .45$,

²² Invariance testing is more easily performed with ESEM than with EFA.

median $r = .44$), reflecting an appropriate level of differentiation among the scales. Table 5.29 sets out all of the target and non-target factor loadings, as well as the factor correlations, for the revised MES-S scales, while Figure 5.2 illustrates the revised factor structure to be used in the balance of this study and Study 2.

Table 5.29
Standardised Factor Loadings and Correlations for the MES-S Three-Factor ESEM

MES-S Items	Factor Loadings		
	Booster Thoughts	Booster Behaviours	Hampering
BT1	.79*	-.06	.05
BT2	.33*	.27*	.05
BT3	.32*	.29*	.08
BB1	.05	.69*	.01
BB2	.01	.76*	.08*
BB3	.38*	.33*	.07
Mf1 ^a	-.10	-.15*	.39*
Mf2 ^a	-.23*	-.16*	.40*
Mf3 ^a	.19*	-.12*	.49*
Gz1 ^a	-.08	.32*	.37*
Gz2 ^a	.08	.07	.63*
Factor Correlations			
Booster Thoughts	1.00*		
Booster Behaviours	.54*	1.00*	
Hampering	.44*	.37*	1.00*

Note. MES-S = Motivation and Engagement Scale-Short; ESEM = exploratory structural equation model; BT = Booster Thoughts; BB = Booster Behaviours; Mf = Mufflers; Gz = Guzzlers. Target loadings highlighted grey.

*indicates significant p -value $< .05$

^aThis item is negatively worded and was reverse-scored prior to analysis.

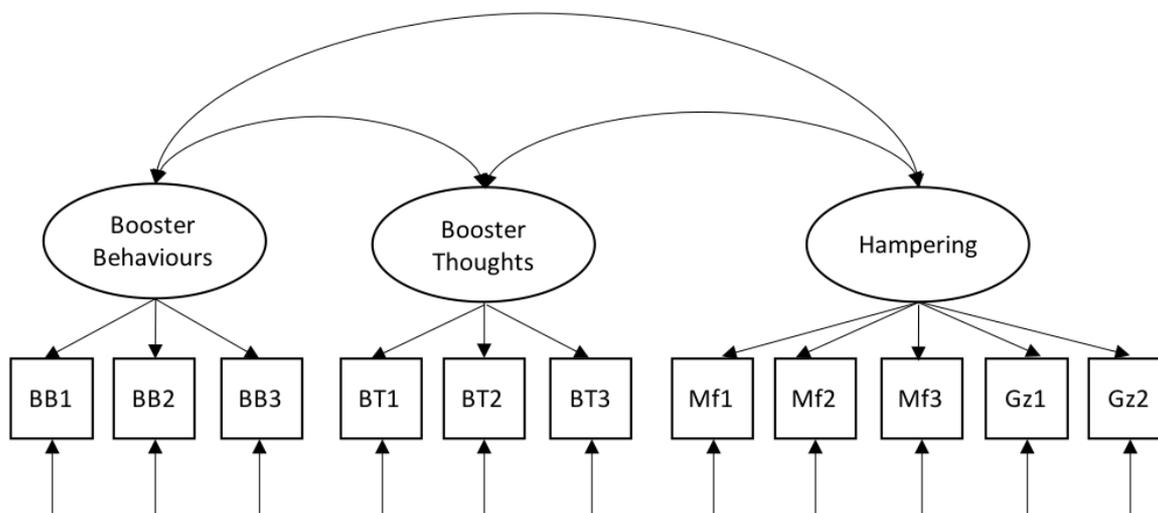


Figure 5.2. Modified MES-S three-factor structure.

Note. BB = Booster Behaviours; BT = Booster Thoughts; GZ = Guzzlers; and Mf = Mufflers. The boxes include the item numbers from the MES-S.

Results of Research Hypothesis 1.3: Invariance analysis for MES-S.

Having selected the three-factor ESEM as the optimal measurement model for the MES-S data, the longitudinal invariance of that data was tested using the selected model. The results of these invariance tests are set out in Table 5.30. The configural model had acceptable fit statistics to support the model at both timepoints. In addition, any degradation in fit statistics from the configural model to the metric model, and from the metric model to the scalar model, were within the acceptable range, suggesting that the MES-S items evidence equivalent relationships to their latent factor over time with their intercepts also being invariant over time. Accordingly, any analysis of mean change over time can be attributed to true change in the constructs of Booster Behaviours, Booster Thoughts, and Hampering, each as measured by the relevant MES-S items.

Table 5.30

Longitudinal Invariance Models Based on MES-S Three-Factor ESEM at T1 and T2: Change in Fit Statistics

Models	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
Configural (1)	.958		.936		.032	
Metric (2)	.969	-.011	.960	-.024	.025	-.007
Scalar (3)	.960	.009	.950	.010	.028	.003

Note. MES-S = Motivation and Engagement Scale-Short; ESEM = exploratory structural equation model; T1 = pre-test; T2 = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). A complete table including chi square test statistics (χ^2), degrees of freedom (df), scaling correlation factor (SCF), and p -value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations = 350.

Conclusion. Research Hypothesis 1.1 was supported only for the Booster Behaviours scale of the MES-S, with reliability estimates below .70 for the other scales. The inter-item correlations were also assessed, and the Booster Thoughts and Guzzlers scales both had good inter-item correlations within the desired range. The Mufflers scale had poor inter-item correlations with each other, and some of those items had higher inter-item correlations with items from other scales, making reliability of the Mufflers scale difficult to support. The four-factor CFA had poor fit, and the high correlation between the Booster Behaviours and Booster Thoughts scales raised issues of multicollinearity. As the four-factor ESEM would not converge, an EFA was investigated. The EFA revealed a three-factor solution to be the best fit, with all of the negatively-worded items in the Guzzlers and Mufflers scales loading onto a single factor. The EFA also demonstrated cross-loadings among some of the items, suggesting that an ESEM might be the preferable model. Taking into account the EFA, while also attempting to retain as much of the hypothesised structure as possible, the MES-S scale structure was modified to combine the Guzzlers and Mufflers scales, while retaining the other two scales. The reliability estimate was still below the threshold for the revised scale, and the inter-item correlations were mixed. Accordingly, Research Hypothesis 1.1 was only supported for the Booster Behaviours and Booster Thoughts scales. The three-factor

ESEM for the MES-S showed good fit to the data, reasonable factor loadings, and factor correlations that evidenced an appropriate level of differentiation among the scales. Therefore, while Research Hypothesis 1.2 was not supported for the hypothesised MES-S four-factor structure, it was supported for the three-factor structure. Invariance testing demonstrated longitudinal invariance for the data with this model, suggesting that Research Hypothesis 1.3 also be accepted for the three-factor ESEM.

Unidimensional Measures: ARS, LRS, SWLS, WEMWBS

Results of Research Hypothesis 1.1: Internal consistency reliability of unidimensional measures. Omega estimates for each unidimensional measure are set out in Table 5.31. Each scale reflects omega estimates, as well as 95% confidence intervals, above the threshold and near to or exceeding a very good rating.

Table 5.31

Reliability Estimates (Omega Coefficients) for the Unidimensional Measures using Long Form Data

Scale	Omega Coefficient	95% CI
ARS (6 items; $n = 1,052$)	.77	.75 / .80
LRS (6 items; $n = 1,041$)	.79	.76 / .81
SWLS (5 items; $n = 1,044$)	.80	.77 / .82
WEMWBS (14 items; $n = 1,022$)	.88	.87 / .90

Note. CI = confidence interval; ARS = Academic Resilience Scale; LRS = Life Resilience Scale; SWLS = Satisfaction with Life Scale; WEMWBS = Warwick-Edinburgh Mental Well-Being Scale; n = number of observations.

Results of Research Hypothesis 1.2: Factor analysis for unidimensional measures. Each unidimensional measure was modelled separately as a unidimensional CFA. The fits statistics for each model are set out in Table 5.32. All models other than the ARS, had an acceptable fit to the data. While the ARS had a TLI and RMSEA outside of the acceptable range, its CFI was acceptable. All factor loadings were significant and greater than the .30 threshold. Only the WEMWBS had factor loadings in the .30 to .40 range (three items). Mean

loadings for each scale are as follows: ARS (mean loading = .61), LRS (mean loading = .62), SWLS (mean loading = .66), and WEMWBS (mean loading = .54).

Table 5.32

Fit Statistics for the Unidimensional Measures using One-Factor CFAs and Long Form Data

One-Factor CFA	CFI	TLI	RMSEA	90% CI
ARS	.907	.845	.095	.079 / .113
LRS	.998	.996	.014	.000 / .039
SWLS	1.000	1.002	.000	.000 / .038
WEMWBS	.954	.945	.041	.034 / .047

Note. CFA = confirmatory factor analysis; CFI = comparative fit index; TLI = Tucker-Lewis fit index; RMSEA = root mean square error of approximation; CI = confidence interval; ARS = Academic Resilience Scale; LRS = Life Resilience Scale; SWLS = Satisfaction with Life Scale; WEMWBS = Warwick-Edinburgh Mental Well-Being Scale. A complete table including chi square test statistics (χ^2), degrees of freedom (df), scaling correlation factor (SCF), and p -value can be found in Appendix O. Number of observations = 1,068.

Results of Research Hypothesis 1.3: Invariance analysis for unidimensional measures. Having accepted the one-factor CFA for each unidimensional measure, the longitudinal invariance of the data for each measure was tested using the selected models. The results of these invariance tests are set out in Table 5.33.

Table 5.33
Longitudinal Invariance Models for Unidimensional Measures Based on One-Factor CFAs at T1 and T2: Change in Fit Statistics

Models	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
ARS						
Configural (1)	.907		.870		.064	
Metric (2)	.899	.008	.872	-.002	.063	-.001
Scalar (3)	.898	.001	.882	-.010	.061	-.002
LRS						
Configural (1)	.994		.991		.015	
Metric (2)	.989	.005	.986	.005	.020	.005
Scalar (3)	.984	.005	.982	.004	.022	.002
SWLS						
Configural (1)	.999		.999		.008	
Metric (2)	.999	.000	.999	.000	.007	-.001
Scalar (3a)	.985	.014	.982	.017	.028	.021
Part. Scalar (3b)	.998	.001	.998	.001	.009	.002
WEMWBS						
Configural (1)	.927		.918		.036	
Metric (2)	.922	.005	.915	.003	.037	.001
Scalar (3)	.916	.006	.912	.003	.038	.001

Note. CFA = confirmatory factor analysis; T₁ = pre-test; T₂ = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). ARS = Academic Resilience Scale; LRS = Life Resilience Scale; SWLS = Satisfaction with Life Scale; WEMWBS = Warwick-Edinburgh Mental Well-Being Scale. For the Partial Scalar (3b) model, the intercept for SL₂ (being the second item on the Satisfaction with Life scale) was free. A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p* value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations = 350.

The configural model for each scale had acceptable fit statistics (other than the TLI for the ARS), supporting each model at both timepoints. The changes in fit statistics from the configural model to the metric model were within the acceptable range for each scale. From the metric model to the scalar model, the changes in fit statistics were within the acceptable range for each scale other than the SWLS. By releasing the constraint on item SL₂ (“The conditions of my life are excellent”), in accordance with the modification indices, partial scalar invariance was achieved. It is suggested, therefore, that the items for each unidimensional measure evidence equivalent relationships to their scale over time and their intercepts are also

invariant or in the case of the SWLS, substantially invariant, over time. Accordingly, any analysis of any analysis of mean change over time can be attributed to true change in the constructs of academic resilience, life resilience, wellbeing, and life satisfaction.

Conclusion. Research Hypothesis 1.1 was supported for each unidimensional measure, with all reliability estimates near to or exceeding a very good rating. The factor analysis revealed the one-factor CFAs to have good fit statistics (with the ARS being borderline), and Research Hypothesis 1.2 was accepted for these models. Longitudinal invariance testing found the configural models on balance to have an acceptable fit, and the changes in fit statistics for the metric and scalar models to be satisfactory (with the exception of the SWLS which achieved only partial scalar invariance). Overall, Research Hypothesis 1.3 was accepted for each unidimensional measure.

Self-Description Questionnaire II – Short (SDQII-S)

Results of Research Hypothesis 1.1: Internal consistency reliability of SDQII-S. Omega estimates for each scale of the SDQII-S are set out in Table 5.34. Each scale reflects omega estimates, as well as 95% confidence intervals, above the threshold and near to or exceeding a very good rating (mean $\omega = .84$, median $\omega = .83$).

Table 5.34
Reliability Estimates (Omega Coefficients) for the SDQII-S and its Scales using Long Form Data

SDQII-S Scale	Omega Coefficient	95% CI	
Non-Academic SC			
Phys Abilities SC ^a (4 items; <i>n</i> = 1,049)	.89	.87	/ .90
Phys Appearance SC (4 items; <i>n</i> = 1,045)	.90	.89	/ .91
Opp-Sex Rel'ships SC ^a (4 items; <i>n</i> = 1,049)	.83	.81	/ .85
Same-Sex Rel'ships SC ^a (5 items; <i>n</i> = 1,046)	.82	.80	/ .85
Parent Rel'ships SC ^a (4 items; <i>n</i> = 1,053)	.86	.84	/ .88
Honesty-Trust SC ^a (6 items; <i>n</i> = 1,046)	.79	.77	/ .81
Emot Stability SC ^a (5 items; <i>n</i> = 1,051)	.79	.77	/ .81
Academic SC			
Math SC ^a (4 items; <i>n</i> = 1,049)	.90	.89	/ .91
Verbal SC ^a (5 items; <i>n</i> = 1,047)	.89	.88	/ .91
School SC ^a (4 items; <i>n</i> = 1,049)	.80	.77	/ .82
Global SC			
Gen Self-Esteem/SC ^a (6 items; <i>n</i> = 1,033)	.81	.79	/ .83

Note. SDQII-S = Self-Description Questionnaire II-S; CI = confidence interval; SC = Self-Concept; Phys = Physical; Opp = Opposite; Rel'ships = Relationships; Trust = Trustworthiness; Emot = Emotional; Gen = General; *n* = number of observations.

^aAll or some of the items for this scale are negatively worded and were reverse-scored prior to analysis.

Results of Research Hypothesis 1.2: Factor analysis for SDQII-S. The SDQII-S was first modelled as a CFA using the a priori 11-factor structure. Correlated uniquenesses were included for two pairs of items that had similar wording: the item “I make friends easily with members of my own sex” was correlated for girl participants with the item “I make friends easily with girls” and for boy participants with the item “I make friends easily with boys.” These correlations were significant, evidencing the appropriateness of their inclusion.

The CFA demonstrated a borderline acceptable fit to the data, as seen in the fit statistics set out in Table 5.35. These results are lower than found in previous research (see Marsh, Ellis, et al., 2005). However, standardised factor loadings were

all significant and, consistent with previous research, ranged from .48 (item Ho4 on Honesty Self-Concept factor) to .95 (item Pab2 on Physical Abilities Self-Concept factor; mean loading = .73, median loading = .72). A number of the factor correlations were low and insignificant (e.g., Math/Verbal Self-Concept factors), with the highest correlation at $r = .77, p < .01$ (School Self-Concept/General Self-Esteem factors) and mean $|r| = .26$, median $|r| = .25$. Consistent with previous findings (see Marsh, Byrne, et al., 1988), the Math and Verbal Self-Concept factors were virtually uncorrelated, but each factor was substantially correlated with the School Self-Concept factor ($r = .57$ and $.55$, respectively), lending further support to the I/E model of self-concept. Also consistent with previous findings (see Marsh & Craven, 1997), the non-academic self-concept factors were more correlated with each other (mean $|r| = .26$) than with the academic self-concept factors (mean $|r| = .17$).

Table 5.35

Fit Statistics for the Hypothesised SDQII-S Factor Models using Long Form Data

Model	CFI	TLI	RMSEA	90% CI
11-factor CFA + CU	.904	.896	.038	.037 / .040
11-factor ESEM + CU	.977	.962	.023	.020 / .025
10-factor ESEM/CFA (Sch)+CU	.955	.932	.031	.029 / .033

Note. SDQII-S = Self-Description Questionnaire II-Short; CFI = comparative fit index, TLI = Tucker-Lewis fit index, RMSEA = root mean square error of approximation; CI = confidence interval; CFA = confirmatory factor analysis; CU = correlated uniqueness; ESEM = exploratory structural equation model; Sch = School Self-Concept. A correlated uniqueness was included for a pair of items from the Same Sex Relationships Self-Concept scale. A complete table including chi square test statistics (χ^2), degrees of freedom (df), scaling correlation factor (SCF), and p -value can be found in Appendix O. Number of observations = 1,068.

An ESEM was modelled as a single set with 11 factors. It was anticipated that allowing the items to cross-load onto the other factors would improve the fit. Fit statistics for the ESEM were excellent and are set out in Table 5.35. The standardised target factor loadings were all significant and well above the threshold of .30 except for three of the four items on the School Self-Concept factor and one item on the General Self-Esteem factor (GS6: "Overall I am a failure). The poor loadings seem to be a consequence of the significant cross-loadings of the School Self-Concept items on the Math Self-Concept, Verbal Self-Concept and General

Self-Esteem factors, as well as the cross-loading of the General Self-Esteem item on the School Self-Concept factor. Standardised target factor loadings for the other items ranged from .41 (item ES4 on Emotional Stability Self-Concept factor) to .94 (item Pab2 on Physical Abilities Self-Concept factor), with a mean target loading (all items) = .67 and median target loading (all items) = .71). The ESEM demonstrated lower factor correlations (mean $|r|$ = .19, median $|r|$ = .17), with the highest correlation at r = .46 (Opposite-Sex Relations Self-Concept/Physical Appearance Self-Concept).

As a result of the poor factor loadings for the School Self-Concept factor, a third model was tested using a CFA for the School Self-Concept factor together with a single set ESEM for all of the other 10 factors. The fit statistics were acceptable, falling between the 11-factor CFA and the 11-factor ESEM (see Table 5.35). All target factor loadings were significant, and all standardised loadings were above the .30 threshold with the exception of the same item on the General Self-Esteem factor (GS6), which had a standardised target factor coefficient of .28 ($p < .01$). Standardised target factor loadings for the other items ranged from .41 (item ES4 on Emotional Stability Self-Concept factor) to .95 (item Pab2 on the Physical Abilities Self-Concept factor), with mean target loading (all items) = .70 and median target loading (all items) = .72. Factor correlations were similar to the ESEM (mean $|r|$ = .22, median $|r|$ = .21), with the exception of the correlations between the School Self-Concept factor and the other academic self-concept factors, as well as with the General Self-Esteem factor, which pairing had the highest correlation (r = .75, $p < .01$). A full set of the factor loadings and factor correlations for this model is included in Appendix Q. Given the acceptable fit of the ESEM/CFA model and parameter estimates that supported all 11 factors, the ESEM/CFA model was selected as the preferred model.

Results of Research Hypothesis 1.3: Invariance analysis for SDQII-S.

Having selected the ESEM/CFA as the optimal measurement model for the SDQII-S data, the longitudinal invariance of that data was using the selected model. The results of these invariance tests are set out in part A of Table 5.36. The configural model required 1,258 parameter estimates (for 102 items and 22 latent factors)

based on a sample size of 350. While the approach used in this study for longitudinal invariance testing is preferable because it takes into account lagged associations between items across assessment waves as well as the within-time covariances, it can result in poor fit statistics for complex models with smaller sample sizes (Vandenberg & Lance, 2000). In addition, the scalar model would not converge. Given the complexity of the SDQII-S model and the convergence issue, longitudinal invariance was also tested in Mplus using a multigroup approach with time as the grouping variable, the complex option, MLR estimation, and MODEL equal to “configural metric scalar,” which produces all three models at once. These results are set out in part B of Table 5.36.

Table 5.36

Two Longitudinal Invariance Models Based on SDQII-S ESEM/CFA at T1 and T2 using (A) Wide Format Data and (B) Long Format Data with a Grouping Variable

A. Longitudinal Invariance Models Based on SDQII-S ESEM/CFA at T1 and T2: Change in Fit Statistics

Models	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
Configural (1)	.870		.836		.040	
Metric (2)	.871	-.001	.852	-.016	.038	-.002
Scalar (3)	---	Did	Not	Converge	---	

Note. SDQII-S = Self-Description Questionnaire II-Short; ESEM = exploratory structural equation model; CFA = confirmatory factor analysis; T1 = pre-test; T2 = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations = 350.

(continues)

Table 5.36 (continued)

Two Longitudinal Invariance Models Based on SDQII-S ESEM/CFA at T1 and T2 using (A) Wide Format Data and (B) Long Format Data with a Grouping Variable

B. Longitudinal Invariance Multigroup Models Based on SDQII-S ESEM/CFA with T1 and T2 as Grouping Variable: Change in Fit Statistics

Models	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
Configural (1)	.929		.891		.043	
Metric (2)	.929	.000	.912	-.021	.039	-.004
Scalar (3)	.927	.002	.911	.001	.039	.000

Note. SDQII-S = Self-Description Questionnaire II-Short; ESEM = exploratory structural equation model; CFA = confirmatory factor analysis; T1 = pre-test; T2 = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations for T1 = 319 and T2 = 292.

Although the original configural model had a CFI and TLI outside the acceptable fit range (while RMSEA was within range), for the multigroup configural model only the TLI was outside of acceptable range. It is suggested, therefore, that the SDQII-S scales demonstrate equivalent structure across both timepoints. For both tests, the fit statistics from the configural model to the metric model demonstrated no degradation in fit statistics, indicating that items evidence equivalent relationships to their scale over time. Fit statistics from the metric to the scalar multigroup model either did not change or had only a very minor degradation, supporting invariance of the item intercepts. Accordingly, it is argued that any analysis of change over time can be attributed to true change in the constructs measured by the SDQII-S.

Conclusion. Research Hypothesis 1.1 was supported for each scale of the SDQII-S, with all reliability estimates above the threshold. While a CFA for the 11 factors provided a good fit to the data, the factor correlations were very high. An 11-factor, one-set ESEM resulted in better fit statistics and reduced factor correlations, however, the School Self-Concept factor was not supported in this model. As a result, a 10-factor ESEM combined with a one-factor School CFA was modelled. The

ESEM/CFA had good fit to the data and good parameter estimates, although the correlations between the School Self-Concept factor and some of the other factors remained high, consistent with the CFA. On the basis of the ESEM/CFA, Research Hypothesis 1.2 was considered satisfied for the SDQII-S data. Longitudinal invariance testing used a multigroup model given the model complexity. This analysis found the configural model to have poor fit statistics, likely related to the complexity of the model and sample size. Comparison of the further constrained metric and scalar models for the SDQII-S was satisfactory. As a consequence, Research Hypothesis 1.3 was accepted for the SDQII-S.

Review of Personal Effectiveness with Locus of Control (ROPELOC)

Results of Research Hypothesis 1.1: Internal consistency reliability of ROPELOC. Omega estimates for each scale of the ROPELOC are set out in Table 5.37. Each scale, other than Active Involvement, reflects an omega estimate above the threshold, with some estimates exceeding .80 in the very good range (mean $\omega = .77$, median $\omega = .78$). Despite the low reliability estimate for the Active Involvement scale, the items in that scale had good inter-item correlations ranging between .31 and .45.

Table 5.37

Reliability Estimates (Omega Coefficients) for the ROPELOC and its Scales using Long Form Data

ROPELOC Scale	Omega Coefficient	95% CI
<i>Personal Abilities and Beliefs</i>		
Self-Confidence (3 items; $n = 1,052$)	.79	.76 / .82
Self-Efficacy (3 items; $n = 1,048$)	.78	.75 / .80
Stress Management (3 items; $n = 1,049$)	.77	.74 / .80
Open Thinking (3 items; $n = 1,057$)	.71	.67 / .74
<i>Social Skills</i>		
Social Effectiveness (3 items; $n = 1,059$)	.82	.80 / .85
Coop Teamwork (3 items; $n = 1,047$)	.79	.76 / .82
Leadership Ability (3 items; $n = 1,036$)	.88	.86 / .90
<i>Organisational Skills</i>		
Time Efficiency (3 items; $n = 1,049$)	.77	.74 / .80
Quality Seeking (3 items; $n = 1,053$)	.74	.70 / .77
Coping with Change (3 items; $n = 1,049$)	.82	.80 / .85
<i>Energy</i>		
Active Involvement (3 items; $n = 1,049$)	.65	.61 / .69
<i>Overall Effectiveness</i>		
Overall Effectiveness (3 items; $n = 1,045$)	.80	.77 / .83
<i>Locus of Control</i>		
Internal LOC (3 items; $n = 1,055$)	.72	.69 / .76
External LOC ^a (3 items; $n = 1,049$)	.71	.68 / .75

Note. ROPELOC = Review of Personal Effectiveness with Locus of Control; CI = confidence interval; Coop = Cooperative; LOC = Locus of Control; n = number of observations.

^a The items for this scale are all negatively worded and were reverse-scored prior to analysis.

Results of Research Hypothesis 1.2: Factor analysis for ROPELOC. The ROPELOC was first modelled as a CFA using the a priori 14-factor structure. As reported in Table 5.38, the fit statistics were excellent or just below. Factor loadings were all significant. Standardised loadings ranged from .56 (AI2 on the Active Involvement factor) to .88 (LA2 on the Leadership Ability factor), with mean loading = .72 and median loading = .74. Factor correlations ranged from $r = .02$ (Self-Efficacy/External Locus of Control) to $r = .83$ (Self-Confidence/Overall Effectiveness), with mean $r = .50$ and median $r = .53$. Substantively, there is a

theoretical argument for overlap between these scales, which are said to each contribute to an overall capacity for life effectiveness. Although factor correlations that exceed .80 or .85 may be found to lack discriminant validity (Brown, 2015), it is important to keep in mind the context of the scale. In addition, although there were two correlations that exceeded .80, there was a wide range of factor correlations and as a whole, the correlations were moderate. Further consideration will be given to the correlations across scales in the analysis of the MTMM matrix below. For a full list of the CFA factor loadings and factor correlations, see Appendix R.

Table 5.38

Fit Statistics for the Hypothesised ROPELOC Factor Models using Long Form Data

Model	CFI	TLI	RMSEA	90% CI
14-factor CFA	.953	.945	.029	.026 / .031
14-factor ESEM	.977	.945	.029	.025 / .032

Note. ROPELOC = Review of Personal Effectiveness with Locus of Control; CFI = comparative fit index, TLI = Tucker-Lewis fit index, RMSEA = root mean square error of approximation; CI = confidence interval; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model. A complete table including chi square test statistics (χ^2), degrees of freedom (df), scaling correlation factor (SCF), and p -value can be found in Appendix O. Number of observations = 1,068.

Given the high factor correlations, an ESEM was also conducted for comparison. This model demonstrated better fit statistics only for the CFI, as reported in Table 5.38. Some of the target factor loadings for the ESEM were not significant, resulting from higher standard errors which were arguably the consequence of testing a less parsimonious model (42 items, 14 factors, and 581 free parameters) with a relatively small sample size. The ESEM standardised target loadings ranged from a loading of .13 (QSI on the Quality Seeking factor) to .89 (SE2 on the Social Effectiveness factor), with a mean loading = .60 and median loading = .62. The factor correlations for the ESEM were reduced, ranging from $r = -.004$ (Self-Efficacy/External Locus of Control) to $r = .66$ (Stress Management/Coping with Change), with mean $|r| = .29$ and median $|r| = .31$, indicating overestimated factor correlations in the CFA. While the ESEM significantly reduced the correlations between factors, it did not result in improved fit statistics for the TLI or RMSEA. In addition, two of the factors (Quality Seeking

and Internal Locus of Control) were not supported by the target loadings of their items. For these reasons, the CFA was selected as the preferred model.

Results of Research Hypothesis 1.3: Invariance analysis for ROPELOC.

Having selected the CFA as the optimal measurement model for the ROPELOC data, the longitudinal invariance of that data was tested using the selected model. The results of these invariance tests are set out in part A of Table 5.39. Similar to the SDQII-S, this model was complex, requiring 672 parameter estimates (for 84 items and 28 latent factors) based on a sample of 350 students. As with the SDQII-S, a multigroup invariance test was also undertaken. The results of the multigroup invariance models are set out in part B of Table 5.39.

Table 5.39

Two Longitudinal Invariance Models Based on ROPELOC CFA at T1 and T2 using (A) Wide Format Data and (B) Long Format Data with a Grouping Variable

A. Longitudinal Invariance Models Based on ROPELOC CFA at T1 and T2: Change in Fit Statistics

Models	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
Configural (1)	.879		.858		.036	
Metric (2)	.879	.000	.860	-.002	.036	.000
Scalar (3)	.876	.003	.857	.003	.036	.000

Note. ROPELOC = Review of Personal Effectiveness with Locus of Control; CFA = confirmatory factor analysis; T1 = pre-test; T2 = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations = 350. (continues)

Table 5.39 (continued)

Two Longitudinal Invariance Models Based on ROPELOC CFA at T1 and T2 using (A) Wide Format Data and (B) Long Format Data with a Grouping Variable

B. Multigroup Invariance Models by Time Based on ROPELOC CFA: Change in Fit Statistics

Models	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
Configural (1)	.925		.912		.039	
Metric (2)	.925	.000	.913	-.001	.038	-.001
Scalar (3)	.923	.002	.912	.001	.039	.001

Note. ROPELOC = Review of Personal Effectiveness with Locus of Control; CFA = confirmatory factor analysis; T1 = pre-test; T2 = immediate post-test; CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease). A complete table including chi square test statistics (χ^2), degrees of freedom (*df*), scaling correlation factor (SCF), and *p*-value for the Satorra-Bentler scaled chi square difference test can be found in Appendix P. Number of observations for T1 = 319 and T2 = 292.

While the original configural model had CFI and TLI statistics that were below the threshold, the multigroup configural model had all fit statistics within the acceptable range. As a result, it is suggested that the ROPELOC scales demonstrate equivalent structure across both timepoints. For both sets of tests, any degradation in fit statistics from the configural model to the metric model, and from the metric model to the scalar model, were within the acceptable range. Accordingly, it is suggested that the ROPELOC items evidence equivalent relationships to their scale over time and their intercepts are also invariant over time. Accordingly, any analysis of change over time can be attributed to true change in the ROPELOC constructs.

Conclusion. Each scale of the ROPELOC, other than Active Involvement, had reliability estimates exceeding .70, and the Active Involvement scale had inter-item correlations within the acceptable range. Accordingly, Research Hypothesis 1.1 was accepted for the ROPELOC scales. The 14-factor CFA and 14-factor, one-set ESEM both fit the ROPELOC data well with similar fit statistics. While the ESEM had meaningfully reduced factor correlations when compared with the CFA, two of the factors were not supported by the target loadings of their items. Accordingly, the CFA was selected as the preferred model and Research Hypothesis 1.2 was

satisfied for the ROPELOC data. The CFA was also found to be invariant over time, supporting Research Hypothesis 1.3.

MTMM Analysis

Research Hypothesis 1.4 proposes that analysis of an MTMM matrix of factor scores for all scales with time as the method variable, will demonstrate the convergent and discriminant validity of responses for each scale. The factor scores were derived from the preferred model for the latent structure for each instrument following the results of the factor structure analysis described above. Analysing correlations for the 41 scales across two occasions (T1 and T2) resulted in an 82 x 82 matrix.

Results of Research Hypothesis 1.4: Construct validity. The pattern of relations among the 41 scales was evaluated to assess the convergent and discriminant validity of responses for each scale.

Results of Research Hypothesis 1.4.1: Convergent validity. To establish convergent validity for the scale responses, the correlations between the factor scores for those scales measured at different occasions (monotrait-heteromethod, MTHM) should be significant and substantial. Table 5.40 shows these convergent validities highlighted. All of these were statistically significant, ranging from $r = .45$ (Consistency of Interest from the Grit-S) to $r = .80$ (Physical Abilities from the SDQII-S), with mean $r = .61$, $SD = .08$. These significant correlations also establish the test-retest stabilities of the scales, further supporting their longitudinal stability.

Results of Research Hypothesis 1.4.2: Discriminant validity.

Discriminant validity refers to the distinctiveness of the different scales.

Discriminant validity is supported when both: (a) the correlations between the factor scores for different scales measured at the same time (heterotrait-monomethod, HTMM); and (b) the correlations between the factor scores for different scales measured at different times (heterotrait-heteromethod, HTHM), are lower than the convergent validities. Table 5.41 shows the HTMM correlations at T1 and Table 5.42 shows the HTMM correlations at T2. The HTHM correlations are all of the entries other than the convergent validities (on the diagonal) in Table 5.40. The absolute values of the 820 HTMM correlations at T1 ranged from $|r| = .00$ to $|r|$

= .91 (mean $|r| = .38$, $SD = .21$), and the absolute values of the 820 HTMM correlations at T2 ranged from $|r| = .00$ to $|r| = .91$ (mean $|r| = .39$, $SD = .21$), with an overall HTMM mean $|r| = .39$, $SD = .21$. The absolute values of the 1,640 HTHM correlations ranged from $r = .00$ to $r = .64$ with mean $|r| = .27$, $SD = .15$. The means of these correlations are lower than the mean convergent validities ($r = .61$), however, there are 12 HTMM correlations at T1 and 13 HTMM correlations at T2 which exceeded the highest MTHM correlation ($r = .80$). All but three of these correlations were between ROPELOC scales.²³ Thus, the ROPELOC factors do not appear to be as well differentiated as the factors from the other scales. This finding is consistent with the factor analysis on the ROPELOC and the theoretical argument for overlap between these scales on the basis they each contribute to an overall capacity for life effectiveness.

Each convergent validity for a scale was also compared with all of the other HTHM correlations involving the same scale. Of the 3,280 comparisons, the convergent validities exceeded their related HTHM correlations in all but 11 cases. The violations involved correlations between School Self-Concept and Agency at T1 (.50) and T2 (.49), General Self-Esteem and Agency at T1 (.52) and T2 (.56), Booster Thoughts and Agency at T2 (.49), and Self-Confidence and Agency at T2 (.49), all of which were higher than the convergent validity for Agency (.48). In addition, the correlation of Open Thinking with Pathways Thinking at T2 (.50) exceeded the convergent validity for Pathways Thinking (.48), the correlation of Self-Efficacy and Coping with Change with Life Resilience at T2 (.60 and .59, respectively) exceeded the convergent validity for Life Resilience (.57), and the correlation of Life Resilience and Self-Efficacy with Coping with Change at T1 (.59 and .58, respectively) exceeded the convergent validity for Coping with Change (.53). Overall, these violations ranged between .01 and .08 with a mean discrepancy of .03. On balance, therefore, it is suggested that the scales evidence sufficient discriminant validity.

²³ The exceptions were the correlations between Agency and Pathways Thinking at T1 ($r = .85$), Global Self-Esteem and School Self-Concept at T2 ($r = .83$), and Global Self-Esteem and Overall Effectiveness at T2 ($r = .83$).

Table 5.40

Monotrait-Heteromethod Correlations (Highlighted on the Diagonal) and Heterotrait- Heteromethod Correlations (All Other Cells)

	AG_1	PATH_1	OPT_1	PESS_1	FOCUS_1	GOALSR_1	EMOTSR_1	COI_1	POE_1	BTH_1	BBH_1	HAMP_1	AR_1	LR_1	SWL_1	WB_1	PAB_1	PAP_1	OPsx_1	SMSX_1
AG_2	.48	.42	.32	.11	.23	.45	.30	.11	.44	.49	.35	.29	.40	.39	.41	.44	.15	.12	.11	.28
PATH_2	.43	.48	.29	.03	.26	.46	.32	.13	.43	.44	.31	.24	.34	.34	.29	.37	.16	.07	.10	.26
OPT_2	.32	.31	.55	.32	.24	.31	.20	.08	.30	.30	.25	.18	.21	.23	.43	.43	.18	.16	.03	.23
PESS_2	.14	.09	.26	.46	.15	.15	.10	.15	.15	.18	.12	.30	.12	.16	.24	.20	.02	-.01	.00	.20
FOCUS_2	.35	.33	.28	.10	.61	.60	.51	.24	.43	.48	.49	.39	.37	.31	.20	.31	.14	.01	-.02	.12
GOALSR_2	.45	.45	.37	.15	.46	.63	.48	.25	.55	.50	.44	.38	.39	.38	.25	.43	.18	.08	-.05	.25
EMOTSR_2	.37	.37	.32	.07	.47	.52	.59	.12	.43	.39	.33	.33	.41	.40	.18	.35	.16	.05	.04	.23
COI_2	.14	.09	.05	.14	.24	.22	.21	.45	.22	.08	.18	.21	.04	.08	.00	-.01	.13	-.07	.02	.10
POE_2	.39	.34	.39	.19	.39	.52	.38	.22	.59	.37	.33	.29	.33	.35	.20	.37	.18	.12	.07	.29
BTH_2	.38	.31	.34	.24	.40	.48	.30	.08	.39	.63	.48	.42	.31	.22	.30	.34	.05	.06	-.03	.24
BBH_2	.33	.28	.32	.18	.38	.45	.26	.20	.37	.44	.53	.38	.17	.15	.19	.20	.07	.04	-.05	.10
HAMP_2	.24	.21	.22	.26	.35	.36	.31	.29	.30	.45	.35	.49	.30	.27	.18	.22	.01	-.06	-.07	.13
AR_2	.38	.33	.29	.15	.32	.47	.39	.15	.39	.39	.24	.30	.58	.56	.31	.43	.07	.08	.06	.18
LR_2	.43	.40	.36	.16	.37	.52	.47	.15	.48	.41	.32	.31	.49	.57	.36	.49	.14	.11	.11	.26
SWL_2	.36	.29	.35	.21	.18	.34	.28	.00	.32	.34	.26	.22	.30	.29	.61	.46	.08	.13	.04	.18
WB_2	.41	.35	.41	.23	.22	.42	.31	.05	.40	.38	.26	.23	.33	.36	.48	.57	.21	.20	.12	.32
PAB_2	.11	.15	.18	.05	.08	.18	.07	.11	.19	.06	.09	.04	.11	.16	.20	.31	.80	.32	.35	.19
PAP_2	.25	.21	.38	.25	.13	.25	.17	.00	.19	.19	.16	.19	.26	.31	.37	.43	.33	.62	.32	.26
OPsx_2	.05	.06	.05	.03	.01	.08	.08	.06	.13	.07	.08	.02	.09	.15	.16	.26	.36	.40	.72	.28
SMSX_2	.19	.13	.13	.16	.07	.14	.13	.20	.20	.27	.13	.24	.13	.13	.17	.23	.08	.10	.21	.61
PR_2	.25	.21	.27	.21	.20	.26	.25	.12	.23	.28	.27	.27	.18	.20	.38	.32	.03	.06	.00	.26
HO_2	.13	.11	.14	.16	.19	.23	.14	.22	.27	.19	.21	.22	.11	.11	.07	.11	.02	-.03	.08	.13
ES_2	.20	.12	.10	.28	.09	.14	.22	.18	.14	.08	-.02	.23	.25	.32	.19	.24	.05	.02	.09	.29
MH_2	.23	.15	.10	.03	.19	.20	.07	.03	.15	.36	.22	.27	.30	.14	.14	.14	.02	.09	.03	.05
VER_2	.30	.23	.16	.16	.20	.21	.19	.09	.28	.28	.21	.19	.13	.12	.18	.18	.01	.09	.21	.29
SCH_2	.50	.40	.34	.22	.38	.45	.30	.12	.41	.61	.38	.43	.43	.31	.37	.40	.08	.15	.08	.36
GS_2	.52	.46	.47	.25	.36	.50	.38	.11	.46	.54	.35	.38	.39	.39	.48	.56	.22	.19	.06	.39
SC_2	.46	.43	.46	.23	.32	.52	.39	.18	.49	.50	.33	.33	.35	.38	.42	.56	.25	.21	.15	.36
SF_2	.44	.42	.36	.14	.34	.51	.46	.08	.41	.39	.23	.27	.46	.47	.37	.51	.17	.17	.04	.25
SM_2	.42	.41	.29	.08	.29	.44	.44	.03	.34	.33	.21	.24	.44	.44	.34	.47	.14	.13	.05	.26
OT_2	.39	.41	.37	.10	.29	.48	.37	.11	.45	.45	.26	.26	.38	.36	.35	.49	.20	.13	.08	.28
SE_2	.26	.27	.29	.15	.08	.27	.20	.16	.31	.24	.14	.10	.21	.24	.26	.43	.31	.22	.37	.38
CT_2	.21	.23	.25	.12	.16	.29	.23	.16	.28	.28	.23	.18	.22	.20	.29	.38	.32	.14	.17	.32
LA_2	.27	.28	.29	.21	.18	.31	.21	.28	.37	.26	.18	.19	.21	.26	.25	.38	.32	.24	.29	.30
TE_2	.40	.38	.39	.20	.40	.55	.37	.17	.47	.37	.40	.37	.30	.36	.29	.43	.22	.20	.06	.20
QS_2	.36	.36	.46	.18	.37	.49	.35	.14	.46	.53	.44	.38	.27	.27	.33	.44	.18	.12	.01	.26
CH_2	.39	.40	.29	.14	.32	.48	.41	.13	.41	.37	.22	.29	.43	.44	.33	.48	.18	.14	.08	.29
AI_2	.32	.32	.34	.12	.19	.39	.26	.14	.39	.37	.24	.18	.27	.28	.33	.47	.41	.20	.24	.28
OE_2	.43	.41	.47	.26	.31	.49	.36	.12	.49	.46	.29	.31	.34	.38	.42	.56	.22	.22	.13	.34
IL_2	.33	.34	.38	.10	.27	.41	.31	.10	.40	.51	.32	.31	.26	.24	.33	.41	.14	.08	.01	.30
EL_2	.11	.09	.06	.19	.11	.13	.14	.19	.19	.24	.11	.27	.13	.15	.04	.07	.00	-.05	-.04	.19

Note. AG = Agency; PATH = Pathways Thinking; OPT = Optimism; PESS = Pessimism; FOCUS = Focus; GOALSR = Goal Self-Regulation; EMOTSR = Emotion Self-Regulation; COI = Consistency of Interest; POE = Perseverance of Effort; BTH = Booster Thoughts; BBH = Booster Behaviors; HAMP = Hampering; AR = Academic Resilience; LR = Life Resilience; SWL = Life Satisfaction; WB = Wellbeing; PAB = Physical Abilities Self-Concept; PAP = Physical Appearance Self-Concept; OPSX = Opposite-SEX Relationships Self-Concept; SMSX = Same-Sex Relationships Self-Concept; PR = Parent Relationships Self-Concept; HO = Honesty-Trustworthiness Self-Concept; ES = Emotional Stability Self-Concept; MH = Math Self-Concept; VER = Verbal Self-Concept; SCH = School Self-Concept; GS = General Self-Esteem/Self-Concept; SC = Self-Confidence; SF = Self-Efficacy; SM = Stress Management; OT = Open Thinking; SE = Social Effectiveness; CT = Cooperative Teamwork; LA = Leadership Ability; TE = Time Efficiency; QS = Quality Seeking; CH = Coping with Change; AI = Active Involvement; OE = Overall Effectiveness; IL = Internal Locus of Control; EL = External Locus of Control. _1 refers to data collected pre-program (T1) and _2 refers to data collected immediately post-progra (T2). Convergent validities are shaded grey.

(continues)

Table 5.40 (continued)

Monotrait-Heteromethod Correlations (Highlighted on the Diagonal) and Heterotrait- Heteromethod Correlations (All Other Cells)

	PR_1	HO_1	ES_1	MH_1	VER_1	SCH_1	GS_1	SC_1	SF_1	SM_1	OT_1	SE_1	CT_1	LA_1	TE_1	QS_1	CH_1	AI_1	OE_1	IL_1	EL_1
AG_2	.28	.20	.15	.26	.19	.49	.56	.49	.47	.41	.45	.29	.30	.33	.38	.40	.44	.38	.47	.41	.08
PATH_2	.23	.20	.06	.17	.22	.44	.48	.47	.44	.40	.50	.31	.34	.33	.41	.42	.41	.44	.43	.45	.00
OPT_2	.28	.14	.12	.09	.13	.31	.41	.42	.30	.22	.34	.25	.26	.29	.32	.36	.29	.33	.42	.33	.05
PESS_2	.22	.29	.17	.06	.12	.18	.21	.19	.10	.07	.09	.00	.00	.11	.16	.15	.07	.06	.20	.16	.38
FOCUS_2	.28	.23	.09	.14	.12	.33	.40	.43	.41	.40	.44	.20	.22	.18	.51	.44	.41	.33	.38	.41	.08
GOALSR_2	.31	.26	.12	.13	.19	.42	.54	.56	.47	.45	.62	.34	.37	.36	.54	.54	.50	.51	.51	.55	.01
EMOTSR_2	.30	.19	.17	.07	.12	.27	.39	.42	.45	.50	.50	.22	.30	.20	.45	.40	.48	.38	.35	.42	.02
COI_2	.14	.27	.16	.02	.07	.04	.04	.08	.10	.13	.03	.05	.08	.02	.21	.06	.10	.04	.07	.02	.25
POE_2	.25	.23	.22	.13	.15	.37	.47	.50	.42	.38	.47	.27	.26	.29	.45	.44	.43	.40	.43	.43	.07
BTH_2	.37	.23	.01	.26	.20	.51	.50	.46	.27	.22	.43	.20	.25	.20	.41	.53	.29	.35	.42	.53	.13
BBH_2	.27	.22	.00	.13	.10	.28	.33	.33	.22	.17	.26	.16	.18	.11	.43	.39	.24	.22	.30	.27	.04
HAMP_2	.32	.34	.11	.23	.16	.32	.32	.32	.24	.21	.22	.08	.05	.09	.35	.31	.20	.14	.31	.32	.37
AR_2	.20	.20	.27	.22	.05	.33	.43	.46	.52	.46	.47	.26	.20	.30	.32	.30	.48	.33	.44	.37	.08
LR_2	.32	.23	.28	.18	.08	.34	.48	.51	.60	.55	.54	.30	.29	.29	.44	.40	.59	.42	.47	.40	.02
SWL_2	.44	.09	.16	.06	.08	.30	.48	.40	.31	.25	.34	.20	.23	.29	.28	.33	.28	.32	.38	.37	.10
WB_2	.39	.17	.22	.10	.12	.36	.54	.48	.41	.37	.49	.34	.40	.37	.38	.42	.43	.47	.45	.45	.03
PAB_2	.11	.13	.10	.02	-.09	.06	.24	.30	.18	.14	.22	.37	.41	.33	.21	.24	.24	.47	.28	.19	.06
PAP_2	.23	.11	.23	.02	.13	.22	.39	.35	.29	.26	.31	.37	.32	.36	.26	.25	.33	.38	.36	.28	.09
OPSX_2	-.02	.09	.10	-.03	.16	.10	.14	.24	.15	.16	.20	.45	.32	.37	.10	.12	.24	.35	.24	.17	.06
SMSX_2	.23	.22	.21	.08	.26	.27	.25	.23	.14	.14	.17	.25	.22	.17	.19	.21	.18	.21	.17	.22	.26
PR_2	.75	.34	.15	.00	.12	.18	.33	.26	.18	.14	.23	.12	.17	.20	.29	.27	.17	.20	.24	.26	.14
HO_2	.26	.67	.10	.02	.15	.13	.11	.20	.09	.07	.12	.15	.10	.19	.22	.18	.07	.15	.20	.19	.23
ES_2	.18	.17	.58	.06	.07	.11	.17	.19	.32	.34	.17	.16	.15	.13	.11	.07	.28	.13	.19	.11	.27
MH_2	.01	.03	.01	.75	-.14	.44	.25	.23	.19	.15	.13	.06	.01	.05	.12	.20	.15	.07	.22	.20	.05
VER_2	.18	.19	.06	-.08	.74	.40	.26	.25	.18	.13	.25	.22	.18	.27	.17	.22	.16	.22	.24	.26	.12
SCH_2	.29	.22	.12	.47	.36	.71	.59	.51	.40	.32	.45	.26	.24	.30	.37	.49	.36	.35	.48	.52	.13
GS_2	.43	.23	.21	.27	.19	.57	.67	.59	.49	.41	.56	.35	.38	.36	.45	.56	.47	.49	.54	.58	.07
SC_2	.35	.24	.18	.19	.22	.49	.61	.66	.47	.40	.60	.48	.46	.47	.50	.57	.50	.59	.59	.59	.11
SF_2	.30	.15	.25	.16	.10	.37	.51	.54	.60	.57	.56	.34	.36	.32	.42	.41	.58	.43	.50	.43	-.04
SM_2	.24	.10	.27	.07	.03	.26	.41	.43	.52	.58	.49	.29	.36	.24	.36	.33	.53	.38	.40	.37	-.04
OT_2	.28	.17	.09	.14	.16	.42	.53	.57	.44	.40	.63	.40	.45	.45	.41	.51	.47	.58	.52	.58	.02
SE_2	.22	.29	.21	-.02	.23	.24	.34	.46	.38	.34	.45	.62	.53	.56	.28	.32	.42	.56	.42	.36	.03
CT_2	.24	.20	.17	.04	.05	.23	.33	.39	.26	.26	.40	.49	.59	.52	.29	.33	.31	.53	.36	.35	.06
LA_2	.24	.33	.19	.08	.29	.35	.39	.48	.29	.23	.46	.56	.49	.73	.35	.32	.33	.58	.44	.35	.06
TE_2	.33	.23	.19	.10	.11	.31	.46	.50	.43	.34	.45	.32	.31	.31	.55	.48	.42	.40	.50	.41	.07
QS_2	.32	.20	.06	.23	.14	.45	.54	.57	.36	.28	.52	.32	.35	.29	.50	.62	.39	.47	.52	.57	.06
CH_2	.28	.18	.26	.10	.07	.31	.44	.49	.48	.49	.53	.36	.37	.33	.40	.40	.53	.44	.48	.45	.03
AL_2	.25	.22	.08	.07	.12	.33	.47	.53	.33	.28	.54	.53	.58	.57	.37	.46	.39	.69	.46	.48	.00
OE_2	.37	.23	.19	.19	.21	.47	.60	.64	.47	.37	.57	.43	.38	.45	.47	.56	.49	.52	.63	.57	.11
IL_2	.28	.15	.03	.20	.20	.47	.51	.54	.32	.25	.54	.32	.35	.35	.38	.55	.36	.48	.61	.61	.12
EL_2	.17	.23	.19	.09	.17	.17	.12	.12	.03	.03	.08	.05	.00	.10	.11	.14	.05	.04	.10	.22	.53

Note. AG = Agency; PATH = Pathways Thinking; OPT = Optimism; PESS = Pessimism; FOCUS = Focus; GOALSR = Goal Self-Regulation; EMOTSR = Emotion Self-Regulation; COI = Consistency of Interest; POE = Perseverance of Effort; BTH = Booster Thoughts; BBH = Booster Behaviors; HAMP = Hampering; AR = Academic Resilience; LR = Life Resilience; SWL = Life Satisfaction; WB = Wellbeing; PAB = Physical Abilities Self-Concept; PAP = Physical Appearance Self-Concept; OPSX = Opposite-SEX Relationships Self-Concept; SMSX = Same-Sex Relationships Self-Concept; PR = Parent Relationships Self-Concept; HO = Honesty-Trustworthiness Self-Concept; ES = Emotional Stability Self-Concept; MH = Math Self-Concept; VER = Verbal Self-Concept; SCH = School Self-Concept; GS = General Self-Esteem/Self-Concept; SC = Self-Confidence; SF = Self-Efficacy; SM = Stress Management; OT = Open Thinking; SE = Social Effectiveness; CT = Cooperative Teamwork; LA = Leadership Ability; TE = Time Efficiency; QS = Quality Seeking; CH = Coping with Change; AI = Active Involvement; OE = Overall Effectiveness; IL = Internal Locus of Control; EL = External Locus of Control. _1 refers to data collected pre-program (T1) and _2 refers to data collected immediately post-program (T2). Convergent validities are shaded grey.

Table 5.41

Heterotrait-Monomethod Correlations at T1

	AG_1	PATH_1	OPT_1	PESS_1	FOCUS_1	GOALSR_1	EMOTSR_1	COI_1	POE_1	BTH_1	BBH_1	HAMP_1	AR_1	LR_1	SWL_1	WB_1	PAB_1	PAP_1	OPSX_1	SMSX_1	
AG_1																					
PATH_1	.85																				
OPT_1	.45	.51																			
PESS_1	.18	.11	.47																		
FOCUS_1	.52	.52	.35	.15																	
GOALSR_1	.61	.64	.47	.14	.74																
EMOTSR_1	.54	.57	.39	.08	.75	.74															
COI_1	.06	.03	.03	.23	.30	.25	.21														
POE_1	.58	.58	.45	.21	.58	.70	.59	.32													
BTH_1	.55	.54	.39	.16	.56	.65	.49	.15	.57												
BBH_1	.44	.42	.34	.11	.67	.61	.50	.21	.53	.71											
HAMP_1	.42	.35	.28	.38	.47	.43	.37	.39	.39	.59	.47										
AR_1	.55	.54	.34	.15	.55	.60	.62	.08	.53	.53	.37	.43									
LR_1	.56	.58	.41	.24	.47	.60	.62	.14	.53	.45	.31	.38	.74								
SWL_1	.50	.47	.51	.26	.25	.41	.33	-.01	.40	.40	.31	.25	.40	.42							
WB_1	.57	.59	.63	.30	.38	.59	.51	.02	.54	.50	.34	.34	.55	.60	.70						
PAB_1	.17	.20	.23	.02	.14	.22	.13	.09	.28	.05	.09	.07	.16	.17	.20	.33					
PAP_1	.29	.28	.31	.17	.10	.23	.16	-.04	.22	.11	.09	.13	.22	.24	.36	.46	.40				
OPSX_1	.16	.17	.07	-.01	.03	.09	.08	.05	.18	.06	.06	.06	.10	.11	.15	.25	.39	.50			
SMSX_1	.35	.30	.30	.34	.14	.20	.22	.16	.24	.26	.14	.33	.22	.29	.28	.36	.17	.23	.24		
PR_1	.31	.28	.37	.17	.28	.34	.33	.14	.31	.31	.28	.29	.27	.30	.47	.43	.10	.07	-.07	.32	
HO_1	.28	.26	.23	.23	.33	.34	.28	.33	.39	.29	.31	.41	.19	.27	.16	.20	.12	.01	.11	.24	
ES_1	.27	.18	.17	.39	.18	.13	.25	.30	.20	.07	.02	.38	.30	.41	.19	.29	.16	.17	.07	.39	
MH_1	.27	.22	.12	.07	.26	.22	.15	.10	.20	.43	.25	.34	.31	.17	.11	.15	-.01	.09	-.01	.09	
VER_1	.39	.35	.21	.20	.27	.29	.25	.17	.33	.41	.27	.24	.20	.22	.20	.24	-.05	.12	.19	.32	
SCH_1	.65	.57	.41	.23	.46	.55	.40	.13	.53	.73	.47	.48	.50	.41	.41	.49	.09	.21	.10	.34	
GS_1	.70	.64	.57	.26	.46	.66	.50	.08	.64	.64	.43	.41	.57	.56	.62	.73	.30	.39	.14	.32	
SC_1	.66	.71	.62	.25	.55	.76	.58	.16	.72	.66	.48	.40	.60	.63	.58	.78	.34	.37	.27	.36	
SF_1	.60	.66	.49	.18	.54	.66	.63	.12	.60	.50	.36	.32	.65	.73	.47	.67	.20	.29	.18	.27	
SM_1	.57	.61	.41	.15	.56	.60	.69	.11	.54	.43	.33	.31	.65	.70	.40	.59	.16	.27	.16	.27	
OT_1	.61	.67	.54	.15	.57	.74	.62	.08	.69	.62	.47	.37	.62	.60	.46	.69	.27	.26	.21	.27	
SE_1	.37	.40	.39	.17	.22	.41	.32	.11	.42	.31	.23	.16	.33	.36	.38	.60	.40	.43	.47	.39	
CT_1	.37	.37	.36	.10	.27	.41	.37	.03	.42	.31	.26	.17	.31	.27	.33	.52	.45	.33	.35	.32	
LA_1	.36	.39	.35	.17	.19	.38	.26	.12	.47	.30	.21	.14	.31	.33	.40	.53	.39	.40	.41	.26	
TE_1	.56	.57	.48	.20	.68	.74	.62	.29	.68	.59	.63	.51	.51	.54	.37	.57	.31	.30	.16	.24	
QS_1	.60	.64	.55	.15	.59	.74	.55	.13	.66	.72	.59	.46	.50	.46	.44	.63	.29	.22	.15	.31	
CH_1	.59	.62	.49	.21	.54	.68	.61	.11	.62	.53	.40	.36	.64	.71	.48	.69	.27	.33	.26	.32	
AI_1	.49	.53	.50	.13	.38	.59	.45	.05	.59	.46	.35	.25	.42	.42	.41	.67	.53	.38	.36	.30	
OE_1	.60	.65	.59	.28	.50	.69	.52	.14	.67	.60	.44	.40	.58	.62	.60	.76	.32	.36	.26	.29	
IL_1	.58	.63	.50	.12	.50	.69	.51	.08	.60	.67	.46	.42	.51	.46	.43	.60	.22	.17	.14	.32	
EL_1	.12	.07	.07	.29	.10	.07	.06	.23	.14	.13	.06	.33	.11	.18	.09	.05	.02	.00	.08	.30	

Note. AG = Agency; PATH = Pathways Thinking; OPT = Optimism; PESS = Pessimism; FOCUS = Focus; GOALSR = Goal Self-Regulation; EMOTSR = Emotion Self-Regulation; COI = Consistency of Interest; POE = Perseverance of Effort; BTH = Booster Thoughts; BBH = Booster Behaviors; HAMP = Hampering; AR = Academic Resilience; LR = Life Resilience; SWL = Life Satisfaction; WB = Wellbeing; PAB = Physical Abilities Self-Concept; PAP = Physical Appearance Self-Concept; OPSX = Opposite-SEX Relationships Self-Concept; SMSX = Same-Sex Relationships Self-Concept; PR = Parent Relationships Self-Concept; HO = Honesty-Trustworthiness Self-Concept; ES = Emotional Stability Self-Concept; MH = Math Self-Concept; VER = Verbal Self-Concept; SCH = School Self-Concept; GS = General Self-Esteem/Self-Concept; SC = Self-Confidence; SF = Self-Efficacy; SM = Stress Management; OT = Open Thinking; SE = Social Effectiveness; CT = Cooperative Teamwork; LA = Leadership Ability; TE = Time Efficiency; QS = Quality Seeking; CH = Coping with Change; AI = Active Involvement; OE = Overall Effectiveness; IL = Internal Locus of Control; EL = External Locus of Control. _1 refers to data collected pre-program (T1).

(continues)

Table 5.41 (continued)
Heterotrait-Monomethod Correlations at T1

	PR_1	HO_1	ES_1	MH_1	VER_1	SCH_1	GS_1	SC_1	SF_1	SM_1	OT_1	SE_1	CT_1	LA_1	TE_1	QS_1	CH_1	AI_1	OE_1	IL_1	EL_1		
AG_1																							
PATH_1																							
OPT_1																							
PESS_1																							
FOCUS_1																							
GOALSR_1																							
EMOTSR_1																							
COI_1																							
POE_1																							
BTH_1																							
BBH_1																							
HAMP_1																							
AR_1																							
LR_1																							
SWL_1																							
WB_1																							
PAB_1																							
PAP_1																							
OPPX_1																							
SMSX_1																							
PR_1																							
HO_1	.38																						
ES_1	.19	.27																					
MH_1	-.03	.08	.04																				
VER_1	.16	.25	.13	-.01																			
SCH_1	.21	.26	.14	.62	.56																		
GS_1	.43	.24	.22	.31	.34	.79																	
SC_1	.35	.33	.22	.25	.33	.64	.79																
SF_1	.30	.25	.33	.21	.26	.52	.68	.76															
SM_1	.24	.22	.38	.21	.21	.45	.58	.64	.89														
OT_1	.32	.26	.16	.20	.32	.60	.74	.82	.77	.69													
SE_1	.21	.25	.21	.07	.25	.34	.47	.66	.46	.44	.53												
CT_1	.22	.22	.18	.06	.14	.33	.48	.54	.45	.44	.60	.69											
LA_1	.22	.26	.18	.04	.30	.38	.52	.62	.42	.31	.56	.71	.63										
TE_1	.35	.41	.21	.20	.27	.51	.65	.73	.63	.59	.66	.43	.45	.44									
QS_1	.36	.34	.09	.28	.30	.63	.73	.84	.58	.47	.77	.44	.52	.37	.74								
CH_1	.31	.27	.31	.23	.25	.53	.70	.76	.87	.86	.80	.58	.52	.42	.67	.64							
AI_1	.25	.25	.11	.10	.22	.45	.63	.75	.54	.49	.81	.73	.81	.70	.58	.68	.64						
OE_1	.37	.33	.22	.23	.32	.60	.77	.91	.74	.60	.79	.59	.48	.57	.70	.78	.77	.66					
IL_1	.31	.29	.07	.25	.33	.62	.67	.81	.55	.45	.83	.43	.46	.41	.58	.88	.59	.67	.73				
EL_1	.12	.28	.32	.04	.16	.10	.06	.16	.00	.02	.03	.08	.03	.10	.11	.12	.06	.02	.15	.21			

Note. AG = Agency; PATH = Pathways Thinking; OPT = Optimism; PESS = Pessimism; FOCUS = Focus; GOALSR = Goal Self-Regulation; EMOTSR = Emotion Self-Regulation; COI = Consistency of Interest; POE = Perseverance of Effort; BTH = Booster Thoughts; BBH = Booster Behaviors; HAMP = Hampering; AR = Academic Resilience; LR = Life Resilience; SWL = Life Satisfaction; WB = Wellbeing; PAB = Physical Abilities Self-Concept; PAP = Physical Appearance Self-Concept; OPSX = Opposite-SEX Relationships Self-Concept; SMSX = Same-Sex Relationships Self-Concept; PR = Parent Relationships Self-Concept; HO = Honesty-Trustworthiness Self-Concept; ES = Emotional Stability Self-Concept; MH = Math Self-Concept; VER = Verbal Self-Concept; SCH = School Self-Concept; GS = General Self-Esteem/Self-Concept; SC = Self-Confidence; SF = Self-Efficacy; SM = Stress Management; OT = Open Thinking; SE = Social Effectiveness; CT = Cooperative Teamwork; LA = Leadership Ability; TE = Time Efficiency; QS = Quality Seeking; CH = Coping with Change; AI = Active Involvement; OE = Overall Effectiveness; IL = Internal Locus of Control; EL = External Locus of Control. _1 refers to data collected pre-program (T1).

Table 5.42

Heterotrait-Monomethod Correlations at T2

	AG_2	PATH_2	OPT_2	PESS_2	FOCUS_2	GOALSR_2	EMOTSR_2	COI_2	POE_2	BTH_2	BBH_2	HAMP_2	AR_2	LR_2	SWL_2	WB_2	PAB_2	PAP_2	OPSX_2	SMSX_2	
AG_2	.80																				
PATH_2	.47	.44																			
OPT_2	.24	.17	.34																		
PESS_2	.42	.49	.39	.15																	
FOCUS_2	.55	.64	.45	.13	.77																
GOALSR_2	.41	.52	.39	.07	.78	.78															
EMOTSR_2	.10	.12	.05	.28	.29	.24	.23														
COI_2	.48	.55	.47	.19	.59	.72	.64	.33													
POE_2	.48	.46	.53	.29	.59	.61	.48	.11	.53												
BTH_2	.41	.40	.47	.14	.65	.61	.53	.25	.50	.64											
BBH_2	.40	.34	.30	.53	.48	.42	.35	.44	.40	.55	.47										
HAMP_2	.59	.52	.42	.19	.55	.56	.56	.16	.54	.45	.36	.44									
AR_2	.62	.61	.47	.17	.58	.64	.69	.16	.62	.48	.44	.39	.76								
LR_2	.48	.41	.49	.35	.40	.45	.44	.09	.42	.52	.35	.32	.46	.52							
SWL_2	.62	.60	.63	.30	.50	.65	.60	.07	.58	.59	.45	.33	.57	.65	.73						
WB_2	.17	.21	.22	.07	.19	.23	.22	.14	.25	.10	.11	.06	.14	.23	.19	.34					
PAB_2	.27	.25	.39	.23	.19	.26	.24	-.02	.28	.25	.20	.13	.33	.28	.39	.49	.36				
PAP_2	.09	.16	.10	.07	.03	.15	.08	-.03	.10	.03	-.03	-.06	.12	.15	.17	.27	.43	.51			
OPSX_2	.26	.29	.15	.28	.13	.29	.20	.30	.35	.24	.14	.31	.17	.27	.18	.31	.18	.18	.25		
SMSX_2	.32	.27	.40	.31	.34	.38	.37	.19	.33	.48	.38	.39	.30	.40	.56	.53	.14	.32	.02	.35	
PR_2	.21	.21	.20	.36	.26	.28	.22	.32	.33	.31	.32	.50	.28	.27	.16	.22	.14	.08	.04	.32	
HO_2	.23	.17	.13	.42	.10	.13	.15	.40	.28	.13	.02	.37	.31	.35	.26	.25	.08	.16	.12	.39	
ES_2	.32	.18	.20	.12	.23	.19	.10	.01	.19	.38	.23	.34	.34	.23	.19	.18	.06	.08	-.02	.04	
MH_2	.24	.29	.28	.22	.27	.31	.24	.09	.28	.35	.27	.21	.11	.17	.23	.30	.03	.22	.26	.27	
VER_2	.61	.53	.49	.31	.51	.56	.41	.08	.53	.72	.47	.52	.52	.49	.51	.58	.15	.30	.12	.32	
SCH_2	.65	.61	.61	.29	.56	.68	.55	.09	.61	.71	.50	.44	.58	.64	.67	.80	.31	.43	.19	.30	
GS_2	.63	.69	.62	.25	.61	.76	.63	.14	.70	.67	.52	.41	.60	.67	.58	.79	.34	.43	.27	.34	
SC_2	.57	.62	.51	.12	.64	.73	.71	.14	.62	.54	.48	.33	.67	.77	.49	.70	.26	.35	.18	.22	
SF_2	.49	.55	.45	.09	.58	.67	.72	.15	.56	.48	.45	.26	.61	.73	.45	.66	.25	.31	.15	.25	
SM_2	.59	.71	.53	.15	.58	.79	.65	.12	.67	.63	.45	.36	.58	.64	.49	.70	.30	.33	.20	.35	
OT_2	.41	.47	.39	.06	.30	.52	.42	.11	.43	.30	.24	.12	.41	.47	.34	.57	.44	.40	.45	.32	
SE_2	.41	.49	.42	.12	.38	.57	.46	.18	.47	.37	.33	.26	.37	.44	.38	.57	.47	.32	.27	.34	
CT_2	.40	.49	.40	.16	.30	.50	.33	.14	.43	.34	.23	.19	.38	.40	.31	.51	.39	.40	.39	.26	
LA_2	.51	.55	.54	.18	.66	.69	.62	.28	.68	.57	.72	.43	.53	.61	.46	.60	.29	.38	.14	.23	
TE_2	.53	.58	.56	.21	.61	.74	.57	.13	.67	.69	.63	.47	.46	.55	.50	.64	.24	.28	.10	.30	
QS_2	.52	.60	.46	.11	.59	.73	.70	.11	.61	.52	.47	.30	.61	.71	.46	.65	.30	.31	.20	.28	
CH_2	.53	.63	.49	.13	.46	.67	.52	.13	.58	.49	.39	.29	.48	.53	.45	.67	.55	.38	.33	.35	
AI_2	.59	.63	.61	.29	.59	.72	.59	.12	.69	.66	.52	.42	.58	.64	.60	.77	.33	.43	.27	.30	
OE_2	.53	.61	.50	.21	.49	.68	.51	.07	.60	.65	.44	.41	.43	.50	.47	.61	.21	.26	.12	.37	
IL_2	.17	.16	.08	.45	.06	.06	.05	.35	.20	.17	.05	.49	.11	.14	.13	.11	.00	.01	-.07	.34	
EL_2																					

Note. AG = Agency; PATH = Pathways Thinking; OPT = Optimism; PESS = Pessimism; FOCUS = Focus; GOALSR = Goal Self-Regulation; EMOTSR = Emotion Self-Regulation; COI = Consistency of Interest; POE = Perseverance of Effort; BTH = Booster Thoughts; BBH = Booster Behaviors; HAMP = Hampering; AR = Academic Resilience; LR = Life Resilience; SWL = Life Satisfaction; WB = Wellbeing; PAB = Physical Abilities Self-Concept; PAP = Physical Appearance Self-Concept; OPSX = Opposite-SEX Relationships Self-Concept; SMSX = Same-Sex Relationships Self-Concept; PR = Parent Relationships Self-Concept; HO = Honesty-Trustworthiness Self-Concept; ES = Emotional Stability Self-Concept; MH = Math Self-Concept; VER = Verbal Self-Concept; SCH = School Self-Concept; GS = General Self-Esteem/Self-Concept; SC = Self-Confidence; SF = Self-Efficacy; SM = Stress Management; OT = Open Thinking; SE = Social Effectiveness; CT = Cooperative Teamwork; LA = Leadership Ability; TE = Time Efficiency; QS = Quality Seeking; CH = Coping with Change; AI = Active Involvement; OE = Overall Effectiveness; IL = Internal Locus of Control; EL = External Locus of Control. _2 refers to data collected immediately post-program (T2).

(continues)

Table 5.42 (continued)
Heterotrait-Monomethod Correlations at T2

	PR_2	HO_2	ES_2	MH_2	VER_2	SCH_2	GS_2	SC_2	SF_2	SM_2	OT_2	SE_2	CT_2	LA_2	TE_2	QS_2	CH_2	AI_2	OE_2	IL_2	EL_2	
AG_2																						
PATH_2																						
OPT_2																						
PESS_2																						
FOCUS_2																						
GOALSR_2																						
EMOTSR_2																						
COI_2																						
POE_2																						
BTH_2																						
BBH_2																						
HAMP_2																						
AR_2																						
LR_2																						
SWL_2																						
WB_2																						
PAB_2																						
PAP_2																						
OPSX_2																						
SMSX_2																						
PR_2																						
HO_2	.35																					
ES_2	.23	.30																				
MH_2	.05	.10	.13																			
VER_2	.19	.23	.13	-.10																		
SCH_2	.33	.30	.22	.60	.52																	
GS_2	.51	.21	.21	.33	.30	.80																
SC_2	.45	.25	.22	.23	.34	.65	.83															
SF_2	.37	.20	.29	.22	.24	.52	.71	.78														
SM_2	.34	.20	.34	.17	.18	.44	.62	.65	.89													
OT_2	.38	.24	.19	.20	.30	.60	.75	.84	.79	.72												
SE_2	.30	.26	.17	.05	.33	.36	.49	.66	.53	.49	.58											
CT_2	.37	.26	.18	.07	.19	.37	.55	.62	.51	.54	.68	.65										
LA_2	.32	.30	.19	.06	.36	.42	.51	.65	.45	.32	.60	.72	.65									
TE_2	.37	.33	.21	.23	.28	.54	.63	.70	.67	.63	.63	.43	.45	.42								
QS_2	.43	.24	.08	.26	.26	.61	.75	.85	.63	.51	.78	.42	.57	.42	.73							
CH_2	.35	.26	.24	.21	.21	.51	.67	.70	.84	.86	.80	.58	.60	.45	.67	.60						
AI_2	.37	.27	.13	.13	.27	.50	.68	.79	.61	.55	.86	.71	.81	.73	.57	.69	.67					
OE_2	.44	.31	.23	.27	.37	.69	.83	.91	.78	.64	.80	.60	.56	.62	.74	.80	.73	.71				
IL_2	.40	.18	.12	.21	.28	.59	.71	.83	.61	.49	.87	.45	.57	.49	.53	.87	.58	.72	.73			
EL_2	.24	.27	.35	.07	.10	.20	.15	.15	.00	.01	.10	-.01	.16	.09	.06	.15	.07	.06	.11	.25		

Note. AG = Agency; PATH = Pathways Thinking; OPT = Optimism; PESS = Pessimism; FOCUS = Focus; GOALSR = Goal Self-Regulation; EMOTSR = Emotion Self-Regulation; COI = Consistency of Interest; POE = Perseverance of Effort; BTH = Booster Thoughts; BBH = Booster Behaviors; HAMP = Hampering; AR = Academic Resilience; LR = Life Resilience; SWL = Life Satisfaction; WB = Wellbeing; PAB = Physical Abilities Self-Concept; PAP = Physical Appearance Self-Concept; OPSX = Opposite-SEX Relationships Self-Concept; SMSX = Same-Sex Relationships Self-Concept; PR = Parent Relationships Self-Concept; HO = Honesty-Trustworthiness Self-Concept; ES = Emotional Stability Self-Concept; MH = Math Self-Concept; VER = Verbal Self-Concept; SCH = School Self-Concept; GS = General Self-Esteem/Self-Concept; SC = Self-Confidence; SF = Self-Efficacy; SM = Stress Management; OT = Open Thinking; SE = Social Effectiveness; CT = Cooperative Teamwork; LA = Leadership Ability; TE = Time Efficiency; QS = Quality Seeking; CH = Coping with Change; AI = Active Involvement; OE = Overall Effectiveness; IL = Internal Locus of Control; EL = External Locus of Control. _2 refers to data collected immediately post-program (T2).

Comparison was also made between the HTMM correlations and HTHM correlations. Having HTMM correlations which are systematically higher than HTHM correlations indicates that there may be method effects associated with the specific occasion of data collection. Results from the MTMM matrix found both the T1 HTMM correlations (mean $|r| = .38$, $SD = .21$) and the T2 HTMM correlations (mean $|r| = .39$, $SD = .21$) to be higher than the HTHM correlations (mean $|r| = .27$, $SD = .15$), suggesting that there may be some amount of method effect associated with the Survey administration at T1 and T2.

Conclusion. With regard to the Campbell-Fiske (1959) criteria, there is support for the convergent and discriminant validity of each of the Survey scales, as well as test-retest stability over time. In particular, all 41 convergent validities were statistically significant and substantial, and support for the two criteria of discriminant validity was met for 3,270 of the 3,281 comparisons. However, comparisons between the HTMM correlations and the HTHM correlations indicated some method effect may be associated with the data assessment at T1 and T2. Finally, investigation of the HTMM correlations revealed the ROPELOC factors to be less well differentiated than the other factors.

Summary

This study aims to establish the robustness of each of the scales that makes up the Survey instrument, and the integrity of the data derived from the administrations of the Survey to the participants (referred to in subsequent chapters as *assessment waves*). To meet this aim, the psychometric properties of each of the scales that form part of the Survey were investigated. Reliability testing found the majority of the scales to have omega and alpha estimates that exceed the recommended threshold of .70. Of the scales that did not reach this threshold (CHS scales, LOT-R scales, Grit-S scales, MES-S scales(except Booster Behaviours), ASRI-R's Emotion Self-Regulation, and ROPELOC's Active Involvement), some had reliability estimates above .60, a value which is still considered acceptable, particularly in the context of psychological constructs (Aron, Coups, & Aron, 2013). In addition, inter-item correlations for these scales were investigated and found to

be acceptable with the exception of the Agency, Emotion Self-Regulation, and negatively-worded MES scales.

Factor analysis started with the a priori factor structure for each measurement instrument modelled as a CFA, while also considering an ESEM. Most of these a priori factor structures were found to have an appropriate fit to the data with either a CFA or ESEM. However, the CHS factor structure required modification to allow one Agency item (Ag3) to cross-load onto the Pathways Thinking factor, as had been foreshadowed by the reliability assessments. In addition, the a priori factor structure of the ASRI was not supported by the data, whether modelled as a CFA or as an ESEM. Based on self-regulation theory and the goals of the THP program, 15 items from the ASRI were selected for a revised measure referred to as ASRI-R, and a three-factor ESEM fit this data well. Further, the a priori factor structure of the MES-S had poor fit when modelled as a CFA, and an ESEM would not converge. An EFA revealed a three-factor solution to be the best fit, with the negatively-worded items loading onto a single factor. This revised structure exhibited good fit to the data when modelled as an ESEM.

All measurement instruments evidenced longitudinal invariance with the exception of the CHS and SWLS, which both achieved partial scalar invariance. Finally, an MTMM matrix analysis based on factor scores for the scales evidenced support for both convergent and discriminant validity, as well as test-retest stabilities, across each of the scales. Factor scores for each scale derived from the best fitting factor analytic models will provide the data for the analyses to be undertaken in Study 2 (see Chapter Six).

Discussion

Strengths

Much of the existing OAE and coaching research does not reassess the psychometric properties of its outcome measures with its data, choosing to rely instead on pre-existing validation evidence. However, the validity of measurement instruments depend as much on the data collected (including the nature of the participants and the measurement protocols) as on the instruments themselves (Marsh & Hau, 2007). Accordingly, it is best practice to reassess the psychometric

properties of measurement instruments with each use. Specific appeals have been made for quantitative research in OAE to be methodologically rigorous (Cason & Gillis, 1994; Hattie et al., 1997; Neill, 2008; Scrutton & Beames, 2015). One of the central strengths of the present investigation is found in the rigorous statistical procedures used to test the psychometric properties of the measurement instruments used in the Survey. Coefficient omega was used in addition to coefficient alpha to assess internal consistency. ESEM was used as an alternative to modification indices for model refinement and then compared with the more parsimonious CFA model. The factor analyses took advantage of long form data to increase statistical power, while also accounting for the non-independence of the observations inherent in long form data. A MTMM matrix with time as the multiple methods was used to assess construct validity and test/re-test stability of the scales. In reporting results, confidence intervals were provided in addition to point estimates, consistent with best practice.

The findings from this study contribute to the existing validation research into the 11 instruments used in the research for this thesis. These findings also extend previous validation research on these instruments to disadvantaged adolescents. While many of the instruments have been developed or used for a youth population, it does not appear that many have been assessed with a disadvantaged adolescent population. Furthermore, these findings also extend prior validation research beyond the traditional measurement properties of validity and reliability by assessing longitudinal invariance for each instrument. Although many of the Survey instruments are the subject of prior validation research, many of them have not been assessed previously for invariance, particularly over time. This analysis, therefore, provides additional research data on the psychometric properties of these instruments for these data.

Limitations

Some limitations should be considered when interpreting the results. First, the presentation of omega estimates did not establish reliability for each scale. While a measure's reliability is said to be an important first step in assessing research results (Furr, 2011; Kline, 2005), reliability estimates assume, rather than

confirm, the unidimensionality of the scales. Factor analysis is said to be the best way in which to confirm how well a set of items represents a single underlying construct. For this reason, the psychometric evaluations conducted in this study rely more heavily on the factor analyses than the reliability estimates. Furthermore, to the extent that these results indicate a concern with the amount of error being captured as part of the measurement for a scale, the further analyses to be undertaken in Study 2 use methods that control for variance due to error.

Second, while the factor analyses supported the a priori hypothesised structure of the majority of the measurement instruments, three of the instruments required structural adjustment (CHS, ASRI, MES-S), and one instrument was modified also at the item level (ASRI). However, the modifications to the CHS were consistent with previous findings (Edwards, Ong, et al., 2007). Moreover, the ASRI is a relatively new measure that was not supported by strong psychometric research. A shorter, modified version of the ASRI was created that better aligns with the THP program aims and demonstrates sound psychometric properties. Furthermore, while the MES-S was a short form of an instrument with strong psychometric properties, the MES-S had not been previously validated. However, it was possible to retain all of the MES-S items using a slightly modified factor structure. In each case, the changes made theoretical sense and were not performed simply to yield a better fitting model. Moreover, the revised ASRI and MES-S, being shorter forms of their original instruments, may provide useful alternatives for research in the future. As it is generally not recommended to use the same data to both develop a model and evaluate its fit, ideally the modified models should be tested on another sample of data (Breckler, 1990; Cudeck & Browne, 1983). In this case, retesting with another sample was not possible given the statistical analyses were already under strain due to sample size. Consequently, retesting these modified factor models is something that should be considered for future research.

Third, while a number of the models tested for longitudinal invariance had acceptable results, the fit statistics for some of the configural models were substantially lower than the corresponding fit statistics found in the factor analysis. While this can be attributed at least in part to the smaller sample size and increased

complexity of these models, this is something that should be reconsidered in future research with larger sample sizes.

Finally, the items from each of the 41 scales were mixed together and administered to the participants in a single instrument (rather than as intact measures). Doing so resulted in modifications to the instruments from their original form, including for some of the instruments, the response wording, loss of opening wording setting the context of the measurement items, and other subtle changes to the wording of individual items. As a consequence, it may be difficult to compare the results of this study with other validation research on particular measurement instruments. Furthermore, it is best practice to assess the structural integrity of the Survey by modelling together all of the individual factor analytic models for the instruments. However, such a model would not converge given the number of items and scales, as well as model complexity, in comparison to the amount of available data. While such a model provides the opportunity to detect overlap among the scales, analysis of the MTMM matrix allows for similar comparisons to be made.

Conclusion

The aim of Study 1 was to assess the psychometric properties of each of the measurement instruments and their scales included in the Survey, based on the Survey data. Overall, the measures demonstrated sound psychometric properties. Where they initially did not evidence acceptable results, modifications that were theoretically supported were made, and these adjusted measures showed acceptable psychometric properties. Notwithstanding these results, it is important to note that establishing the robustness of a measurement instrument is an ongoing process. Study 2 examines the relations among factors and will provide additional information about the validity of the measurement instruments. Nonetheless, the results of this study lay a solid foundation for the analyses to be undertaken in Study 2.

CHAPTER SIX

STUDY 2: QUANTITATIVE INVESTIGATION OF THP PROGRAM OUTCOMES

Introduction

The Helmsman Project (THP) provides a novel program for socioeconomically disadvantaged adolescents that aims to foster personal and social development by cultivating a range of skills and qualities, including hope, resilience, and self-regulation. The ultimate goals of the program are to improve the participants' educational engagement and wellbeing, thereby assisting participants to flourish and reach their full potential.

As detailed in Chapter Four, the primary program offered by THP integrates developmental coaching with two outdoor adventure experiences (Adventure Program). This program is grounded in the theory and research related to outdoor adventure education (OAE). Further detail on the theoretical underpinnings of the program can be found in Chapter Two. OAE programs, in general, come in many forms and serve a variety of populations with diverse aims. While the research findings on the impact of OAE programs have been largely positive, much of the research has been found to be methodologically weak, owing to a lack of control groups, lack of randomisation, and lack of longitudinal follow-up (Cason & Gillis, 1994; Hattie et al., 1997; Neill, 2008; Scrutton & Beames, 2015). Moreover, research has found only some programs to be effective and then only in relation to some outcomes (Hattie et al., 1997; Sheard & Golby, 2006). OAE researchers and practitioners argue that what is needed is a better understanding of which program variables and theoretical elements are most related to positive program outcomes (Hans, 2000; McKenzie, 2000; Neill, 2008; Sheard & Golby, 2006; Sibthorp & Arthur-Banning, 2004).

As mentioned in Chapter Two, the program variable that has been given the most attention by researchers is program facilitation. In particular, skilled facilitators have been found to be of primary importance to the experiential

learning process and critical to successful outcomes (Ewert & McAvoy, 2000; Gass & Gillis, 1995; Hattie et al., 1997; Kemp, 2006; Luckner & Nadler, 1997; Martin & Leberman, 2005; Martin & Legg, 2002; McKenzie, 2000; Sibthorp et al., 2011; Sibthorp et al., 2007). Each of the outdoor adventure components of an Adventure Program is facilitated by providers experienced in delivering the relevant mode of adventure. In addition, THP explicitly incorporated developmental coaching into its program with the intention of enhancing the experiential learning process and overall impact on outcomes, resulting in a novel type of OAE program with a reduced role for program facilitators found in traditional OAE programs. With coaching psychology being a relatively recent field of study, the research on coaching efficacy is limited and what exists is similarly plagued by a lack of both experimental design and methodologically rigorous analysis (Grant, 2012a; Grant & Cavanagh, 2007; Grant et al., 2010; Lowman, 2005). The emerging research that does exist to support the effectiveness of coaching is particularly focused on executive and life coaching for adults. Accordingly, more research is required with respect to developmental coaching for adolescents.

The Adventure Program is an OAE program for adolescents that focuses on outcomes related to the personal and social development of its participants. In addition to the novel developmental coaching component and skilled facilitation, the program has a number of key elements based on prior research and theory: (a) an experiential learning framework that provides opportunities for experimentation and reflection; (b) experimentation through a structure of focused goal setting and goal striving; and (c) controlled exposure to challenge, together with appropriate support. While these key elements form part of each Adventure Program, the programs differ in the type of adventure experience offered, as well as the size of the adventure experience group. Two of the adventure experiences are sailing-based: one being a small-sized group on a yacht known as the *Arctos*, and the other being a large-sized group on a tall ship known as the *James Craig*. The third adventure experience is a medium-sized group, land-based adventure with Outward Bound Australia (referred to in this thesis as Outward Bound or OB). Further details on each of these programs is included in Appendix B. Aside from the adventure experiences, each of the Adventure Programs was conducted through a group of 8-

10 participants. A small-sized group was considered important given the individualisation in the program focus, as well as the theories relating smaller-sized groups to stronger program outcomes (McKenzie, 2000, 2003; Walsh & Golins, 1976).

For research purposes, THP also has provided a program that is based solely on developmental coaching and skill development, without the outdoor adventure experiences (Coaching Only Program). The Coaching Only Program otherwise incorporates the same elements as described above for the Adventure Programs. For a detailed comparison of the Adventure Programs with the Coaching Only Program, see Appendix C. The principal purpose of initiating the Coaching Only Program was to examine the incremental benefits that the outdoor adventure experiences provide for the THP program.

After presenting a broad overview of the research aims underlying this study, the research hypotheses and questions are outlined, including the rationale for those research hypotheses and questions. The research hypotheses refer to the a priori predictions about the outcomes of the research that will be tested with the participant data. Where there is no strong empirical basis for making a prediction about an outcome, a research question is posed instead. Following the research aims, hypotheses, and questions, the specific methodology and procedures used for this study are then elaborated on, after which the results are presented in the order of the research hypotheses and questions. These results are then briefly discussed before summarising the study. A more detailed discussion of the complete results for this thesis are included in Chapter Eight.

Research Aims, Hypotheses, and Questions

Research Aims

The principal aim of this study is to evaluate the quantitative effects of the novel Adventure Programs on a range of outcomes, using a randomised controlled trial (RCT) and rigorous statistical analysis. A further aim is to compare the individual effects of the different Adventure Programs, as well as to compare the effects of the Adventure Programs (taken together) with the Coaching Only

Program. This investigation into the various effects of the THP programs will cover the following primary research aims:

1. *Experimental effects of THP programs on outcome variables*: to examine the short-term, long-term, and follow-up effects of the various THP programs on participants' hope, self-regulation, resilience, motivation, wellbeing, self-concept, and life effectiveness skills, compared to the waitlist control group,²⁴ including whether there are any interaction effects between pre-intervention (T1) aptitude in the outcome variable of interest and participation in a THP program, and whether the relevance of the outcome variable to the program design was related to program impact on that outcome variable;
2. *Differences in effects between Adventure Program and Coaching Only Program*: to examine any differences in experimental effects between the Adventure Programs (taken together) when compared to the Coaching Only Program; and
3. *Alternative replication of short-term effects of THP program on outcome variables*: to consider the replicability of the results using pre-post within-subjects comparisons of the extended baseline pre-test data from the control group (T1-T3) with the immediate post-test data from the control group following participation in a THP program (T4).

Research Hypotheses and Questions: Statement and Rationale

The research hypotheses and questions in this study are based on existing theory and research in connection with each of the outcome variables, as well as OAE and coaching psychology more generally. For ease of reference, research hypotheses and questions appear in the order most appropriate to their analysis, and they are numbered sequentially beginning with the number of this study.

²⁴ For ease of reference, the waitlist control group may be referred to in this thesis as the *control group*.

Research Hypothesis 2.1: Short-term effects of THP programs on outcome variables.

Research Hypothesis 2.1.1: Positive short-term effects of THP programs on all outcome variables. Existing research has demonstrated OAE and coaching psychology separately to have positive effects on various aspects of adolescents' personal and social development (see Chapter Two). Both the coaching and outdoor adventure components of the THP program seek to integrate existing evidence-based research and practice from these fields. Consequently, it was expected that the Adventure Programs (both individually and taken together), as well as the Coaching Only Program, would demonstrate direct positive experimental effects on the 41 outcome variables, immediately following completion of the program (T2). This hypothesis will be evidenced in the RCT analysis by statistically significant higher factor scores of the program participants at T2 for the outcome measures when compared to those factor scores for the control group, while controlling for any differences in the outcome variables prior to the program (T1) and certain other individual differences (for more information on these individual differences, refer to the section below headed "Additional covariates" in "Methodology and Procedures").

Research Hypothesis 2.1.2: Greater positive short-term effects of THP programs on outcome variables with most relevance to THP programs.

Previous research has found program effects across a range of outcome measures to be larger for those scales that were most relevant to the goals of the program (Marsh et al., 1986a, 1986b). Given the breadth of measurement scales included in the research, it was hypothesised that the Adventure Programs (both individually and taken together), as well as the Coaching Only Program, would demonstrate greater direct positive experimental effects on those outcome variables with the most relevance to the THP program design and aims. This hypothesis will be evidenced in the RCT analysis by comparison of effects of program participation at immediate post-test (T2) when compared to the control group, for each group of scales categorised by their relevance to the program design and aims (high, moderate, and low; for more information on the method of scale relevance rating,

refer to the section below headed “Scale relevance” in “Methodology and Procedures”). In addition, consideration will be given to the effect sizes of program participation at immediate post-test (T2), when compared to the control group, in relation to the relevance ratings for the scales (high = 1, moderate = 2, and 3 = low).

Research Hypothesis 2.2: Long-term maintenance of positive effects of THP programs on outcome variables. Although there is a call for more longitudinal data in OAE research, existing OAE research and theory suggest that the positive developmental outcomes arising from OAE programs can endure (Bowen & Neill, 2013; Gillis & Speelman, 2008; Hattie et al., 1997; Neill, 2008). Accordingly, it was hypothesised that the Adventure Programs (both individually and taken together), as well as the THP Coaching Only Program, would continue to demonstrate direct positive experimental effects on the outcome variables approximately three months following completion of the program (T3). This hypothesis will be evidenced in the RCT analysis by statistically significant higher factor scores of the program participants at T3 for the outcome variables when compared to those factor scores for the control group, controlling for any differences in the outcome variables prior to the program (T1) and certain other individual differences.

Research Question 2.3: Follow-up effects of THP programs on outcome variables. Although longitudinal OAE research is limited, there has been meta-analytic evidence to suggest that participants may experience additional growth following an OAE experience (Hattie et al., 1997). It also has been suggested that given the deeply reflective nature of coaching, some effects may not emerge until after coaching has concluded (Spence et al., 2019). On the other hand, for many extracurricular programs, including OAE, the more common result is a loss of benefits once the program has finished (see, e.g., Gillis & Speelman, 2008; Hatch & McCarthy, 2005; Neill, 2008; Schary, Wozniak, Jenny, & Morrow, 2016). These results may be further confounded by effects such as post-group euphoria or incomplete transfer of learning. However, the design of the THP program was intended to facilitate the transfer of learning through intact groups and the opportunity for a post-program project.

Owing to the lack of follow-up research in this area and the novel program design, there were no specific hypotheses about the specific experimental effects during the three-month period following program completion. Instead, the following research question is posed: To what extent do the THP programs demonstrate any new or additional significant positive experimental effects (known as *sleeper effects*), or dissipation or maintenance of short-term experimental effects, on the outcome variables between completion of the program (T2) and the date three-months after completion of the program (T3)? This research question will be evaluated by considering the factor scores of the program participants for the outcome variables at T3 when compared to those factor scores for the control group, controlling for any differences in the outcome variables both prior to the program (T1) and immediately following program completion (T2), as well as certain other individual differences.

Research Hypothesis 2.4: Aptitude-treatment interaction effects.

There is some limited research to indicate that participation in extracurricular activities benefits disadvantaged students more than students from higher socioeconomic backgrounds (Marsh, 1992a; Marsh & Kleitman, 2002). Given that the THP program was designed expressly for the purpose of improving outcomes for disadvantaged students, it was hypothesised that those students who began a THP program with lower baseline levels (also referred to as *aptitude*) of an outcome variable of interest would benefit more from the THP program in connection with that outcome variable (both in the short and longer term) than participants who started the program at or above the average baseline level of that outcome variable. This hypothesis will be evidenced in the RCT analysis by statistically significant aptitude-treatment interaction effects between the main effect of group (THP program vs. control) and the participants' level of the outcome variable at T1 (the moderating variable), together with significant positive simple main effects for low-aptitude participants (for more information on the interaction analysis methodology, refer to the section below headed "Interaction effects" in "Methodology and Procedures"). Consideration will be given to aptitude-treatment interactions in both the short- and long-term analyses.

Research Question 2.5: Differences in effects between Adventure Programs and Coaching Only Program. Research indicates that both OAE and coaching are associated with benefits for the personal and social development of adolescents (see Chapter Two). While average effects for OAE programs have been found comparable to other types of programs aimed at enhancing participants' self-perceptions and other outcomes (Hattie et al., 1997; Neill, 2008; Neill & Richards, 1998), there is no research from a single study that considers the incremental benefit of the outdoor adventure experience that is an inherent element of OAE. Thus, rather than hypothesising about the relative effects of these programs, the following research question is posed: To what extent do the Adventure Programs (taken together) demonstrate statistically significant differences in their short-term, long-term, and follow-up experimental effects on the outcome variables, that favour the Adventure Programs when compared to those effects generated by the Coaching Only Program? This research question will be evaluated by considering whether there are statistically significant higher experimental effect sizes for the Adventure Programs for the outcome variables when compared to those effect sizes for the Coaching Only Program.

Research Question 2.6: Replication of RCT results with waitlist control group data. Data collected from the waitlist control group after their participation in a THP program provides an opportunity to consider the short-term effect of the THP program using an alternative within-subjects design. The results of the within-subjects pre-post analysis can be compared to the results of the RCT analysis in order to consider the replicability of those results. Given the variability in OAE outcomes that result in part from the range of research designs and methods (Cason & Gillis, 1994), the following research question is posed: To what extent do the within-subjects comparisons of the waitlist control group's extended baseline data with their immediate post-test data (T4), replicate the results from the RCT short-term analysis? This research question will be evaluated by considering significant results of the within-subjects pre-post analysis against the significant experimental effects of the RCT analysis.

Summary

While research evidence on OAE and coaching programs suggests each has a role to play in developing important life skills for disadvantaged adolescents that may improve educational outcomes, more evidence is required to understand what aspects of these programs are most effective and for what outcomes. Additionally, the research that exists is undermined by criticisms of a lack of methodological rigour. The THP program has been grounded in existing research and theory in OAE and coaching psychology, and the research design with an RCT and longitudinal follow-up meets some of the criticisms of existing research. The additional data from the control group participants when they subsequently participate in a THP program allows for evaluation of the replicability of the RCT results using a within-subjects pre-post design. Accordingly, the research aims, hypotheses, and questions in this study address substantial gaps in both the OAE and coaching psychology literature. Moreover, THP's unique Adventure Programs that integrate OAE and developmental coaching provide an opportunity to test a novel approach, as well as to examine the incremental benefits that the outdoor adventure component provides for the THP program. The next section outlines the methodology and procedures used in this study, with results presented in the section that follows.

Methodology and Procedures

Participants

The sample for this study consisted of 362 students in Year 9, 51.7% of which were female. Each participant had the opportunity to take part in a THP program, either as a member of the intervention group or through the waitlist control group. Table 6.1 provides a summary overview of the different Adventure Program modes, as well as the Coaching Only Program. These programs and the recruitment process are described in detail in Chapter Four. Participants were from 11 high schools all located in Western Sydney, an area with high rates of disadvantage. Self-reported demographic information on these participants is included in Appendix E.

Table 6.1
 THP Program Overview by Program Mode

	Adventure Programs			Coaching Only Program
	Arctos	James Craig	Outward Bound	
Adventure Type	Small yacht	Tall ship	Hiking	No adventure
Experiences	Sail boat independently& plan/ prepare meals	Assist to sail boat, group tasks, non-sailing activities	Hiking, rock climbing, abseiling, meal prep, survival skills	Presentation skills seminar, individual project
Program Group Size	8-10	8-10	8-10	8-10
Experience Group Size	8-10	40	16	8-10
Experience Inter-School	No	Yes	Yes	No
Framework	Experiential learning cycle			
Support	Developmental coaching			
Other Opportunities	Community Project			

Note. THP = The Helmsman Project. Adventure Programs refer to those THP programs with an adventure component and Coaching Only Program refers to the THP program without an adventure component. Arctos = Arctos Adventure Program; James Craig = James Craig Adventure Program; Outward Bound = Outward Bound Adventure Program.

Measures

For the research in this study, participants completed a self-report questionnaire (Survey). The Survey includes items from 11 instruments measuring a range of outcomes. Background to the constructs and scales underlying the measurement instruments is provided in Chapter Three, and a detailed analysis of the psychometric properties of those measures is the subject of Chapter Five. For the purposes of this study, participants in the intervention group completed the Survey just prior to entry into a THP program (T1), immediately after completion of the program (T2), and approximately three months after program completion (T3). Participants in the waitlist control group completed the Survey together with their corresponding intervention group at T1, T2, and T3 (extended baseline), as well as when they completed a THP program (T4).

Research Design

This study proposed a rigorous experimental design with an RCT using a waitlist control group and repeated measures. Having a control group reduces certain threats to the validity of results by mitigating potential threats such as maturation, familiarity with the measurement instrument, and regression toward the mean over time (Field et al., 2012). The use of a waitlist control group ensures all participants receive the benefit of the intervention, thus overcoming critical ethical issues in RCT design, particularly in educational settings. The data from the waitlist control group is used as a basis of comparison with the intervention group's data from one pre-test (T1) and two post-tests (T2 and T3). The pre-test provides a baseline by which to measure and analyse change over time. A high-level diagram of the overall research design is set out in Figure 6.1 (for more information on the differences among the various THP programs, see Chapter Four and the related appendices).

The waitlist control group's data (T1, T2, and T3) also serves as an extended baseline against which to compare later results for the control group when they subsequently experience a THP program (T4). Therefore, the study includes both RCT between-group comparisons between the intervention and control groups, as well as alternative pre-post within-subjects comparisons of the control group, providing an opportunity to test the replicability of the RCT results with an alternative experimental design.

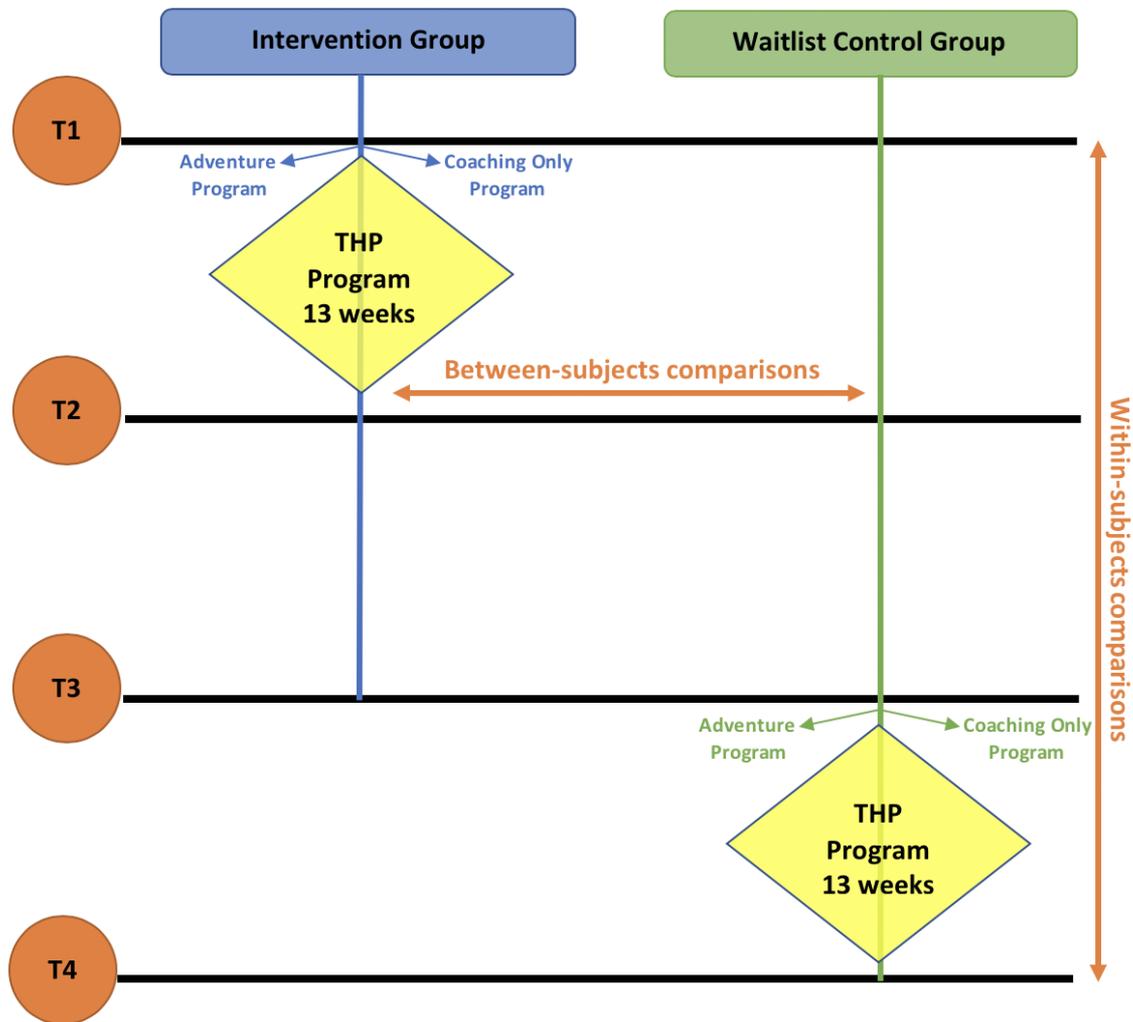


Figure 6.1. Research design for Study 2.

Note. For the intervention group, T1 = pre-test; T2 = immediate post-test; T3 = three months post-test. For the waitlist control group, T1, T2, and T3 = extended baseline pre-test (collected at the same time as intervention group T1, T2, and T3) and T4 = immediate post-test.

Randomisation

RCTs provide the most reliable evidence for evaluating the efficacy of an intervention. Accordingly, this research proposed an RCT research design. In RCTs, the random assignment of participants to the intervention and control groups allows for the strongest causal inferences by reducing unsystematic variation as much as possible (Field et al., 2012). Accordingly, it is important to provide details of the method of random assignment of participants in an RCT design.

For this research study, given the overnight nature of the adventure components of the Adventure Programs, schools required that program groups be single-gender, and this requirement impacted student allocation between groups

and the matching of intervention and control groups across gender. Moreover, each school in the program allocated a teacher to support each program group (Cohort Teacher), and schools also required that the gender of the Cohort Teacher match the gender of the program group. As a result, the gender of the Cohort Teacher sometimes influenced the allocation of students to one group or another. However, the schools who made the group allocations were at all times unaware of which group represented the intervention or control group for research purposes. The program type to which program groups were allocated at times was impacted also by the availability of the particular adventure experience. Consequently, while students of each gender were allocated to both intervention and control groups, some of the program types ended up with unbalanced gender numbers. Table 6.2 provides a breakdown of the participants by group and gender. To prevent confounding consequences of the gender imbalances, the effect of gender was controlled for in the analyses. A number of additional covariates were also included to control for other pre-treatment variables, and these are described in more detail below (see the sections headed “Pre/Post intervention effects” and “Additional covariates”).

Table 6.2
Breakdown of Participants by Group and Gender

	Arctos	James Craig	Outward Bound	Coaching Only	Adv (Total)	Total
Total Participants	110	101	89	62	300	362
Male	33	68	44	30	145	175
Female	77	33	45	32	155	187
Intervention	59	54	45	28	158	186
Intervention Male	6	38	16	12	60	72
Intervention Female	53	16	29	16	98	114
Control	51	47	44	34	142	176
Control Male	27	30	28	18	85	103
Control Female	24	17	16	16	57	73

Note. Arctos = Arctos Adventure Program; James Craig = James Craig Adventure Program; Outward Bound = Outward Bound Adventure Program; Coaching Only = Coaching Only Program; Adv = Adventure Programs (taken together); Control = waitlist control group.

Data Analysis

Treatment of missing data. The types of missing data and the common approaches to dealing with that missing data, have been outlined in detail in Chapter Four. As is common in a study involving students and a repeated measures design, many participants had missing data for an entire assessment wave, due primarily to absence from school, change of school, withdrawal from a program, inability to collect data at a school for a wave, or in the case of a waitlist control participant, late admission to the waitlist. Table 6.3 shows the percentage of participants missing data for all items at each assessment wave relevant for this study (T1 -T4), broken down by program group and between intervention and control participants.²⁵ While there were participants in attendance for administration of a Survey who did not provide complete responses to that Survey, all partially completed Surveys were at least 95% complete and considered sufficient for inclusion in this study.

²⁵ Due to the extent of missing data at T5, this wave of data was not included in the analyses for Study 2.

Table 6.3
 Percentage of Participants Missing All Data at each Wave: T1 to T4

Wave	Group	Participants (N)	Missing All Data	
			Number	Percentage
T1	Intervention	186	7	3.76%
	Arctos	59	2	3.39%
	James Craig	54	2	3.70%
	Outward Bound	45	1	2.22%
	Coaching Only	28	2	7.14%
	Control	176	36	20.45%
	Arctos	51	10	19.61%
	James Craig	47	7	14.89%
	Outward Bound	44	11	25.00%
	Coaching Only	34	8	23.53%
T2	Intervention	186	38	20.43%
	Arctos	59	7	11.86%
	James Craig	54	14	25.93%
	Outward Bound	45	9	20.00%
	Coaching Only	28	8	28.57%
	Control	176	32	18.18%
	Arctos	51	6	11.76%
	James Craig	47	8	17.02%
	Outward Bound	44	7	15.91%
	Coaching Only	34	11	32.35%
T3	Intervention	186	42	22.58%
	Arctos	59	11	18.64%
	James Craig	54	20	37.04%
	Outward Bound	45	7	15.56%
	Coaching Only	28	4	14.29%
	Control	176	27	15.34%
	Arctos	51	6	11.76%
	James Craig	47	7	14.89%
	Outward Bound	44	6	13.64%
	Coaching Only	34	8	23.53%
T4	Control	176	76	43.18%
	Arctos	51	23	45.10%
	James Craig	47	20	42.55%
	Outward Bound	44	17	38.64%
	Coaching Only	34	16	47.06%

Note. Arctos = Arctos Adventure Program; James Craig = James Craig Adventure Program; Outward Bound = Outward Bound Adventure Program; Coaching Only = Coaching Only Program (without adventure); and Control = waitlist control group. For intervention participants, T1 = pre-test; T2 = immediate post-test; T3 = three months post-test; and for Control, T1, T2, and T3 = extended baseline pre-test (assessed with intervention participants, with T3 being immediate pre-test); T4 = immediate post-test. T4 data was collected only from the waitlist control group.

As indicated in Chapter Four, because the missing data for this Survey is either *missing completely at random* (MCAR) or *missing at random* (MAR), the missing data was handled primarily by using *full information maximum likelihood* (FIML).²⁶ In most cases, FIML is regarded to be the best and most appropriate way to handle missing data in structural equation modelling (SEM; Allison, 2003; Brown, 2015), including where there is a large amount of missing data (Enders, 2010).

Negatively-worded items. Some of the Survey scales include negatively-worded items, either in whole or in part. For consistency, all of these items were reverse-scored prior to conducting the factor analysis and calculating factor scores in the psychometric analysis undertaken for Chapter Five. Accordingly, larger positive effect sizes indicate favourable effects across all latent outcome variables (for example, a larger positive effect size for Pessimism indicates a decrease in pessimism).

Scale relevance. The research on which this thesis is based emanates from a research project managed by the Institute for Positive Psychology and Education, at the Australian Catholic University, in partnership with THP and co-funded by the Australian Research Council. The research team, led by Professor Herbert Marsh, in conjunction with THP staff, selected a comprehensive set of research outcome constructs and measurement instruments based on previous theory and research related to the broad aims of THP at inception. In accordance with previous recommendations, attention was paid not only to primary outcomes, but also to anticipated secondary and other consequential outcomes (Ewert & McAvoy, 2000). Moreover, the research team was made up of a large number of partner investigators, each with their own theoretical perspectives. Owing to the nature of the study, it was decided to include in the research each of the scales recommended by each partner investigator.

Owing to the large number of scales in the research and as a way to focus the analysis, each scale was formally reviewed in the context of the ultimate design of

²⁶ FIML is also referred to as *maximum likelihood* and *direct maximum likelihood* in the literature (Enders, 2010).

the THP program and the explicit aims of the coaching and adventure experiences developed as part of the program. This approach is consistent with the recommendations of OAE researchers, who suggest that program outcomes be specifically related to program goals and objectives (Hattie et al., 1997). Using a quasi-Delphi survey approach, each measurement scale was rated for its relevance to the THP program design elements and aims (1 = high relevance; 2 = moderate relevance; and 3 = low relevance) by three different raters: (a) the author of this thesis (and a coach on the Arctos Adventure Program), (b) the THP Coach Coordinator (and a coach on several of the Coaching Only Programs), and (c) the THP Coach Supervisor and THP program developer (and a coach on the James Craig and Outward Bound Adventure Programs). There was no training or discussion of the rating of scales among the raters prior to the initial rating. Consequently, inter-rater reliability was only moderate with ICC = .65 [95% CI: .34, .81].²⁷ Following this initial rating of the scales, the responses were shared among the group of raters. The raters then met via telephone conference to discuss the rating process and realised that the criteria to be used for rating the scales had not been clear. Following further discussion, the raters agreed that in addition to the connection between the underlying construct of interest and the THP program design elements and aims, it was important also to consider the individual scale items, in particular, the breadth and temporal orientation of those items. Where relevant, there would be a preference for more concise items querying a current way of being and therefore more sensitive to change (e.g., “When I have a problem, I can come up with lots of ways to solve it,” Pathways Thinking), as opposed to items tapping into a broader and longer-term assessment of oneself (e.g., “So far I have gotten the important things I want in life,” Life Satisfaction). The raters also agreed it was important to consider the semantics of the scale items, as they had become aware that some participants had indicated difficulty in understanding certain words in the Survey. Following this agreement on process, the raters discussed each scale until they came to consensus on a single rating for each scale. While it would

²⁷ The ICC estimate and its 95% confidence interval were calculated in R using the ICC function from the psych package (Revelle, 2018), based on a mean rating ($k = 3$), consistency, 2-way mixed-effects model.

have been preferable for the raters to re-rate the scales independently, certain timing and logistical factors made it impractical to do so. No specific consideration was given to the Coaching Only Program in this exercise. However, it is noted that while scales such as Physical Ability Self-Concept are less relevant for the Coaching Only Program, overall the majority of the scales are considered equally relevant to the Coaching Only Program as they are to the Adventure Programs.

Factor scores. In the psychometric analysis the subject of Chapter Five, factor analytic models were estimated for the Survey measures. From these models, factor scores were derived for each scale. Further detail on calculating the factor scores is provided in Chapter Four. These factor scores were used for all of the analyses in this study as proxies for the latent outcome variables measured by the Survey scales.

Multiple regression and multiple-group analyses. One advantage of SEM is that multiple relations between variables can be examined simultaneously. Moreover, SEM with latent variables (as opposed to manifest variables) takes into account measurement error, and this leads to a more precise estimation of the parameters of the structural model.

Multiple regression analysis. In the RCT analysis, structural parameters in relation to the latent outcome variables were estimated using ordinary linear regression analysis with the factor scores, often referred to as *factor score regression*. A separate multiple regression analysis was conducted for each group of scales sharing the same relevance rating. This was preferred to a single analysis incorporating all scales due to the number of parameters in the models and the small sample size for each THP program group. For the RCT analysis, a multiple regression analysis was conducted for each set of comparisons between repeated measures: pre-test to immediate post-test (T1-T2; Short-Term Analysis), pre-test to follow-up test (T1-T3; Long-Term Analysis), and immediate post-test to follow-up test (T2-T3; Follow-Up Analysis). Each of these models used factor scores for the relevant latent outcome variables as the dependent variables and a series of dummy variables representing each THP program group as the independent variables. The control group served as the baseline group. A number of covariates were also

included to control for a number of pre-treatment variables, and these are described in more detail below. Comparisons between the Adventure Programs (taken together) and the control group, as well as between the Adventure Programs (taken together) and the Coaching Only Program, were tested in Mplus using equations within the model constraint option.

Multiple-group analysis. For the pre-post within-subjects analysis, there was less data available with fewer waitlist control group participants ($n = 176$) and 43.18% of those participants with missing data at T4 (immediate program post-test for the control group). Accordingly, rather than group the scales together by scale relevance, each scale was analysed separately in order to reduce the number of parameters in the models. A single multiple-group SEM using the MLR estimator was conducted for each scale with factor scores for the control group's three extended baseline tests (T1, T2, and T3), as well as the immediate post-test (T4). For each model, the dependent latent variable was regressed on the covariates for each assessment wave, and a mean was assessed for the dependent variable for each THP program. For each dependent variable, the model constraint option in Mplus was used to test whether there was a significant difference between the relevant control participants' immediate post-test factor scores (T4) and their extended baseline scores for each THP program individually, as well as the Adventure Programs (taken together). Differences between the Adventure Programs (taken together) and the Coaching Only Program also were assessed within the model constraint option of Mplus.

Clustered data. Given the multiple cohort research design (with 54 cohorts), it is important to account for the potential within-group effects of this clustered data (e.g., non-independence; Sibthorp, Witter, Wells, Ellis, & Voelkl, 2004). Multilevel modelling (also referred to as hierarchical linear modelling) is an appropriate method for clustered data because it allows intercepts and slopes to vary by groups (Gelman & Hill, 2007). Initially, a multilevel model was tested for the short-term analysis with the high relevance scales. However, the number of parameters in the model relative to the number of clusters resulted in issues with model identification, potentially prejudicing the standard errors in the model.

Particularly in psychology-related research, the complex design option in Mplus provides an alternative approach for producing cluster-robust standard errors without the additional steps and assumptions inherent in multilevel models (McNeish, Stapleton, & Silverman, 2017). Accordingly, the multiple regression models used the Mplus complex design option with cohort as the cluster variable, in order to appropriately adjust standard errors to account for the clustered data structure. The complex design option applies a robust maximum likelihood estimator (MLR), which is robust against violations of normality assumptions. Using the complex design option in the within-subjects multiple-group analysis with a smaller sample and fewer cohorts resulted in issues with model identification, possibly prejudicing the standard errors in the model. Consequently, clustering was not taken into account in the within-subjects analysis.

Pre/Post intervention effects. The Short-Term and Long-Term Analyses controlled for any significant differences in the dependent variables at T1, so as to control for any pre-treatment differences. Such effects may occur, for example, from a sense of excitement or fear in connection with the proposed intervention (Hattie et al., 1997). The Follow-Up Analysis controlled for any significant differences in the dependent variables at both T1 and T2, so as to consider only the incremental change between T2 and T3. Doing so also serves to control for any post-test bias that may have occurred as a result of excitement or sense of achievement immediately following completion of the program (Marsh et al., 1986a; see also Neill, 2008).

Additional covariates. It is suggested that variables that theoretically reflect individual differences in the outcomes of interest should be included in the analysis in order to reduce some of the unexplained variation in the outcome variables, thereby reducing the standard errors of the intervention effect and increasing the precision of the effect estimates (de Boer, Waterlander, Kuijper, Steenhuis, & Twisk, 2015; Gelman & Hill, 2007; Senn, 2013). Consequently, a number of independent variables were included in the models as covariates in addition to pre-test scores for the outcome variables. In particular, variables were included to control for individual differences in gender, previous outdoor adventure

experience, socioeconomic status, and pre-existing levels of “flourishing” (see below for a description of how this variable was derived). Gender and previous outdoor adventure experience are both dichotomous independent variables (female = 0; male = 1; no = 0; yes = 1). Previous outdoor experience was measured through the program application, which asked an applicant whether they had previously been on an adventure education program or camp. Socioeconomic status is a variable consisting of summated yes/no self-rated responses (no=1; yes=2) for each participant as to whether or not the participant had any of the following in their home: own room, study space, computer, internet, dictionary, and dishwasher. Accordingly, scores for a participant on this variable ranged from 6 to 12 (with lower scores intended to represent lower socioeconomic status). The pre-existing flourishing variable consisted of a factor score for each participant derived for the first principal component from a principal component analysis conducted in SPSS on the T1 (pre-test) factor scores for all of the outcome variables used in this study. The principal component analysis used an oblique rotation and extracted three factors based on a scree plot of the data. The first principal component explained 43.1% of the total variance.

Data transformations. Several data transformations were undertaken to improve the interpretation of the results (Gelman & Hill, 2007). While the factor scores were already standardised, other covariate variables, such as socioeconomic status, were standardised across the entire sample. This was particularly important for the models that included interaction terms (Gelman & Hill, 2007).

Interaction effects. Interaction effects represent the multiplicative effect of predictor variables on the outcome variables. When an interaction effect is present, the impact of one variable depends on the level of the other variable (referred to as the *moderator* variable). Part of the power of multiple regression analysis is the ability to estimate and test interaction effects when the predictor variables are either categorical or continuous. If interaction effects are significant, then the interpretation of a single predictor variable in isolation may be misleading. Consequently, it has been suggested that multiple effects should be studied in intervention research (Pedhazur & Schmelkin, 1991). Because the THP program was

specifically designed for disadvantaged students, it was considered important to understand the extent to which a participant's initial baseline level (or aptitude) for an outcome variable (represented by the participant's pre-test (T1) factor score for the scale measuring the outcome variable) impacted the program's effectiveness in relation to that outcome variable. Consequently, additional multiple regression analyses were conducted for the outcome variables grouped by scale relevance for the Short-Term and Long-Term Analyses, incorporating additional predictor variables in the regression models defined as the product of the pre-test (T1) factor scores for the relevant outcome variables and program group (referred to as an *aptitude-treatment interaction*). A significant aptitude-treatment interaction effect for an outcome variable is indicated by the interactive effect of program participation and pre-test aptitude on the post-test outcome, controlling for the main effects of program participation and pre-test aptitude on the post-test outcome. Where significant aptitude-treatment interaction effects were found, simple main effects of the relevant programs were calculated for three levels of the pre-test (T1) factor scores for the relevant outcome variables. It is common to consider conditional values that are one standard deviation below the mean of the relevant variable, equal to the mean of that variable, and one standard deviation above the mean of that variable (Bauer & Curran, 2005; Preacher, Curran, & Bauer, 2006).

Effect sizes. As discussed in Chapter Four, an effect size (ES) generally describes the magnitude of a relationship or an effect between two or more variables, in a standardised way. The size of the difference between a THP program and the control group, or between repeated measures for the control group, is evidenced by the regression or other parameter coefficient. As recommended by Aiken and West (1991), all of the outcome variables in the models were standardised, so the unstandardised coefficients used in the evaluation of the effects, effectively represent standardised effect sizes (see also Marsh, 1992a; Montgomery, Peck, & Vining, 2012). In general, this statistic is said to represent a small effect around .10, a medium effect around .30, and a large effect over .50 (Cohen, 1988, 1992). The standard error also is reported as a measure of the

accuracy of the effect size estimate. Moreover, graphs are included that illustrate the 95% confidence interval for the effect size estimates. Finally, while the significance threshold for this research was nominally set at .05 (without any adjustment for the number of comparisons), statistical significance is indicated at the .05, .01, and .001 alpha levels. A more detailed evaluation of significance also can be derived from the effect size estimate and standard error.

Summary

This section described the methodology and procedures specific to the analyses carried out in this study. The methodology and procedures outlined in this section, together with the details provided in Chapter Four, demonstrate the rigorous approach that was taken to test the proposed research hypotheses and questions for this study. The following section presents the results of the analyses undertaken in connection with those research hypotheses and questions.

Results

This section outlines the results from the quantitative analyses of the effects of the THP program on a range of outcomes for its disadvantaged, adolescent participants. Based on the objectives of the THP program, it was anticipated that the program would positively influence the outcome measures over both the short term and long term, with an open question as to whether there would be any additional effects during the three-month period following completion of the program. Moreover, those outcome variables having the closest connection to the THP program design and aims, were hypothesised to demonstrate the greatest effects. It was also anticipated that participants who started the program with lower levels of an outcome variable would experience greater effects, particularly for the most relevant scales. There was an open question about the extent to which the Adventure Programs would demonstrate effects that exceeded the effects of the Coaching Only Program.

This section begins with a brief presentation of some preliminary descriptive statistics and analysis. Following that introduction is a high-level overview of the results, after which more detailed results are presented in the order of the research hypotheses and questions outlined above. Owing to the focus on scale relevance

and the grouping of scales by relevance for the related analyses, the scales are presented in these results by their category of relevance (being high, moderate, or low).

Preliminary Descriptive Statistics and Analysis

Descriptive statistics (means, standard deviation (*SD*), skewness, kurtosis, and standard error (*SE*) are presented in Appendix S for each scale and for each of the relevant assessment waves by reference to each of the following groups of research participants:

- Adventure Programs (taken together) (T1-T3);
- Arctos Adventure Program (T1-T3);
- James Craig Adventure Program (T1-T3);
- Outward Bound Adventure Program (T1-T3);
- Coaching Only Program (T1-T3);
- waitlist control group (T1-T3);
- control group in Adventure Programs (taken together) (T1-T4);
- control group in Arctos Adventure Program (T1-T4);
- control group in James Craig Adventure Program (T1-T4);
- control group in Outward Bound Adventure Program (T1-T4); and
- control group in Coaching Only Program (T1-T4).

T1 data was analysed to consider whether there were significant statistical differences between the control and various THP program groups (and between the Adventure Programs and Coaching Only Program groups) prior to entry into a THP program. T1 factor scores for the outcome variables (grouped by scale relevance) were regressed on a series of dummy variables indicating allocation to the various THP programs or the control group (with the control group as the reference group), as well as the other covariates described in the “Methodology and Procedures” section above. Table 6.4 sets out a high-level overview of this analysis, indicating whether the difference was significant (using † and – signs to indicate the direction and level of the significance) or not (indicated by *ns*). A detailed table containing the effect sizes, standard errors, and significance level for this analysis is presented in Appendix T. Of the 246 comparisons, only 15 demonstrated significance (6.10%,

approximately the alpha level used to test significance with $\alpha = .05$). For each group of scales, the scales with moderate relevance had a significant mean TI difference for the Outward Bound program and the scales with low relevance had a significant mean TI difference for the comparison of the Adventure Programs (taken together) with the Coaching Only Program, but each had only a small effect size ($ES = -.060$ and $ES = -.085$, respectively). Any differences will be controlled for in the further analyses.

Table 6.4
Differences at Baseline (T1) between Groups

Scale	Baseline (T1) Differences Between Groups					
	Adv vs. C	Arctos vs. C	JC vs. C	OB vs. C	CO vs. C	Adv vs. CO
<i>High Relevance</i>						
Agency	ns	ns	ns	ns	ns	ns
Pathways Thinking	†	ns	ns	†	ns	ns
Goal Self-Regulation	ns	ns	ns	ns	ns	ns
General Self-Esteem/SC	ns	ns	ns	ns	ns	ns
Self-Confidence	ns	ns	ns	ns	ns	ns
Self-Efficacy	ns	ns	ns	ns	ns	ns
Open Thinking	ns	ns	ns	ns	ns	ns
Cooperative Teamwork	ns	ns	ns	ns	ns	ns
Active Involvement	ns	ns	ns	ns	ns	ns
<i>Mean for High Relevance</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>
<i>Moderate Relevance</i>						
Optimism	ns	ns	ns	ns	--	†
Emotional Self-Regulation	ns	ns	ns	-	ns	ns
Booster Behaviors	ns	ns	ns	ns	ns	ns
Booster Thoughts	ns	ns	ns	ns	ns	ns
Wellbeing	ns	ns	ns	†	ns	ns
Same-Sex Relationships SC	ns	ns	ns	ns	ns	ns
Emotional Stability SC	ns	ns	ns	--	ns	ns
Stress Management	†	ns	†	ns	ns	ns
Social Effectiveness	ns	ns	ns	ns	ns	ns
Leadership Ability	ns	ns	ns	ns	ns	ns
Coping with Change	ns	ns	ns	ns	ns	ns
Internal LOC	ns	ns	ns	ns	ns	ns
<i>Mean for Moderate Relevance</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>---</i>	<i>ns</i>	<i>ns</i>
<i>Low Relevance</i>						
Pessimism ^a	ns	ns	--	ns	ns	--
Focus	ns	ns	ns	ns	ns	ns
Consistency of Interest	ns	ns	ns	ns	ns	ns
Perseverance of Effort	ns	ns	ns	ns	ns	ns
Hampering ^a	ns	ns	ns	ns	ns	ns
Life Resilience	†	ns	ns	ns	ns	ns
Academic Resilience	ns	ns	ns	ns	†	ns
Life Satisfaction	ns	ns	ns	ns	ns	ns
Physical Abilities SC	ns	ns	ns	ns	ns	ns
Physical Appearance SC	ns	ns	ns	ns	ns	ns
Opposite-Sex Relationships SC	ns	ns	ns	ns	ns	ns
Parent Relationships SC	ns	ns	ns	ns	ns	ns
Honesty-Trustworthiness SC	ns	ns	ns	ns	ns	ns
Math SC	ns	ns	ns	ns	ns	ns
Verbal SC	ns	ns	ns	ns	ns	-
School SC	ns	ns	ns	ns	ns	ns
Time Efficiency	ns	ns	ns	ns	ns	ns
Quality Seeking	ns	ns	ns	ns	ns	ns
Overall Effectiveness	ns	ns	ns	ns	ns	†
External LOC ^a	ns	ns	ns	ns	ns	ns
<i>Mean for Low Relevance</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>-</i>

Note. Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; C = waitlist control group; SC = Self-Concept; LOC = Locus of Control. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean difference for that group of scales. † indicates a significant positive difference (favoring THP program participants when compared to control participants or Adv when compared to CO) and - indicates a significant negative difference (favoring control participants when compared to the THP program participants or CO when compared to Adv): † or - ($p < .05$); †† or -- ($p < .01$); ††† or --- ($p < .001$); ns (no significant difference).

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome

Results of Research Hypotheses and Questions

Overview of Results. This study evaluated the short-term, long-term, and follow-up effects of the THP programs on 41 different outcome variables. Table 6.5 provides a visual overview of these results for each THP program individually, as well as for the Adventure Programs taken together. A separate comparison was also made between the Adventure Programs (taken together) and the Coaching Only Program. The Short-Term Analysis considered the factor scores for each scale immediately after program completion (T2) controlling for differences in the relevant scale prior to the program (T1), referred to in Table 6.5 as *T2*. The Long-Term Analysis considered the factor scores for each scale approximately three months after program completion (T3) controlling for differences in the relevant scale prior to the program (T1) but ignoring any differences immediately after program completion (T2), referred to in Table 6.5 as *T3a*. The Follow-Up Analysis considered the factor scores for each scale approximately three months after program completion (T3) controlling for both differences in the relevant scale prior to the program (T1) as well as immediately after the program (T2), focusing only on the incremental effects that occurred during the three-month period after program completion, referred to in Table 6.5 as *T3b*. Table 6.5 indicates only whether the difference between groups was significant (using † and – signs to indicate the direction and level of the significance) or not (indicated by *ns*). In this table and the specific outcome tables that follow, a significant result in the positive direction for an analysis indicates that at the relevant assessment wave (i.e., T2 for the Short-Term Analysis and T3 for the Long-Term and Follow-up Analyses), there were significant gains in scores in the particular outcome variable over the relevant earlier scores (i.e., T1 for the Short- and Long-Term Analyses and T2 for the Follow-Up Analyses) for participants in the relevant THP program(s) when compared to the control group (or the Adventure Programs when compared to the Coaching Only Program). A significant result in the negative direction indicates significant reductions in those scores for the relevant participants, and a non-significant result indicates that scores were not significantly different for participants in the groups the subject of the comparison.

Table 6.5
 Overview of Results for Short-Term, Long-Term, and Follow-Up Analyses of THP Program Effects

Scale	Summary THP Program Effects																	
	Adv vs. C			Arctos vs. C			JC vs. C			OB vs. C			CO vs. C			Adv vs. CO		
	T2	T3a	T3b	T2	T3a	T3b	T2	T3a	T3b	T2	T3a	T3b	T2	T3a	T3b	T2	T3a	T3b
<i>High Relevance</i>																		
Agency	††	ns	ns	ns	ns	ns	†	ns	ns	†	ns	ns	ns	ns	ns	†††	ns	ns
Pathways Thinking	††	ns	ns	ns	ns	ns	†††	ns	†	ns	ns							
Goal Self-Regulation	ns	ns	ns	ns	ns	ns	††	ns	ns	-								
General Self-Esteem/SC	ns	ns	ns	ns	ns	ns	†	ns	ns	ns								
Self-Confidence	†	ns	ns	ns	ns	ns	†††	ns	ns	ns								
Self-Efficacy	††	ns	ns	ns	ns	ns	†††	ns	ns	ns	†	ns	ns	ns	ns	ns	ns	ns
Open Thinking	†	ns	ns	ns	ns	ns	††	ns	†	ns	ns							
Cooperative Teamwork	ns	†	ns	ns	ns	ns	ns	†	ns	††	†	ns						
Active Involvement	†	ns	ns	ns	ns	ns	††	†	ns	††	ns	ns						
<i>Mean High Relevance</i>	<i>†</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>†††</i>	<i>ns</i>	<i>†</i>	<i>ns</i>	<i>ns</i>							
<i>Moderate Relevance</i>																		
Optimism	ns	ns	ns	ns	ns	ns	†††	†	ns	ns	ns	ns	ns	†	†	ns	-	---
Emotional Self-Regulation	ns	ns	ns	ns	ns	ns	†	ns	ns	ns								
Booster Behaviours	ns	†	ns	ns	ns	ns	†††	ns	ns	ns								
Booster Thoughts	ns	ns	ns	ns	ns	ns	†	†	ns	ns	ns	ns	ns	ns	†	ns	ns	ns
Wellbeing	††	ns	ns	ns	ns	ns	†††	†	ns	ns	ns							
Same-Sex Rel'ships SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Emotional Stability SC	ns	ns	ns	ns	ns	ns	ns	ns	-	ns	ns	ns						
Stress Management	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Social Effectiveness	†	ns	ns	ns	ns	-	†	†	ns	ns	ns	ns	†	ns	ns	ns	ns	ns
Leadership Ability	ns	ns	-	ns	ns	ns	†††	†	ns	ns	ns	-	ns	ns	ns	ns	ns	ns
Coping with Change	ns	ns	ns	ns	ns	ns	†	ns	ns	ns								
Internal LOC	ns	ns	ns	ns	ns	ns	†††	†	ns	†	ns	ns						
<i>Mean Moderate Relevance</i>	<i>†</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>†††</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>								

Note. Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; C = waitlist control group; T2 = group differences in the factor scores for the scale at T2, controlling for differences at T1 as well as other covariates in the model; T3a = group differences in the factor scores for the scale at T3, controlling for differences at T1 as well as other covariates in the model; T3b = group differences in the factor scores for the scale at T3, controlling for differences at T1 and T2 as well as other covariates in the model; SC = Self-Concept; Rel'ships = Relationships; LOC = Locus of Control. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean difference for that group of scales. † indicates a significant positive effect for THP program participants when compared to control group (or significant difference in favor of Adv when compared to CO) and - indicates a significant negative effect for THP program participants when compared to control group (or significant difference in favor of CO when compared to Adv): † or - (p < .05); †† or --- (p < .01); ††† or --- (p < .001); ns (no significant effect).

(continues)

Table 6.5 (continued)

Overview of Results for Short-Term, Long-Term, and Follow-Up Analyses of THP Program Effects

Scale	Summary THP Program Effects																	
	Adv vs. C			Arctos vs. C			JC vs. C			OB vs. C			CO vs. C			Adv vs. CO		
	T2	T3a	T3b	T2	T3a	T3b	T2	T3a	T3b	T2	T3a	T3b	T2	T3a	T3b	T2	T3a	T3b
<i>Low Relevance</i>																		
Pessimism ^a	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Focus	††	ns	ns	ns	ns	ns	†††	ns	ns	ns								
Consistency of Interest	ns	ns	ns	ns	ns	ns	ns	ns	-	ns	ns	ns	---	ns	ns	†††	ns	ns
Perseverance of Effort	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	---	ns	ns	†††	ns	ns
Hampering ^a	†	ns	ns	ns	ns	ns	†	ns	ns	†	ns	ns	ns	ns	ns	ns	ns	ns
Life Resilience	ns	ns	ns	ns	ns	ns	††	ns	ns	ns	ns	†	ns	ns	ns	ns	ns	ns
Academic Resilience	ns	ns	ns	ns	ns	ns	ns	ns	-	ns	ns	ns						
Life Satisfaction	ns	ns	ns	ns	ns	ns	ns	†	ns	ns	ns							
Physical Abilities SC	ns	ns	ns	††	ns	--	ns	ns	†									
Physical Appearance SC	ns	ns	ns	ns	ns	ns	ns	ns	†††	ns	ns	ns						
Opposite-Sex Rel'ships SC	ns	†	†	ns	ns	ns	ns	†††	†	ns	ns	ns	ns	†	†††	ns	ns	ns
Parent Rel'ships SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Honesty-Trust SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	†	†††	ns	ns	ns	ns	ns	ns
Math SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Verbal SC	†	ns	ns	ns	ns	ns	†††	ns	ns	ns	ns	ns	†	†	ns	ns	ns	ns
School SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Time Efficiency	ns	ns	ns	ns	ns	ns	†††	ns	ns	ns								
Quality Seeking	†	ns	ns	ns	ns	ns	†††	ns	†	ns	ns							
Overall Effectiveness	ns	ns	ns	ns	ns	ns	†††	ns	ns	ns								
External LOC ^a	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	---	-	ns	††	†	ns
<i>Mean Low Relevance</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>††</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>								

Note. Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; C = waitlist control group; T2 = group differences in the factor scores for the scale at T2, controlling for differences at T1 as well as other covariates in the model; T3a = group differences in the factor scores for the scale at T3, controlling for differences at T1 as well as other covariates in the model; T3b = group differences in the factor scores for the scale at T3, controlling for differences at T1 and T2 as well as other covariates in the model; SC = Self-Concept; Rel'ships = Relationships; LOC = Locus of Control. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean difference for that group of scales. † indicates a significant positive effect for THP program participants when compared to control group (or significant difference in favor of Adv when compared to CO) and - indicates a significant negative effect for THP program participants when compared to control group (or significant difference in favor of CO when compared to Adv); † or - ($p < .05$); †† or -- ($p < .01$); ††† or --- ($p < .001$); ns (no significant effect).

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

Results of Research Hypothesis 2.1: Short-term effects of THP programs on outcome variables. Research Hypothesis 2.1 predicted that immediately following completion of the program (T2), participants in each of the THP programs would report statistically significant higher factor scores in each of the scales measuring the outcome variables when compared to those factor scores for the control group, controlling for pre-test differences. Furthermore, it was predicted that those scales of greatest relevance to the THP program design and goals would evidence the greatest effects.

The results of the Short-Term Analysis for each outcome variable, grouped by scale relevance, are presented in Table 6.6. A mean effect size for the scales grouped by relevance, is also presented. Figure 6.2 illustrates the point estimate of the short-term effect size and 95% confidence interval for each outcome variable (grouped by scale relevance) across each program or group of programs.

Table 6.6
Short-Term Program Effects (T1-T2)

Scale	Program Effects: ES (SE)				
	Adv vs. C	Arctos vs. C	JC vs. C	OB vs. C	CO vs. C
<i>High Relevance</i>					
Agency	.266 (.087)**	.178 (.139)	.275 (.124)*	.344 (.164)*	-.120 (.095)
Pathways Thinking	.247 (.093)**	.119 (.142)	.490 (.128)***	.131 (.169)	-.082 (.156)
Goal Self-Regulation	.152 (.094)	.074 (.124)	.370 (.142)**	.011 (.083)	-.126 (.181)
General Self-Esteem/SC	.084 (.112)	-.081 (.163)	.309 (.145)*	.023 (.183)	-.065 (.176)
Self-Confidence	.215 (.100)*	.049 (.160)	.541 (.114)***	.053 (.142)	.069 (.133)
Self-Efficacy	.226 (.088)**	.134 (.139)	.406 (.120)***	.139 (.085)	.104 (.242)
Open Thinking	.213 (.108)*	-.010 (.144)	.430 (.149)**	.219 (.149)	-.056 (.136)
Cooperative Teamwork	.192 (.104)	.193 (.119)	.276 (.160)	.107 (.144)	-.038 (.082)
Active Involvement	.218 (.101)*	.029 (.148)	.412 (.142)**	.211 (.138)	-.027 (.097)
<i>Mean for High Relevance</i>	.210 (.081)*	.076 (.119)	.390 (.116)***	.138 (.113)	-.038 (.102)
<i>Moderate Relevance</i>					
Optimism	.097 (.079)	-.076 (.093)	.430 (.114)***	-.064 (.151)	.078 (.119)
Emotion Self-Regulation	.022 (.088)	-.082 (.120)	.353 (.154)*	-.204 (.121)	-.283 (.210)
Booster Behaviours	.198 (.105)	.051 (.124)	.508 (.141)***	.034 (.184)	.010 (.306)
Booster Thoughts	.116 (.095)	-.026 (.147)	.326 (.133)*	.049 (.124)	.086 (.178)
Wellbeing	.337 (.112)**	.148 (.172)	.630 (.150)***	.234 (.205)	.225 (.179)
Same-Sex Relationships SC	.004 (.100)	-.080 (.139)	.026 (.138)	.065 (.169)	.091 (.123)
Emotional Stability SC	-.038 (.079)	.059 (.126)	-.067 (.138)	-.106 (.152)	.037 (.198)
Stress Management	.124 (.099)	.010 (.151)	.274 (.149)	.089 (.150)	-.023 (.255)
Social Effectiveness	.184 (.090)*	-.027 (.147)	.336 (.160)*	.244 (.128)	.234 (.080)*
Leadership Ability	.134 (.077)	-.114 (.113)	.380 (.084)***	.137 (.160)	.191 (.109)
Coping with Change	.135 (.094)	-.036 (.142)	.278 (.142)*	.162 (.139)	.010 (.204)
Internal Locus of Control	.213 (.118)	-.053 (.178)	.479 (.120)***	.214 (.179)	-.053 (.130)
<i>Mean for Moderate Relevance</i>	.127 (.061)*	-.019 (.093)	.329 (.097)***	.071 (.099)	.050 (.119)
<i>Low Relevance</i>					
Pessimism ^a	.043 (.082)	-.028 (.119)	.082 (.117)	.075 (.119)	-.170 (.212)
Focus	.223 (.085)**	.071 (.108)	.455 (.137)***	.142 (.113)	-.223 (.325)
Consistency of Interest	.001 (.088)	.007 (.133)	.099 (.173)	-.103 (.121)	-.523 (.116)***
Perseverance of Effort	.035 (.091)	.075 (.075)	.319 (.164)	-.287 (.148)	-.386 (.109)***
Hampering ^a	.173 (.078)*	.004 (.104)	.265 (.130)*	.249 (.104)*	-.247 (.236)
Life Resilience	.057 (.084)	-.005 (.122)	.326 (.106)**	-.148 (.133)	-.182 (.192)
Academic Resilience	.050 (.091)	-.095 (.106)	.222 (.156)	.024 (.151)	-.360 (.271)
Life Satisfaction	.041 (.099)	.017 (.143)	.223 (.154)	-.118 (.158)	.231 (.191)
Physical Abilities SC	.151 (.080)	.253 (.083)**	.150 (.155)	.049 (.108)	.023 (.105)
Physical Appearance SC	.038 (.098)	.042 (.123)	.123 (.179)	-.051 (.111)	.101 (.176)
Opposite-Sex Relationships SC	.059 (.093)	.176 (.115)	.097 (.115)	-.095 (.149)	.097 (.159)
Parent Relationships SC	.004 (.081)	-.010 (.116)	.164 (.106)	-.142 (.122)	.070 (.100)
Honesty-Trustworthiness SC	.031 (.094)	.069 (.106)	-.026 (.174)	.049 (.150)	-.224 (.216)
Math SC	.082 (.082)	.098 (.117)	-.015 (.139)	.164 (.137)	-.100 (.260)
Verbal SC	.169 (.082)*	.181 (.103)	.330 (.092)***	-.004 (.111)	.278 (.082)***
School SC	.169 (.092)	.119 (.115)	.246 (.139)	.142 (.159)	.033 (.198)
Time Efficiency	.163 (.087)	.029 (.117)	.467 (.141)***	-.006 (.146)	-.142 (.276)
Quality Seeking	.241 (.099)*	.070 (.140)	.549 (.105)***	.103 (.145)	.002 (.106)
Overall Effectiveness	.149 (.095)	-.004 (.130)	.449 (.119)***	.001 (.148)	.137 (.173)
External Locus of Control ^a	-.087 (.073)	-.122 (.091)	-.091 (.109)	-.048 (.118)	-.338 (.095)***
<i>Mean for Low Relevance</i>	.090 (.054)	.047 (.064)	.222 (.093)**	.000 (.087)	-.096 (.141)

Note. T₁ = pre-test; T₂ = immediate post-test; ES = standardised effect size; SE = standard error; Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. SC = Self-Concept. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean effect for that group of scales. Significant effects are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

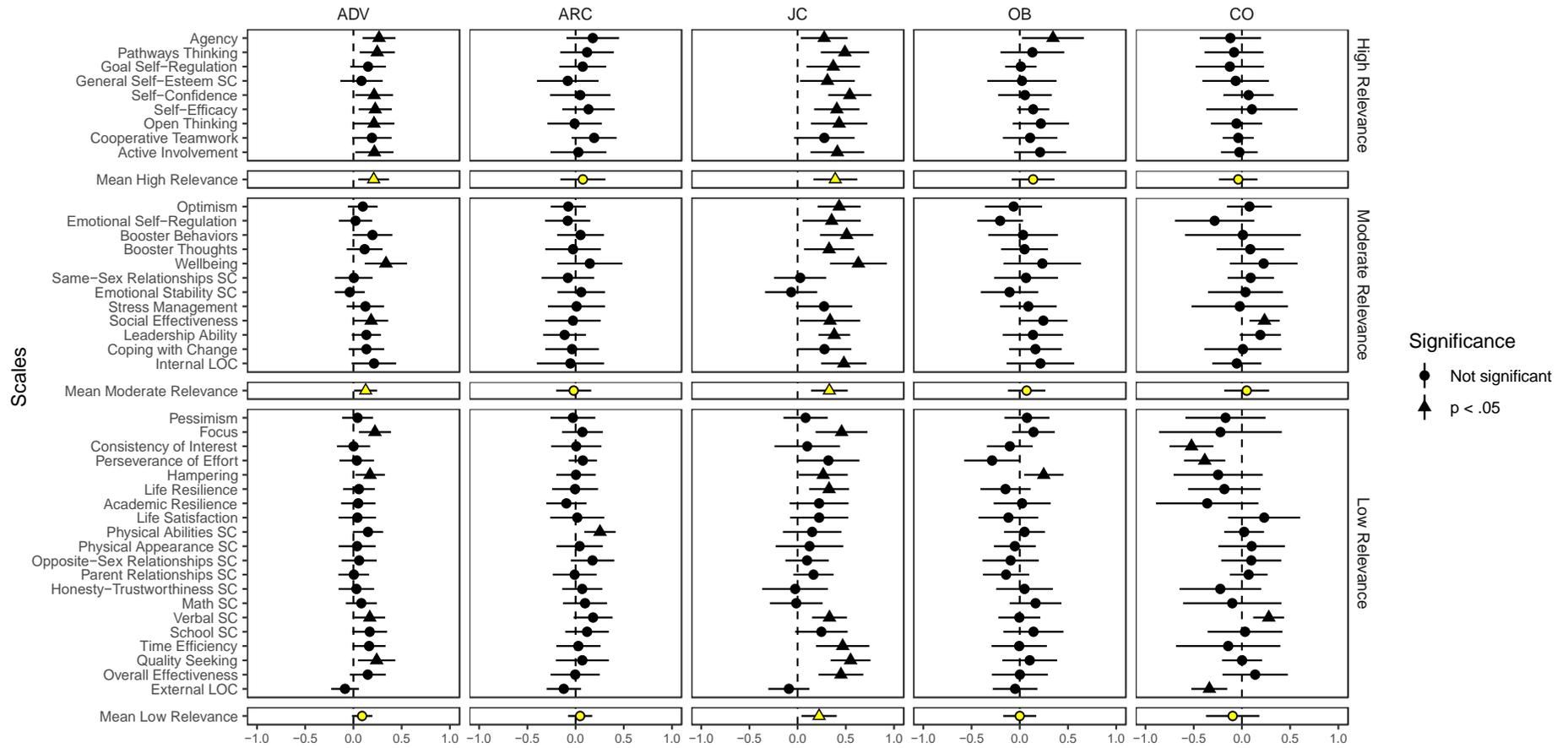


Figure 6.2. THP program effect sizes with 95% confidence intervals for Short-Term Analysis (T1-T2) with scales grouped by relevance.

Note. T1 = pre-test; T2 = immediate post-test; ADV = Adventure Programs (taken together); ARC = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; SC = Self-Concept; LOC = Locus of Control. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean effect for that group of scales. Effect sizes are standardised.

Results of Research Hypothesis 2.1.1: Positive short-term effects of THP programs on all outcome variables. For the Adventure Programs (taken together), 12 scales evidenced significant positive short-term effects when compared to the control group. The effect sizes for significant effects were small to moderate, ranging from $ES = .169$ (Verbal Self-Concept, a low relevance scale) to $ES = .337$ (Wellbeing, a moderate relevance scale), with an average effect size for all significant short-term effects for the Adventure Programs equal to $.226$. However, the individual Adventure Programs demonstrated disparate results, with the James Craig Adventure Program²⁸ evidencing many more significant positive short-term effects than any of the other programs. Participants in the James Craig program reported significantly higher scores at T2 for 24 of the scales when compared to the control group, with mostly moderate-to-large effect sizes for these significant effects, ranging from $ES = .265$ (Hampering, a low relevance scale) to $.630$ (Wellbeing, a moderate relevance scale), with an average effect size for all significant positive effects for the James Craig program equal to $.408$. The Outward Bound Adventure Program participants reported significantly higher scores than the control group at T2 for two scales (Agency: $ES = .344$ and Hampering: $ES = .249$). The Arctos Adventure Program participants demonstrated only a single scale with a significant effect for Physical Abilities Self-Concept ($ES = .253$). The Coaching Only Program participants demonstrated significantly higher scores at T2 for two of the scales when compared to the control group (Social Effectiveness: $ES = .234$ and Verbal Self-Concept: $ES = .278$), however, for three of the scales they evidenced significantly lower scores (Perseverance of Effort: $ES = -.523$, Consistency of Interest: $ES = -.386$, and External Locus of Control: $ES = -.338$).

While not all THP programs evidenced significant effects for all outcome variables, these results were partially consistent with the a priori prediction in Research Hypothesis 2.1.1. Participants across the Adventure Programs reported significant outcomes in many important qualities and skills related to flourishing,

²⁸ For ease of reference, individual Adventure Programs are sometimes referred to by their mode followed by “program” or “participants” (or both) rather than “Adventure Program” (e.g., James Craig program).

including hope, self-confidence, self-efficacy, wellbeing, and aspects of self-regulation, motivation, and self-concept.

Results of Research Hypothesis 2.1.2: Greater positive short-term effects of THP programs on outcome variables with most relevance to THP programs. Participants in the Adventure Programs (taken together) evidenced significant positive short-term effects for six of the nine high relevance scales (66.67%), two of the 12 moderate relevance scales (16.67%), and four of the 20 low relevance scales (20%). Table 6.6 also includes a mean effect size for each group of scales categorised by relevance (high, moderate, and low). The Adventure Program participants demonstrated overall significant positive mean short-term effects for the high and moderate relevance scales, with a greater mean effect size for the high relevance scales (ES = .210) compared to the moderate relevance scales (ES = .127), lending support to Research Hypothesis 2.1.2.

Of the individual programs, participants in the James Craig Adventure Program demonstrated significant positive short-term effects for eight of the nine high relevance scales (88.89%), nine of the 12 moderate relevance scales (75%), and seven of the 20 low relevance scales (35%). Significant positive mean short-term effects were found for all three groups of scales, with the highest mean effect size for the high relevance scales (ES = .390), followed by the moderate relevance scales (ES = .329), and then the low relevance scales (ES = .222), again supporting Research Hypothesis 2.1.2. None of the other programs individually demonstrated significant mean short-term effects for any group of scales categorised by relevance.

Correlating the short-term effect sizes with scale relevance demonstrated a moderate correlation for the Adventure Programs (taken together; $r = .48$), as well as the James Craig ($r = .39$) and Outward Bound ($r = .41$) programs. The Arctos ($r = .04$) and Coaching Only ($r = .19$) programs reflected only a small correlation between effect sizes and scale relevance, and this result is likely because these programs had very few significant effects.

Overall, where participants in the THP programs reported significant positive effects, these effects were greater for those scales determined to be more

closely connected to the THP program design and aims, consistent with Research Hypothesis 2.1.2.

Results of Research Hypothesis 2.2: Long-term maintenance of positive effects of THP programs on outcome variables. Research Hypothesis 2.2 predicted that approximately three months following completion of the program (T3), participants in each of the THP programs would also report statistically significantly higher factor scores for the scales representing the outcome variables when compared to those factor scores for the control group, controlling for pre-test differences. In other words, Research Hypothesis 2.2 predicted that the positive effects anticipated immediately after program completion would be maintained approximately three months later.

The results of the Long-Term Analysis, modelled and grouped by scale relevance, are set out in Table 6.7. A mean effect size for the scales grouped by relevance, is also presented. Figure 6.3 illustrates the point estimate of the long-term effect size and 95% confidence interval for each outcome variable (grouped by scale relevance) across each program or group of programs. The specific follow-up effects between T2 and T3 will be considered further in the results for Research Question 2.3, below.

Table 6.7
Long-Term Program Effects (T1-T3)

Scale	Program Effects: ES (SE)				
	Adv vs. C	Arctos vs. C	JC vs. C	OB vs. C	CO vs. C
<i>High Relevance</i>					
Agency	.130 (.094)	.106 (.145)	.193 (.177)	.089 (.134)	.008 (.110)
Pathways Thinking	.147 (.089)	.066 (.111)	.281 (.189)	.094 (.125)	-.031 (.088)
Goal Self-Regulation	.106 (.108)	.060 (.126)	.238 (.223)	.021 (.130)	.129 (.148)
General Self-Esteem/SC	.099 (.109)	.118 (.145)	.137 (.200)	.042 (.141)	-.029 (.129)
Self-Confidence	.107 (.088)	.033 (.110)	.274 (.141)	.014 (.144)	.002 (.130)
Self-Efficacy	.150 (.096)	-.048 (.139)	.300 (.196)	.197 (.086)*	.062 (.198)
Open Thinking	.190 (.098)	-.011 (.127)	.366 (.188)	.215 (.126)	.008 (.176)
Cooperative Teamwork	.187 (.089)*	.109 (.110)	.352 (.162)*	.100 (.125)	.001 (.079)
Active Involvement	.155 (.087)	-.037 (.107)	.353 (.149)*	.149 (.132)	-.045 (.140)
<i>Mean for High Relevance</i>	<i>.141 (.079)</i>	<i>.044 (.103)</i>	<i>.277 (.163)</i>	<i>.102 (.097)</i>	<i>.012 (.109)</i>
<i>Moderate Relevance</i>					
Optimism	.011 (.070)	-.092 (.094)	.254 (.118)*	-.127 (.087)	.191 (.082)*
Emotion Self-Regulation	.022 (.106)	-.143 (.137)	.190 (.229)	.018 (.116)	.008 (.135)
Booster Behaviours	.222 (.101)*	.100 (.111)	.397 (.224)	.169 (.116)	.093 (.196)
Booster Thoughts	.196 (.106)	.044 (.136)	.419 (.204)*	.124 (.133)	.104 (.128)
Wellbeing	.184 (.095)	.044 (.125)	.469 (.185)*	.039 (.145)	.088 (.137)
Same-Sex Relationships SC	.052 (.086)	-.082 (.139)	.228 (.161)	.011 (.080)	.063 (.127)
Emotional Stability SC	-.088 (.097)	.002 (.133)	-.148 (.145)	-.119 (.172)	.006 (.107)
Stress Management	.131 (.104)	-.108 (.138)	.330 (.229)	.171 (.123)	.042 (.145)
Social Effectiveness	.049 (.089)	-.205 (.108)	.296 (.120)*	.055 (.132)	.063 (.099)
Leadership Ability	.027 (.088)	-.134 (.114)	.266 (.129)*	-.053 (.159)	.062 (.079)
Coping with Change	.100 (.108)	-.083 (.146)	.265 (.242)	.119 (.114)	-.005 (.147)
Internal Locus of Control	.101 (.091)	.002 (.149)	.309 (.121)*	-.007 (.134)	-.097 (.139)
<i>Mean for Moderate Relevance</i>	<i>.084 (.069)</i>	<i>-.055 (.083)</i>	<i>.273 (.150)</i>	<i>.033 (.085)</i>	<i>.052 (.097)</i>
<i>Low Relevance</i>					
Pessimism ^a	-.012 (.084)	.051 (.097)	-.049 (.157)	-.039 (.121)	-.026 (.079)
Focus	.130 (.093)	.171 (.119)	.159 (.174)	.060 (.150)	-.021 (.178)
Consistency of Interest	-.066 (.083)	.061 (.139)	-.115 (.139)	-.145 (.096)	-.146 (.179)
Perseverance of Effort	.046 (.095)	.050 (.104)	.050 (.172)	.038 (.160)	-.014 (.077)
Hampering ^a	.062 (.081)	.097 (.177)	.023 (.175)	.065 (.095)	.005 (.078)
Life Resilience	.088 (.100)	-.056 (.107)	.144 (.203)	.177 (.115)	-.148 (.165)
Academic Resilience	.038 (.097)	.066 (.107)	-.101 (.202)	.150 (.128)	-.121 (.170)
Life Satisfaction	.082 (.092)	-.025 (.128)	.273 (.133)*	-.004 (.154)	-.048 (.135)
Physical Abilities SC	.013 (.057)	.106 (.081)	.064 (.092)	-.130 (.078)	-.131 (.110)
Physical Appearance SC	.030 (.083)	-.042 (.100)	.217 (.121)	-.084 (.149)	.085 (.128)
Opposite-Sex Relationships SC	.228 (.081)**	.219 (.125)	.331 (.103)***	.133 (.113)	.318 (.085)***
Parent Relationships SC	.053 (.081)	.006 (.107)	.195 (.103)	-.043 (.113)	.044 (.106)
Honesty-Trustworthiness SC	.135 (.075)	.084 (.095)	.063 (.110)	.258 (.124)*	.092 (.191)
Math SC	.036 (.090)	.147 (.107)	-.084 (.186)	.046 (.163)	.073 (.077)
Verbal SC	.109 (.100)	.194 (.130)	.165 (.142)	-.032 (.134)	.296 (.118)*
School SC	.123 (.105)	.238 (.141)	.075 (.186)	.056 (.159)	.092 (.089)
Time Efficiency	.076 (.087)	.072 (.091)	.166 (.157)	-.011 (.115)	.061 (.180)
Quality Seeking	.101 (.086)	.097 (.106)	.251 (.145)	-.044 (.127)	.062 (.067)
Overall Effectiveness	.081 (.086)	.018 (.096)	.231 (.145)	-.005 (.124)	.011 (.105)
External Locus of Control ^a	.001 (.078)	-.006 (.102)	.119 (.112)	-.110 (.110)	-.308 (.134)*
<i>Mean for Low Relevance</i>	<i>.068 (.051)</i>	<i>.077 (.054)</i>	<i>.109 (.104)</i>	<i>.017 (.079)</i>	<i>.009 (.067)</i>

Note. T1 = pre-test; T3 = three months post-test; ES = standardised effect size; SE = standard error. Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. SC = Self-Concept. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean effect for that group of scales. Significant effects are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

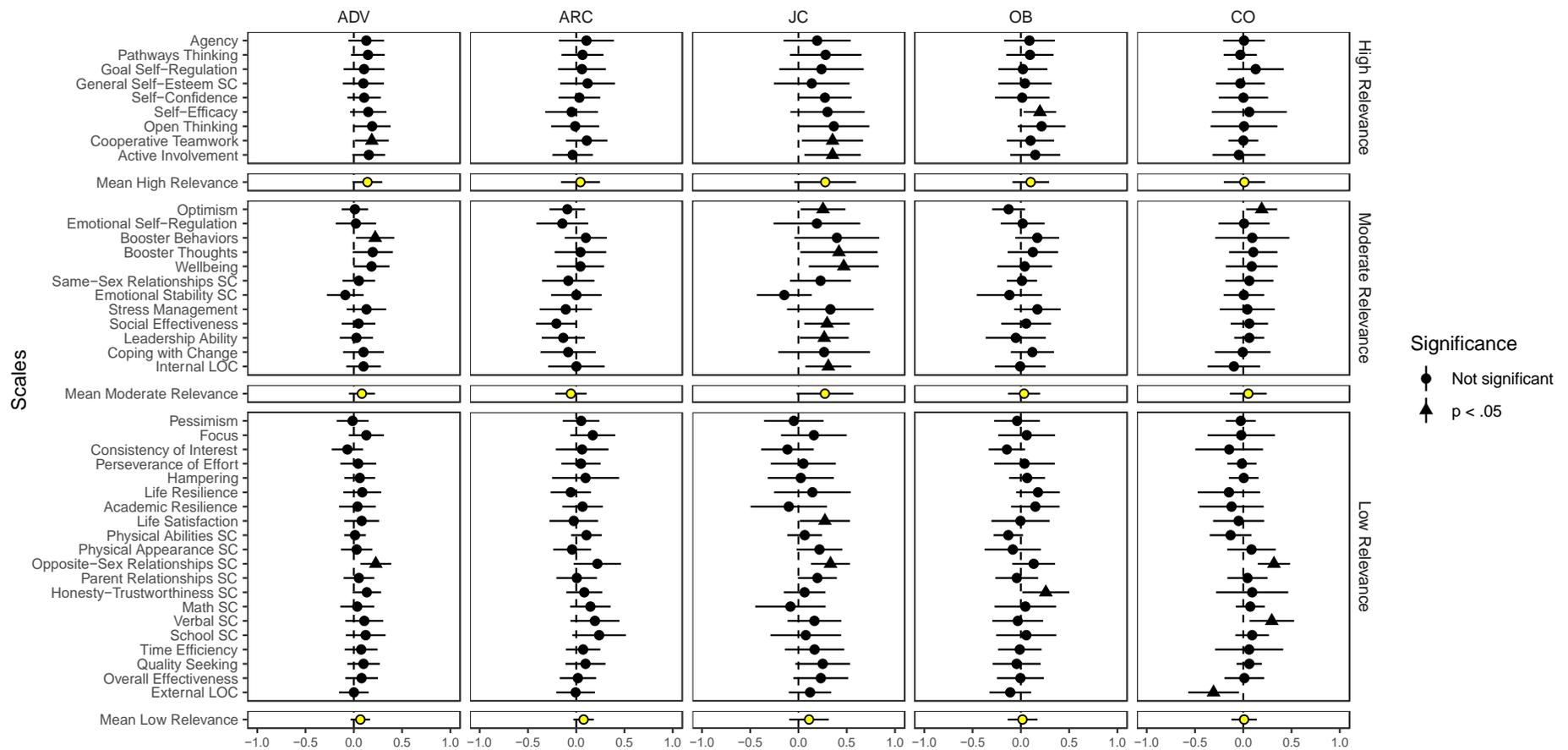


Figure 6.3. THP program effect sizes with 95% confidence intervals for Long-Term Analysis (T1-T3) with scales grouped by relevance.

Note. T1 = pre-test; T3 = three months post-test; ADV = Adventure Programs (taken together); ARC = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; SC = Self-Concept; LOC = Locus of Control. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean effect for that group of scales. Effect sizes are standardised.

For the Adventure Programs (taken together), three scales evidenced significant small positive long-term effects when compared to the control group: one scale of high relevance (Cooperative Teamwork: $ES = .187$), one scale of moderate relevance (Booster Behaviours: $ES = .222$), and one scale of low relevance (Opposite Sex Relationships Self-Concept: $ES = .228$). None of the outcome variables with significant effects reported by participants in the Adventure Programs in the short term maintained those significant positive effects over the longer term. There were also no significant mean effects for any of the high, moderate, or low relevance scales as a group of scales.

Results for the long-term effects of the individual THP programs were consistent with the short-term results, with the James Craig program participants demonstrating the most significant long-term effects. The James Craig participants reported significantly higher factor scores at T3 for 10 of the scales when compared to the control group, with significant effect sizes ranging from $ES = .254$ (Optimism, a moderate relevance scale) to $ES = .469$ (Wellbeing, a moderate relevance scale). Only one of the high relevance scales demonstrating a significant short-term effect maintained a significant effect longer term (Active Involvement: $ES = .353$), however, Cooperative Teamwork evidenced a new significant positive effect in the Long-Term Analysis ($ES = .352$). Other than the scales for Emotion Self-Regulation, Booster Behaviours, and Coping with Change, all of the moderate relevance scales demonstrating a short-term significant effect, maintained a significant effect long term. None of the seven low relevance scales with short-term significant effects maintained those significant effects long term. However, two new scales, Opposite Sex Relationships Self-Concept and Life Satisfaction, evidenced significant positive long-term effects. There were no significant mean effects for the James Craig program for any of the high, moderate, or low relevance scales as a group of scales.

The Outward Bound participants demonstrated a significant positive long-term effect for two scales (both of which were different to the significant short-term effects found for that program): Self-Confidence ($ES = .197$) and Honesty-Trustworthiness Self-Concept ($ES = .258$), one high and one low relevance scale. Accordingly, the short-term significant effects demonstrated by Outward Bound participants were not maintained longer-term. The Arctos participants did not

report any significant effects long-term, and therefore, the significant effect for Physical Abilities Self-Concept was not maintained longer-term. The Coaching Only participants reported significant positive long-term effects for Optimism (ES = .191), Opposite-Sex Relationships Self-Concept (ES = .318), and Verbal Self-Concept (ES = .296). There was also a significant negative long-term effect for External Locus of Control (ES = -.308). While External Locus of Control and Verbal Self-Concept were also reported as short-term significant effects, the other significant positive effect for Social Effectiveness and significant negative effects for Consistency of Interest and Perseverance of Effort were not maintained longer-term. There were no significant mean effects for any group of scales grouped by relevance for any of these THP programs.

Overall, eight of the 40 significant positive short-term effects were maintained longer term (20%) and only one of the three significant negative short-term effects was maintained, providing partial support for Research Hypothesis 2.2. There were 10 scales for which the short-term effect was not significant, but for which a long-term significant effect was demonstrated (with all of these effects being positive).

Results of Research Question 2.3: Follow-up effects of THP programs on outcome variables. Research Hypotheses 2.1 and 2.2 predicted the THP program effects would be significant and positive in both the short and long term. However, given the limited longitudinal research in the area of outdoor adventure education, the more specific effects of the THP programs on the outcome variables between program completion and the follow-up approximately three months later were left as an open research question. Therefore, Research Question 2.3 asked about the nature of the THP program effects during this follow-up period, with particular consideration to be given to whether those effects increased, dissipated, or were maintained over time.

The results of the Follow-up Analysis, modelled and grouped by scale relevance, are set out in Table 6.8. This analysis controls for any score differences at both T1 and T2 and, therefore, focuses only on the change in scores between T2 and T3. A mean effect size for the scales grouped by relevance, is also presented. Figure

6.4 illustrates the point estimate of the follow-up effect size and 95% confidence interval for each outcome variable (grouped by scale relevance) across each program or group of programs.

Table 6.8
Follow-Up Program Effects (T2-T3)

Scale	Program Effects: ES (SE)				
	Adv vs. C	Arctos vs. C	JC vs. C	OB vs. C	CO vs. C
<i>High Relevance</i>					
Agency	.004 (.073)	.052 (.120)	-.022 (.138)	-.017 (.065)	.093 (.112)
Pathways Thinking	.023 (.072)	.006 (.108)	.016 (.151)	.046 (.093)	.019 (.057)
Goal Self-Regulation	-.005 (.083)	-.026 (.107)	-.021 (.165)	.033 (.120)	.177 (.073)
General Self-Esteem/SC	.014 (.077)	.094 (.100)	-.078 (.120)	.025 (.110)	.060 (.064)
Self-Confidence	-.003 (.074)	-.002 (.092)	-.010 (.102)	.004 (.126)	.039 (.073)
Self-Efficacy	.019 (.074)	-.143 (.110)	.064 (.127)	.135 (.087)	.061 (.122)
Open Thinking	.063 (.073)	-.032 (.106)	.081 (.118)	.141 (.118)	.063 (.126)
Cooperative Teamwork	.048 (.068)	-.015 (.096)	.147 (.100)	.012 (.115)	.064 (.051)
Active Involvement	.038 (.073)	-.068 (.092)	.114 (.103)	.069 (.124)	.002 (.101)
<i>Mean for High Relevance</i>	<i>.022 (.055)</i>	<i>-.015 (.087)</i>	<i>.032 (.097)</i>	<i>.050 (.084)</i>	<i>.064 (.055)</i>
<i>Moderate Relevance</i>					
Optimism	-.075 (.057)	-.133 (.073)	.031 (.090)	-.122 (.078)	.190 (.080)*
Emotion Self-Regulation	-.042 (.086)	-.171 (.141)	-.046 (.148)	.091 (.100)	.105 (.074)
Booster Behaviours	.092 (.073)	.023 (.084)	.094 (.138)	.160 (.110)	.157 (.083)
Booster Thoughts	.105 (.078)	.028 (.107)	.199 (.166)	.088 (.103)	.152 (.072)*
Wellbeing	.003 (.071)	-.029 (.092)	.073 (.116)	-.034 (.113)	.069 (.067)
Same-Sex Relationships SC	.039 (.084)	-.046 (.118)	.187 (.146)	-.024 (.107)	.046 (.111)
Emotional Stability SC	-.122 (.071)	-.070 (.097)	-.186 (.092)*	-.111 (.113)	-.046 (.120)
Stress Management	.047 (.079)	-.164 (.105)	.138 (.148)	.168 (.098)	.117 (.090)
Social Effectiveness	-.103 (.074)	-.259 (.085)**	.058 (.094)	-.109 (.113)	-.089 (.076)
Leadership Ability	-.125 (.054)*	-.120 (.074)	-.049 (.070)	-.204 (.102)*	-.060 (.074)
Coping with Change	.004 (.080)	-.138 (.119)	.063 (.136)	.087 (.109)	.021 (.076)
Internal Locus of Control	-.022 (.070)	-.014 (.112)	.030 (.102)	-.081 (.098)	-.015 (.082)
<i>Mean for Moderate Relevance</i>	<i>-.016 (.045)</i>	<i>-.091 (.062)</i>	<i>.049 (.084)</i>	<i>-.008 (.066)</i>	<i>.054 (.037)</i>
<i>Low Relevance</i>					
Pessimism ^a	.033 (.063)	.102 (.087)	-.045 (.117)	.040 (.080)	.045 (.130)
Focus	.037 (.061)	.072 (.104)	-.009 (.090)	.050 (.091)	.013 (.123)
Consistency of Interest	-.109 (.070)	.039 (.119)	-.229 (.097)*	-.136 (.087)	.033 (.183)
Perseverance of Effort	-.039 (.070)	-.038 (.092)	-.133 (.117)	.054 (.119)	.042 (.095)
Hampering ^a	-.037 (.062)	.031 (.098)	-.127 (.098)	-.016 (.078)	.106 (.124)
Life Resilience	.042 (.078)	-.097 (.111)	-.035 (.109)	.256 (.113)*	-.080 (.096)
Academic Resilience	-.015 (.074)	.032 (.097)	-.222 (.109)*	.145 (.106)	-.028 (.087)
Life Satisfaction	.081 (.075)	.010 (.092)	.147 (.092)	.085 (.121)	-.053 (.096)
Physical Abilities SC	-.029 (.046)	-.028 (.073)	.010 (.071)	-.068 (.075)	-.226 (.076)**
Physical Appearance SC	.079 (.065)	-.007 (.087)	.285 (.066)***	-.040 (.112)	.063 (.115)
Opposite-Sex Relationships SC	.153 (.062)*	.129 (.092)	.221 (.093)*	.110 (.085)	.240 (.068)***
Parent Relationships SC	.079 (.072)	.022 (.101)	.147 (.085)	.068 (.097)	-.037 (.096)
Honesty-Trustworthiness SC	.113 (.059)	-.004 (.080)	.081 (.096)	.263 (.069)***	.086 (.108)
Math SC	.000 (.065)	.080 (.079)	-.030 (.100)	-.050 (.094)	.090 (.136)
Verbal SC	.059 (.064)	.107 (.092)	.009 (.101)	.062 (.080)	.285 (.152)
School SC	.062 (.072)	.131 (.103)	-.009 (.106)	.063 (.085)	.076 (.110)
Time Efficiency	-.041 (.069)	-.041 (.097)	-.079 (.093)	-.004 (.075)	.006 (.124)
Quality Seeking	-.016 (.074)	-.008 (.091)	-.008 (.120)	-.032 (.097)	.048 (.107)
Overall Effectiveness	-.006 (.075)	-.055 (.092)	-.021 (.122)	.060 (.098)	-.057 (.122)
External Locus of Control ^a	.033 (.070)	.057 (.093)	.123 (.098)	-.082 (.098)	-.202 (.140)
<i>Mean for Low Relevance</i>	<i>.024 (.032)</i>	<i>.027 (.051)</i>	<i>.004 (.047)</i>	<i>.041 (.043)</i>	<i>.022 (.070)</i>

Note. T₂ = immediate post-test; T₃ = three months post-test; ES = standardised effect size; SE = standard error; Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. SC = Self-Concept. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean effect for that group of scales. Significant effects are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

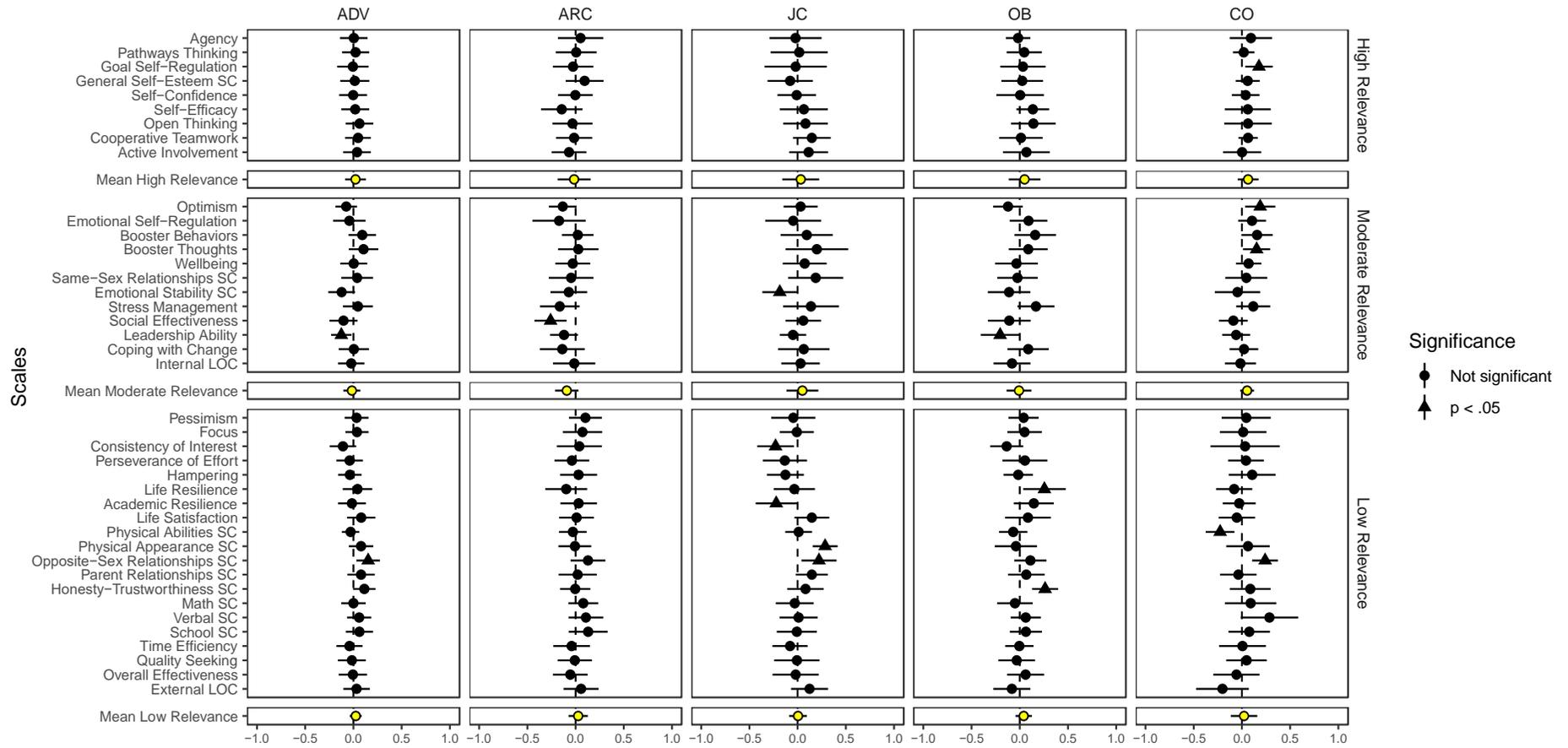


Figure 6.4. THP program effect sizes with 95% confidence intervals for Follow-Up Analysis (T2-T3) with scales grouped by relevance.

Note. T2 = immediate post-test; T3 = three months post-test; ADV = Adventure Programs (taken together); ARC = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; SC = Self-Concept; LOC = Locus of Control. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean effect for that group of scales. Effect sizes are standardised.

None of the high relevance scales demonstrated a significant follow-up effect at T3, whether as a gain or a decline over T2 scores when compared to the control group. Eight scales demonstrated significant positive follow-up effects, representing a gain over T2 scores when compared to the control group. Of these scales, three scales demonstrated no significant short-term effects, but did evidence significant long-term effects, indicating that growth in these outcomes occurred after program completion (i.e., a sleeper effect). Sleeper effects were reported by participants for the Adventure Programs (taken together) for Opposite-Sex Relationships Self-Concept (ES = .153), for the James Craig program for Opposite-Sex Relationships Self-Concept (ES = .221), for the Outward Bound program for Honesty-Trustworthiness Self-Concept (ES = .263), and for the Coaching Only Program for Optimism (ES = .190) and Opposite-Sex Relationships Self-Concept (ES = .240). It is interesting to note these follow-up effects for Opposite-Sex Relationships Self-Concept, for which there were no significant positive short-term effects but significant positive long-term effects for the Adventure Programs (taken together), as well as the James Craig and Coaching Only Program. There was also a significant follow-up effect reported by the James Craig participants for Physical Appearance Self-Concept (ES = .285), although neither the short nor long-term effects were significant. This result is similar to the significant follow-up effect reported by the Coaching Only Program participants for Booster Thoughts (ES = .152), for which there were no significant short- or long-term effects. The Outward Bound participants also reported a significant positive follow-up effect in Life Resilience (ES = .256), for which there were no significant short- or long-term effects. Six of the moderate and low relevance scales evidenced significant negative follow-up effects, however, none of these scales evidenced either short-term or long-term significant effects.

Overall, participant factor scores for the outcome variables were relatively stable during the 3-month period immediately following completion of the THP program, when compared to the control group. None of the scales most relevant to the THP program design and aims evidenced any significant follow-up effects. However, there appeared to be a sleeper effect for Opposite-Sex Relationships Self-Concept that was consistent across a number of programs.

Results of Research Hypothesis 2.4: Aptitude-treatment interaction

effects. Research Hypothesis 2.4 predicted that the analysis of the THP program effects on each outcome variable would also evidence a statistically significant aptitude-treatment interaction effect between the main effect of treatment (i.e., program participation) and the participants' pre-test aptitude in that outcome variable (represented by the factor scores for that variable at T1). Accordingly, it was anticipated that participants who were lower in the outcome variables of interest prior to participation in a THP program, would benefit more from the program through higher levels of the relevant outcome variables after the program, when compared to the control group participants at a similar aptitude level.

Overall, there were 47 significant aptitude-treatment interactions of a possible 410 aptitude-treatment interactions (11.46%): 32 significant interaction effects of a possible 205 (15.61%) in the Short-Term Analysis; and 15 significant interaction effects of a possible 205 (7.30%) in the Long-Term Analysis. These significant results are presented in Table 6.9. A copy of the visual summary of results presented in Table 6.5 for the Short-Term and Long-Term Analyses is reproduced below in Table 6.10 with an overlay representing the significant aptitude-treatment interaction effects (indicated by a double-lined box). Over the two sets of wave comparisons for the five different group comparisons, those scales with any significant interaction effects demonstrated on average 1.85 significant interaction effects, however, the Verbal Self-Concept scale had five significant interaction effects with all being in the hypothesised direction (lower T1 levels of an outcome variable demonstrating higher T2 or T3 scores for that outcome variable when compared to the control group). Moreover, 37 of the 47 interactions (78.72%) were in respect of scales that did not demonstrate significant group main effects in the initial analyses (25 in respect of the Short-Term Analysis and 12 in respect of the Long-Term Analysis). Accordingly, it is relevant to consider the implications of these aptitude-treatment interaction effects in more detail.

Table 6.9

Aptitude-Treatment Interaction Effects: Parameter Coefficients for Significant Effects

Scale	Parameter	Aptitude-Treatment Interaction Effects (T1 Scale Factor Score x Program Group): ES (SE)									
		<u>Adv vs. C</u>		<u>Arctos vs. C</u>		<u>James Craig vs. C</u>		<u>Outward Bound vs. C</u>		<u>Coaching Only vs. C</u>	
		T2	T3	T2	T3	T2	T3	T2	T3	T2	T3
<i>High Relevance</i>											
Agency	b1	.249 (.086)**		.147 (.151)		.228 (.103)*					
	b2	.292 (.122)*		.224 (.130)		.299 (.121)*					
	b3	-.178 (.087)*		.185 (.079)*		-.367 (.125)**					
Pathways Thinking	b1	.247 (.093)**				.467 (.112)***	.262 (.178)				
	b2	.348 (.115)**				.333 (.115)**	.085 (.112)				
	b3	-.190 (.065)**				-.286 (.070)***	-.223 (.078)**				
Cooperative Teamwork	b1	.193 (.104)		.162 (.121)				.104 (.119)			
	b2	.600 (.121)***		.571 (.118)***				.571 (.118)***			
	b3	-.263 (.078)***		-.234 (.081)**				-.413 (.158)**			
<i>Moderate Relevance</i>											
Optimism	b1	.101 (.088)						-.049 (.133)			
	b2	.520 (.077)***						.531 (.079)***			
	b3	-.248 (.088)**						-.467 (.109)***			
Booster Behaviours	b1		.219 (.095)*						.161 (.105)		
	b2		.234 (.089)**						.207 (.088)*		
	b3		.190 (.086)*						.346 (.085)***		
Wellbeing	b1							.261 (.195)			
	b2							.472 (.115)***			
	b3							-.279 (.134)*			
Stress Management	b1				.339 (.230)			.122 (.145)			
	b2				.238 (.236)			.487 (.218)*			
	b3				.185 (.095)*			.308 (.146)*			
Leadership Ability	b1							.133 (.164)			
	b2							.715 (.068)***			
	b3							-.274 (.103)**			

Note. T1 = pre-test; ES = standardised effect size; SE = standard error; Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. T2 = significant Aptitude-Treatment interaction effects for the Short-Term Analysis (pre-test to immediate post-test) and T3 = significant Aptitude-Treatment interaction effects for the Long-Term Analysis (pre-test to three months post-test). High and Moderate Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. Only parameter coefficients for significant interaction effects are shown: b1 = main effect coefficient for program group; b2 = main effect coefficient for T1 scale factor score; b3 = Aptitude-Treatment interaction effect coefficient. Significant parameters are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

(continues)

Table 6.9 (continued)

Aptitude-Treatment Interaction Effects: Parameter Coefficients for Significant Effects

Scale	Parameter	<u>Aptitude-Treatment Interaction Effects (T1 Scale Factor Score x Program Group): ES (SE)</u>									
		<u>Adv vs. C</u>		<u>Arctos vs. C</u>		<u>James Craig vs. C</u>		<u>Outward Bound vs. C</u>		<u>Coaching Only vs. C</u>	
		T2	T3	T2	T3	T2	T3	T2	T3	T2	T3
<i>Low Relevance</i>											
Pessimism ^a	b1					-.022 (.109)					
	b2					.501 (.109)***					
	b3					-.340 (.129)**					
Focus	b1							.177 (.104)			-.348 (.300)
	b2							.262 (.100)**			.262 (.100)**
	b3							.320 (.115)**			.446 (.148)**
Consistency of Interest	b1										-.597 (.136)***
	b2										.321 (.083)***
	b3										.580 (.251)*
Perseverance of Effort	b1				.016 (.105)	.369 (.160)*					-.074 (.094)
	b2				.359 (.101)***	.381 (.100)***					.359 (.101)***
	b3				-.209 (.104)*	.307 (.110)**					.295 (.095)**
Life Resilience	b1	.106 (.089)		.049 (.132)							
	b2	.112 (.090)		.101 (.091)							
	b3	.164 (.071)*		.209 (.092)*							
Academic Resilience	b1		-.012 (.100)		-.028 (.103)						
	b2		.478 (.106)***		.484 (.101)***						
	b3		-.231 (.089)**		-.304 (.136)*						
Life Satisfaction	b1										
	b2										
	b3										
Physical Abilities SC	b1										
	b2										
	b3										

Note. T1 = pre-test; ES = standardized effect size; SE = standard error; Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. T2 = significant Aptitude-Treatment interaction effects for the Short-Term Analysis (pre-test to immediate post-test) and T3 = significant Aptitude-Treatment interaction effects for the Long-Term Analysis (pre-test to three months post-test). Low Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. SC = Self-Concept. Only parameter coefficients for significant interaction effects are shown: b1 = main effect coefficient for program group; b2 = main effect coefficient for T1 scale factor score; b3 = Aptitude-Treatment interaction effect coefficient. Significant parameters are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for this scale were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

(continues)

Table 6.9 (continued)

Aptitude-Treatment Interaction Effects: Parameter Coefficients for Significant Effects

Scale	Parameter	<u>Aptitude-Treatment Interaction Effects (T1 Scale Factor Score x Program Group): ES (SE)</u>									
		<u>Adv vs. C</u>		<u>Arctos vs. C</u>		<u>James Craig vs. C</u>		<u>Outward Bound vs. C</u>		<u>Coaching Only vs. C</u>	
		T2	T3	T2	T3	T2	T3	T2	T3	T2	T3
<i>Low Relevance</i>											
Opposite-Sex Rel'ships SC	b1			.239 (.113)*							.157 (.159)
	b2			.572 (.080)***							.572 (.080)***
	b3			.239 (.107)*							.429 (.132)***
Math SC	b1				.171 (.109)			.193 (.123)			
	b2				.557 (.067)***			.655 (.078)***			
	b3				.256 (.103)*			-.181 (.091)*			
Verbal SC	b1	.158 (.080)*		.156 (.102)	.172 (.129)			-.028 (.122)			.419 (.104)***
	b2	.732 (.081)***		.728 (.078)***	.668 (.084)***			.728 (.078)***			.668 (.084)***
	b3	-.192 (.063)**		-.217 (.080)**	-.192 (.071)**			-.276 (.089)**			-.411 (.079)***
School SC	b1										-.051 (.203)
	b2										.524 (.144)**
	b3										.220 (.103)*
Time Efficiency	b1										-.193 (.258)
	b2										.248 (.102)*
	b3										.232 (.084)**
Quality Seeking	b1										-.040 (.113)
	b2										.277 (.111)*
	b3										.234 (.089)**
Overall Effectiveness	b1										.144 (.159)
	b2										.301 (.128)*
	b3										.150 (.072)*
External Locus of Control ^a	b1										-.312 (.086)***
	b2										.388 (.090)***
	b3										.460 (.196)*

Note. T1 = pre-test; ES = standardised effect size; SE = standard error; Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. T2 = significant Aptitude-Treatment interaction effects for the Short-Term Analysis (pre-test to immediate post-test) and T3 = significant Aptitude-Treatment interaction effects for the Long-Term Analysis (pre-test to three months post-test). Low Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. SC = Self-Concept and Rel'ships = Relationships. Only parameter coefficients for significant interaction effects are shown: b1 = main effect coefficient for program group; b2 = main effect coefficient for T1 scale factor score; b3 = Aptitude-Treatment interaction effect coefficient. Significant parameters are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for this scale were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

Table 6.10
 Summary Results of Effects of THP Programs including Aptitude-Treatment Interaction Effects

Scale	Summary Program Effects									
	Adv vs. C		Arc vs. C		JC vs. C		OB vs. C		CO vs. C	
	T2	T3	T2	T3	T2	T3	T2	T3	T2	T3
<i>High Relevance</i>										
Agency	++	ns	ns	ns	+	ns	+	ns	ns	ns
Pathways Thinking	++	ns	ns	ns	+++	ns	ns	ns	ns	ns
Goal Self-Regulation	ns	ns	ns	ns	++	ns	ns	ns	ns	ns
General Self-Esteem/SC	ns	ns	ns	ns	+	ns	ns	ns	ns	ns
Self-Confidence	+	ns	ns	ns	+++	ns	ns	ns	ns	ns
Self-Efficacy	++	ns	ns	ns	+++	ns	ns	+	ns	ns
Open Thinking	+	ns	ns	ns	++	ns	ns	ns	ns	ns
Cooperative Teamwork	ns	+	ns	ns	ns	+	ns	ns	ns	ns
Active Involvement	+	ns	ns	ns	++	+	ns	ns	ns	ns
<i>Moderate Relevance</i>										
Optimism	ns	ns	ns	ns	+++	+	ns	ns	ns	+
Emotion Self-Regulation	ns	ns	ns	ns	+	ns	ns	ns	ns	ns
Booster Behaviours	ns	+	ns	ns	+++	ns	ns	ns	ns	ns
Booster Thoughts	ns	ns	ns	ns	+	+	ns	ns	ns	ns
Wellbeing	++	ns	ns	ns	+++	+	ns	ns	ns	ns
Same-Sex Relationships SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Emotional Stability SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Stress Management	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Social Effectiveness	+	ns	ns	ns	+	+	ns	ns	+	ns
Leadership Ability	ns	ns	ns	ns	+++	+	ns	ns	ns	ns
Coping with Change	ns	ns	ns	ns	+	ns	ns	ns	ns	ns
Internal Locus of Control	ns	ns	ns	ns	+++	+	ns	ns	ns	ns
<i>Low Relevance</i>										
Pessimism ^a	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Focus	++	ns	ns	ns	+++	ns	ns	ns	ns	ns
Consistency of Interest	ns	ns	ns	ns	ns	ns	ns	ns	---	ns
Perseverance of Effort	ns	ns	ns	ns	ns	ns	ns	ns	---	ns
Hampering ^a	+	ns	ns	ns	+	ns	+	ns	ns	ns
Life Resilience	ns	ns	ns	ns	++	ns	ns	ns	ns	ns
Academic Resilience	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Life Satisfaction	ns	ns	ns	ns	ns	+	ns	ns	ns	ns
Physical Abilities SC	ns	ns	++	ns	ns	ns	ns	ns	ns	ns
Physical Appearance SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Opposite-Sex Rel'ships SC	ns	+	ns	ns	ns	+++	ns	ns	ns	+
Parent Relationships SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Honesty-Trustworthiness SC	ns	ns	ns	ns	ns	ns	ns	+	ns	ns
Math SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Verbal SC	+	ns	ns	ns	+++	ns	ns	ns	+	+
School SC	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
Time Efficiency	ns	ns	ns	ns	+++	ns	ns	ns	ns	ns
Quality Seeking	+	ns	ns	ns	+++	ns	ns	ns	ns	ns
Overall Effectiveness	ns	ns	ns	ns	+++	ns	ns	ns	ns	ns
External Locus of Control ^a	ns	ns	ns	ns	ns	ns	ns	ns	---	-

Note. Adv = Adventure Programs (taken together); Arc = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; C = waitlist control group. T₂ = group differences in the factor scores for the scale at T₂, controlling for differences at T₁ as well as other covariates in the model and T₃ = group differences in the factor scores for the scale at T₃, controlling for differences at T₁ as well as other covariates in the model. SC = Self-Concept; Rel'ships = Relationships; LOC = Locus of Control. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean difference for that group of scales. + indicates a significant positive effect for THP program participants when compared to control group and - indicates a significant negative effect for THP program participants when compared to control group: + or - ($p < .05$); ++ or -- ($p < .01$); +++ or --- ($p < .001$); ns (no significant effect). Double-lined box indicates a significant Aptitude-Treatment

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

Example of aptitude-treatment interaction effect. The Cooperative Teamwork scale provides a useful example of an aptitude-treatment interaction effect because there was a significant interaction effect in the Short-Term Analysis for the Arctos Adventure Program participants, but this scale did not demonstrate a significant main short-term effect for program participation when compared to the control group. In this example, y represents T2 Cooperative Teamwork (the dependent variable), x represents the treatment predictor variable (i.e., the program group variable), and z represents T1 Cooperative Teamwork (the moderator variable). The other covariates in the regression model have been given a value of zero for this analysis (i.e., the mean value, as these variables are standardised to have mean = 0). The regression equation of interest is as follows:

$$\hat{y} = \hat{b}_0 + \hat{b}_1x + \hat{b}_2z + \hat{b}_3xz$$

where \hat{b}_0 is the intercept of the equation and \hat{b}_1 , \hat{b}_2 , and \hat{b}_3 are the regression coefficients for treatment (i.e., program participation), T1 Cooperative Teamwork (i.e., aptitude), and aptitude-treatment interaction, respectively. The relevant coefficients from the regression analysis that included Cooperative Teamwork as a dependent variable, are set out in Table 6.11. If there were no interaction term, \hat{b}_1 would be interpreted as the unique effect of program participation on T2 Cooperative Teamwork. However, the significant interaction term means that the effect of program participation on T2 Cooperative Teamwork differs for different values of T1 Cooperative Teamwork. Therefore, the effect of program participation is limited to \hat{b}_1 only when T1 Cooperative Teamwork equals zero (i.e., mean T1 Cooperative Teamwork), and for other values of T1 Cooperative Teamwork, the effect of program participation depends on the value of T1 Cooperative Teamwork and \hat{b}_3 (i.e., $\hat{b}_1 + \hat{b}_3 \cdot \text{T1 Cooperative Teamwork}$; referred to as the *simple slopes* or *simple effects*).

Table 6.11
Significant Aptitude-Treatment Interactions on T2 Cooperative Teamwork for Arctos Adventure Program Participants

Parameter	Arctos Adventure Program: ES (SE)
Main effect of group (b1)	.162 (.121)
Main effect of T1 Cooperative Teamwork (b2)	.571 (.118)***
Aptitude-treatment interaction (b3)	-.234 (.081)**

Note. T2 = immediate post-test; ES = standardised effect size; SE = standard error; T1 = pre-test. ** $p < .01$; *** $p < .001$.

In order to better understand the nature of the relationship between T1 Cooperative Teamwork and the effect of program participation, it is helpful to consider the simple main effect of program participation at different levels of T1 Cooperative Teamwork (referred to as *conditional values*).

The differing effects for participants with low, medium, and high levels of T1 Cooperative Teamwork are represented graphically in Figure 6.5, with the x-axis representing the different baseline levels of T1 Cooperative Teamwork, the y-axis reflecting the levels of T2 Cooperative Teamwork, and a separate line for each of the control group and Arctos Adventure Program participants, demonstrating the differences in T2 Cooperative Teamwork between the Arctos Adventure Program participants and the control group at each level of T1 Cooperative Teamwork.

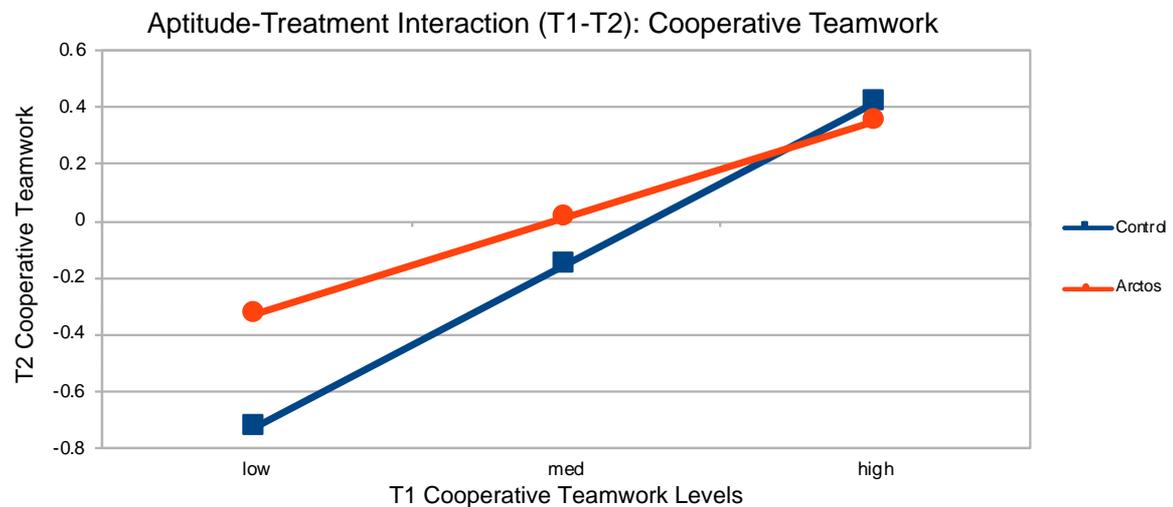


Figure 6.5. Factor scores for T2 Cooperative Teamwork at different T1 Cooperative Teamwork baseline levels for the waitlist control and Arctos Adventure Program groups.

Note. Control = waitlist control group; Arctos = Arctos Adventure Program; low = -1; med = 0; and high = +1. For Arctos, T1 = pre-test; T2 = immediate post-test. For Control, T1 and T2 = extended baseline pre-test.

Participants in the Arctos Adventure Program with a mean level of T1 Cooperative Teamwork (conditional value = 0) demonstrate an effect equal to \hat{b}_1 (ES = .162, $SE = .121$, $p > .05$), which reflects a T2 Cooperative Teamwork level that is not significantly different to the control group with a similar T1 Cooperative Teamwork baseline level. However, Arctos participants with a low baseline level of T1 Cooperative Teamwork (conditional value = -1) demonstrate an effect equal to $\hat{b}_1 + \hat{b}_3*(-1)$ (ES = .396, $SE = .146$, $p < .01$), which reflects a significantly higher T2 Cooperative Teamwork score when compared to the relevant control group, being an effect that is greater than the main effect, consistent with Research Hypothesis 2.4. Arctos participants with a high baseline level of T1 Cooperative Teamwork (conditional value = 1) demonstrate an effect equal to $\hat{b}_1 + \hat{b}_3*(1)$ (ES = -.072, $SE = .144$, $p > .05$), which reflects no significant difference in T2 Cooperative Teamwork scores when compared to the relevant control group.

Calculation of simple main effects. For each of the 47 significant aptitude-treatment interactions, simple main effects were calculated for the three conditional values used in the example above. These results are presented below by wave in Table 6.12 for the Short-Term Analysis and Table 6.13 for the Long-Term Analysis. The aptitude-treatment interaction analysis found 18 significant positive

simple main effects for program participants with low aptitude in the relevant outcome variable, 15 in the Short-Term Analysis and three in the Long-Term Analysis. Eight of these significant positive effects related to high relevance scales, three to moderate relevance scales, and seven to low relevance scales. Five of these significant positive effects related to participants in the Adventure Programs (taken together), four related to the Arctos participants, four related to the James Craig participants, four related to the Outward Bound participants, and one related to the Coaching Only participants. Of these 18 significant positive simple main effects for low aptitude participants, seven also demonstrated significant positive simple main effects for participants with a mean level of aptitude. However, for three of these 18 positive simple main effects for low aptitude participants, significant negative simple main effects were found for program participants with a high level aptitude in those outcome variables. There were also five significant negative simple main effects for program participants with low aptitude in the relevant outcome variable. These negative simple main effects all related to low relevance scales, and four related to participants in the Coaching Only Program.

Aptitude-treatment interaction effects for the Short-Term Analysis

(T1-T2). The simple main effects for the Short-Term Analysis, arranged by scale and grouped by scale relevance, are set out in Table 6.12. Of the nine high relevance scales, three scales evidenced significant aptitude-treatment interaction effects (with seven interactions in total) for the Adventure Programs (Agency, Pathways Thinking, and Cooperative Teamwork), as well as the Arctos (Cooperative Teamwork), James Craig (Agency and Pathways Thinking), and Outward Bound (Cooperative Teamwork) programs. All of these interactions demonstrated significant positive simple main effects for low-aptitude participants, which were higher than the effects for other participants. The interactions for Agency and Pathways Thinking also evidenced significant positive simple main effects for medium-aptitude participants, although with lower effect sizes. However, the interaction for Cooperative Teamwork also evidenced a significant negative simple main effect for high-aptitude Outward Bound participants.

Table 6.12
Aptitude-Treatment Interaction Effects for Short-Term Analysis (T1-T2)

Scale	Wave	Program	Simple Main Effects: ES (SE)		
			Low T1	Med T1	High T1
<i>High Relevance</i>					
Agency	T1-T2	Adventure Programs	.427 (.126)***	.249 (.086)**	.071 (.119)
Agency	T1-T2	James Craig	.595 (.183)***	.228 (.103)*	-.139 (.138)
Cooperative Teamwork	T1-T2	Adventure Programs	.456 (.144)**	.193 (.104)	-.070 (.114)
Cooperative Teamwork	T1-T2	Arctos	.396 (.146)**	.162 (.121)	-.072 (.144)
Cooperative Teamwork	T1-T2	Outward Bound	.517 (.233)*	.104 (.119)	-.309 (.155)*
Pathways Thinking	T1-T2	Adventure Programs	.437 (.124)***	.247 (.093)**	.057 (.102)
Pathways Thinking	T1-T2	James Craig	.753 (.147)***	.467 (.112)***	.180 (.116)
<i>Moderate Relevance</i>					
Optimism	T1-T2	Adventure Programs	.350 (.138)*	.101 (.088)	-.147 (.109)
Optimism	T1-T2	Outward Bound	.418 (.166)*	-.049 (.133)	-.517 (.177)**
Wellbeing	T1-T2	Outward Bound	.540 (.254)*	.261 (.195)	-.018 (.218)
Stress Management	T1-T2	Outward Bound	-.186 (.258)	.122 (.145)	.429 (.136)**
Leadership Ability	T1-T2	Outward Bound	.406 (.215)	.133 (.164)	-.141 (.169)
<i>Low Relevance</i>					
Pessimism ^a	T1-T2	James Craig	.318 (.154)*	-.022 (.109)	-.362 (.182)*
Focus	T1-T2	Outward Bound	-.143 (.172)	.177 (.104)	.497 (.136)***
Focus	T1-T2	Coaching Only	-.794 (.322)*	-.348 (.300)	.098 (.192)
Consistency of Interest	T1-T2	Coaching Only	-1.176 (.328)***	-.597 (.136)***	-.017 (.234)
Perseverance of Effort	T1-T2	James Craig	.062 (.188)	.369 (.160)*	.676 (.200)***
Life Resilience	T1-T2	Adventure Programs	-.059 (.108)	.106 (.089)	.270 (.119)*
Life Resilience	T1-T2	Arctos	-.160 (.130)	.049 (.132)	.257 (.186)
Life Resilience	T1-T2	Outward Bound	-.496 (.226)*	-.073 (.127)	.350 (.204)
Life Satisfaction	T1-T2	Outward Bound	.138 (.165)	-.131 (.151)	-.400 (.201)*
Physical Abilities SC	T1-T2	Arctos	.372 (.108)***	.233 (.080)**	.094 (.103)
Opposite-Sex Rel'ships SC	T1-T2	Arctos	.000 (.167)	.239 (.113)*	.478 (.143)***
Opposite-Sex Rel'ships SC	T1-T2	Coaching Only	-.272 (.197)	.157 (.121)	.586 (.159)***
Math SC	T1-T2	Outward Bound	.374 (.161)*	.193 (.123)	.011 (.145)
Verbal SC	T1-T2	Adventure Programs	.350 (.091)***	.158 (.080)*	-.033 (.112)
Verbal SC	T1-T2	Arctos	.372 (.116)***	.156 (.102)	-.061 (.141)
Verbal SC	T1-T2	Outward Bound	.248 (.143)	-.028 (.122)	-.304 (.172)
School SC	T1-T2	Coaching Only	-.270 (.284)	-.051 (.230)	.169 (.152)
Time Efficiency	T1-T2	Coaching Only	-.425 (.313)	-.193 (.258)	-.193 (.258)
Quality Seeking	T1-T2	Coaching Only	-.274 (.167)	-.040 (.113)	.194 (.116)
Overall Effectiveness	T1-T2	Coaching Only	-.006 (.190)	.144 (.159)	.294 (.156)

Note. ES = standardised effect size; SE = standard error. Adventure Programs = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; Coaching Only = Coaching Only Program; Low T1 = simple main effect for T1 factors scores = -1; Med T1 = simple main effect for T1 factors scores = 0; High T1 = simple main effect for T1 factor scores = +1; SC = Self-Concept; Rel'ships = Relationships; T1 = pre-test factor scores; T2 = immediate post-test factor scores. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. Significant effects are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

Of the 12 moderate relevance scales, four scales evidenced significant aptitude-treatment interaction effects (with five interactions in total) for the Adventure Programs (Optimism), as well as the Outward Bound program (Optimism, Wellbeing, Stress Management, and Leadership Ability). The

interactions for Optimism and Wellbeing demonstrated significant positive simple main effects for low-aptitude participants, which were higher than the effects for other participants. However, one of the interactions for Optimism also evidenced significant negative simple main effects for high-aptitude Outward Bound participants. Stress Management reflected an interaction result inconsistent with the hypothesised result, with significant positive simple main effects for high-aptitude Outward Bound participants and no significant simple main effects for low- or medium-aptitude participants.

Of the 20 low relevance scales, 14 scales evidenced significant aptitude-treatment interaction effects (with 20 interactions in total) across all programs. However, only four of these scales demonstrated significant positive simple main effects for low-aptitude participants: Pessimism for James Craig; Physical Abilities Self-Concept for Arctos; Math Self-Concept for Outward Bound, and Verbal Self-Concept for the Adventure Programs and Arctos. Some of these scales also demonstrated significant positive simple main effects for medium-aptitude participants, however, Pessimism demonstrated significant negative simple main effects for high-aptitude James Craig participants. Moreover, a number of scales reflected interactions with results inconsistent with the hypothesis, with either significant positive simple main effects for medium- or high-aptitude participants and no significant simple main effect for low-aptitude participants (Focus for Outward Bound participants, Perseverance of Effort for James Craig participants, Life Resilience for Adventure Program participants, and Opposite-Sex Relationships Self-Concept for Arctos and Coaching Only participants), or significant negative simple main effects for low-aptitude participants (Focus and Consistency of Interest for Coaching Only participants and Life Resilience for Outward Bound participants).

Considering just the Adventure Programs (taken together) and the Coaching Only Program, the Adventure Programs evidenced six significant short-term aptitude-treatment interaction effects, five of which were in the hypothesised direction with low-aptitude participants demonstrating significant positive simple main effects. Moreover, there were no significant negative simple main effects for participants at any level of aptitude in the Adventure Programs. On the other hand,

the Coaching Only Program evidenced seven significant short-term aptitude-treatment interaction effects, none of which were in the hypothesised direction and two of which demonstrated significant negative simple main effects for low- and medium-aptitude participants.

In total, there were 32 significant aptitude-treatment interaction effects in connection with the group main effects at immediate post-test (T2) out of a possible 205 effects. Of those 32 effects, 15 evidenced significant positive simple main effects for the low-aptitude participants that were greater than the effects for other participants, consistent with Research Hypothesis 2.4, with nine of those effects being in respect of scales that did not demonstrate significant group main effects in the Short-Term Analysis. Two of these were for the Arctos participants (Cooperative Teamwork and Verbal Self-Concept) and four were for the Outward Bound participants (Cooperative Teamwork, Optimism, Wellbeing, and Math Self-Concept), being individual programs which otherwise demonstrated weaker results in the Short-Term Analysis.

Aptitude-treatment interaction effects for the Long-Term Analysis (T1-T3). The simple main effects for the Long-Term Analysis, arranged by scale relevance, are set out in Table 6.13. Of the nine high relevance scales, two scales evidenced significant aptitude-treatment interaction effects in the Long-Term Analysis (Agency for Arctos participants and Pathways Thinking for James Craig participants). Only the interaction for Pathways Thinking demonstrated a significant positive simple main effect for low-aptitude James Craig participants. The Agency scale reflected an interaction result inconsistent with the hypothesised result, with significant positive simple main effects for high-aptitude Arctos participants and no significant simple main effects for low-aptitude participants. There were no significant negative simple main effects in the Long-Term Analysis for participants at any aptitude level for any high relevance scale.

Table 6.13
Aptitude-Treatment Interaction Effects for Long-Term Analysis (T1-T3)

Scale	Wave	Program	Simple Main Effects: ES (SE)		
			Low T1	Med T1	High T1
<i>High Relevance</i>					
Agency	T1-T3	Arctos	-.038 (.176)	.147 (.151)	.332 (.164)*
Pathways Thinking	T1-T3	James Craig	.485 (.210)*	.262 (.178)	.040 (.176)
<i>Moderate Relevance</i>					
Booster Behaviours	T1-T3	Adventure Programs	.029 (.127)	.219 (.098)*	.409 (.126)***
Booster Behaviours	T1-T3	Outward Bound	-.184 (.133)	.161 (.105)	.507 (.137)***
Stress Management	T1-T3	Arctos	-.245 (.159)	-.059 (.138)	.126 (.174)
<i>Low Relevance</i>					
Perseverance of Effort	T1-T3	Arctos	.225 (.129)	.016 (.105)	-.194 (.165)
Perseverance of Effort	T1-T3	Coaching Only	-.369 (.133)**	-.074 (.094)	.220 (.134)
Academic Resilience	T1-T3	Adventure Programs	.219 (.135)	-.012 (.100)	-.244 (.132)
Academic Resilience	T1-T3	Arctos	.276 (.152)	-.028 (.103)	-.333 (.188)
Academic Resilience	T1-T3	James Craig	.067 (.268)	-.154 (.195)	-.375 (.166)*
Math SC	T1-T3	Arctos	-.084 (.134)	.171 (.109)	.427 (.164)**
Verbal SC	T1-T3	Arctos	.364 (.158)*	.172 (.129)	-.020 (.136)
Verbal SC	T1-T3	Coaching Only	.830 (.148)***	.419 (.104)***	.007 (.111)
Quality Seeking	T1-T3	Coaching Only	-.219 (.140)	.025 (.076)	.269 (.074)***
External Locus of Control ^a	T1-T3	Coaching Only	-.772 (.225)***	-.312 (.086)***	.148 (.204)

Note. ES = standardised effect size; SE = standard error. Adventure Programs = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; Coaching Only = Coaching Only Program; Low T1 = simple main effect for T1 factors scores = -1; Med T1 = simple main effect for T1 factors scores = 0; High T1 = simple main effect for T1 factor scores = +1; SC = Self-Concept; T1 = pre-test factor scores; T2 = immediate post-test factor scores. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. Significant effects are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

Of the 12 moderate relevance scales, two scales evidenced significant aptitude-treatment interaction effects in the Long-Term Analysis (with three interactions in total) for the Adventure Programs (Booster Behaviours), as well as the Outward Bound (Booster Behaviours) and Arctos (Stress Management) programs. However, none of these interactions demonstrated significant positive simple main effects for low-aptitude participants. The Booster Behaviours scale reflected interaction results inconsistent with the hypothesis, with significant positive simple main effects for medium- and high-aptitude participants and no significant simple main effects for low-aptitude participants. However, there were no significant negative simple main effects for participants at any aptitude level for any moderate relevance scale.

Of the 20 low relevance scales, six scales evidenced significant aptitude-treatment interaction effects in the Long-Term Analysis (with 10 interactions in

total) across all programs other than Outward Bound. Only the Verbal Self-Concept scale demonstrated significant positive simple main effects for low-aptitude Arctos and Coaching Only participants. A number of the other scales reflected interaction results inconsistent with the hypothesised result, with either significant positive simple main effects for high-aptitude participants and no significant simple main effect for low-aptitude participants (Math Self-Concept for Arctos and Quality Seeking for Coaching Only) or significant negative simple main effects for low-aptitude participants (Perseverance of Effort and External Locus of Control for Coaching Only). Additionally, the Academic Resilience scale demonstrated significant negative simple main effects for high-aptitude James Craig participants.

Considering just the Adventure Programs (taken together) and the Coaching Only Program, the Adventure Programs evidenced two significant aptitude-treatment interaction effects, one of which was in the hypothesised direction but without significant simple main effects. The Coaching Only Program evidenced four significant interaction effects, with one in the hypothesised direction, but two reflecting significant negative simple main effects for low-aptitude participants.

In total, there were 15 significant aptitude-treatment interaction effects in connection with the group main effects approximately three months post program (T3) out of a possible 205 effects. Of those 15 effects, three evidenced significant positive simple main effects for the low-aptitude participants that were greater than the effects for other participants, consistent with Research Hypothesis 2.4, with two of those effects being in respect of scales that did not demonstrate significant group main effects in the Long-Term Analysis. One of these effects was for the Verbal Self-Concept scale for the Arctos participants (similar to the short-term interaction effects), and the other was for the Pathways Thinking scale for the James Craig participants.

Conclusion. There were a number of outcome variables where participants with a lower pre-test score in the outcome variable reported significantly higher T2 or T3 scores when compared to their respective control group, and this effect was greater than for participants with higher pre-test scores (as hypothesised). Some of these significant aptitude-treatment interaction effects in the hypothesised

direction were in respect of outcome variables that did not demonstrate a significant main group effect in the primary Short-Term and Long-Term Analyses. This was particularly the case for the Short-Term Analysis. On the other hand, some of these significant positive results for low-aptitude participants were accompanied by significant negative results for high-aptitude participants. Moreover, other results from the aptitude-treatment interaction analyses were in contrast to Research Hypothesis 2.4 with either participants with higher pre-test scores reporting significantly higher T2 or T3 scores when compared to their respective control group (without significantly higher scores for lower-aptitude participants) or those with lower pre-test scores reporting significantly lower T2 or T3 scores than their respective control group (particularly in the Coaching Only Program). Notwithstanding that some of these results are not in the direction hypothesised, almost half of the results demonstrate participants in some THP programs with low pre-test aptitude in some outcome variables, reporting significantly higher scores in those outcome variables following program participation, and in contrast to participants with higher aptitude in those outcome variables. These results are particularly interesting for the Arctos and Outward Bound programs, where the original analyses did not demonstrate many significant positive effects.

Results of Research Question 2.5: Differences in effects between Adventure Programs and Coaching Only Program. Research Hypotheses 2.1 and 2.2 predicted the THP program effects would be significant and positive in both the short and long term. However, given the novel nature of the Adventure Programs, there was an open question about the differential contribution to those effects by the adventure and coaching components of the programs. The Coaching Only Program was developed expressly for research purposes in order to evaluate the incremental impact of the adventure component within the integrated coaching and adventure THP program. Accordingly, Research Question 2.5 asked whether there would be any significant differences in effects in the Short-Term, Long-Term, and Follow-Up Analyses between the Adventure Programs and Coaching Only Program. The results of this analysis are presented in Table 6.14.

Table 6.14
Comparison of Adventure Programs and Coaching Only Program Effects for all Waves

Scale	Program Effects of Adventure vs. Coaching Only: ES (SE)		
	T2	T3a	T3b
<i>High Relevance</i>			
Agency	.386 (.097)***	.122 (.133)	-.088 (.122)
Pathways Thinking	.328 (.163)*	.178 (.110)	.004 (.076)
Goal Self-Regulation	.278 (.183)	-.022 (.166)	-.182 (.085)*
General Self-Esteem/SC	.149 (.184)	.128 (.148)	-.047 (.059)
Self-Confidence	.145 (.138)	.105 (.140)	-.042 (.087)
Self-Efficacy	.122 (.247)	.088 (.207)	-.042 (.121)
Open Thinking	.268 (.135)*	.182 (.182)	.001 (.130)
Cooperative Teamwork	.230 (.083)**	.186 (.090)*	-.016 (.067)
Active Involvement	.244 (.095)**	.200 (.143)	.036 (.107)
<i>Mean for High Relevance</i>	.239 (.106)*	.129 (.122)	-.042 (.063)
<i>Moderate Relevance</i>			
Optimism	.019 (.119)	-.180 (.086)*	-.265 (.079)***
Emotion Self-Regulation	.305 (.210)	.014 (.146)	-.148 (.078)
Booster Behaviours	.187 (.307)	.129 (.204)	-.065 (.087)
Booster Thoughts	.030 (.176)	.091 (.143)	-.047 (.086)
Wellbeing	.112 (.191)	.096 (.140)	-.066 (.066)
Same-Sex Relationships SC	-.088 (.128)	-.011 (.135)	-.007 (.121)
Emotional Stability SC	-.075 (.207)	-.095 (.109)	-.076 (.110)
Stress Management	.148 (.256)	.089 (.155)	-.069 (.088)
Social Effectiveness	-.049 (.084)	-.015 (.078)	-.014 (.065)
Leadership Ability	-.057 (.114)	-.036 (.094)	-.064 (.080)
Coping with Change	.124 (.202)	.105 (.154)	-.017 (.079)
Internal Locus of Control	.267 (.114)*	.198 (.140)	-.007 (.099)
<i>Mean for Moderate Relevance</i>	<i>.077 (.120)</i>	<i>.032 (.100)</i>	<i>-.070 (.042)</i>
<i>Low Relevance</i>			
Pessimism ^a	.213 (.201)	.013 (.084)	-.012 (.134)
Focus	.445 (.320)	.151 (.178)	.024 (.128)
Consistency of Interest	.524 (.120)***	.080 (.190)	-.142 (.190)
Perseverance of Effort	.422 (.118)***	.060 (.099)	-.080 (.112)
Hampering ^a	.420 (.241)	.057 (.084)	-.144 (.124)
Life Resilience	.240 (.189)	.237 (.177)	.121 (.109)
Academic Resilience	.411 (.271)	.159 (.177)	.012 (.085)
Life Satisfaction	-.190 (.194)	.129 (.134)	.134 (.089)
Physical Abilities SC	.128 (.112)	.144 (.100)	.198 (.082)*
Physical Appearance SC	-.063 (.187)	-.054 (.142)	.016 (.127)
Opposite-Sex Relationships SC	-.038 (.166)	-.091 (.088)	-.087 (.080)
Parent Relationships SC	-.066 (.096)	.009 (.098)	.116 (.092)
Honesty-Trustworthiness SC	.254 (.226)	.043 (.190)	.028 (.094)
Math SC	.183 (.263)	-.037 (.103)	-.090 (.139)
Verbal SC	-.109 (.082)	-.187 (.109)	-.225 (.155)
School SC	.136 (.209)	.030 (.103)	-.014 (.115)
Time Efficiency	.305 (.281)	.015 (.177)	-.047 (.121)
Quality Seeking	.239 (.100)*	.039 (.081)	-.064 (.119)
Overall Effectiveness	.011 (.177)	.071 (.116)	.052 (.129)
External Locus of Control ^a	.252 (.087)**	.309 (.129)*	.235 (.130)
<i>Mean for Low Relevance</i>	<i>.186 (.145)</i>	<i>.059 (.071)</i>	<i>.002 (.070)</i>

Note. ES = standardised effect sizes; SE = standard error. Adventure = Adventure Programs (taken together); Coaching Only = Coaching Only Program. T2 = group differences in the factor scores for the scale at T2, controlling for differences at T1 as well as other covariates in the model; T3a = group differences in the factor scores for the scale at T3, controlling for differences at T1 as well as other covariates in the model; T3b = group differences in the factor scores for the scale at T3, controlling for differences at T1 and T2 as well as other covariates in the model; SC = Self-Concept. Low Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for the Low Relevance scales represents the mean effect for that group of scales. Positive ES favors Adventure and negative ES favors Coaching Only. Significant differences are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

Of the 123 effect comparisons (41 comparisons for each of the Short-Term, Long-Term, and Follow-Up Analyses), there were 16 significant differences in effects (11.38%). In the Short-Term Analysis, there were 10 scales with significant differences in effects, all of which reflected significantly greater effects for the participants in the Adventure Programs. Five of these scales were high relevance scales (Agency, Pathways Thinking, Open Thinking, Cooperative Teamwork, and Active Involvement), one was a moderate relevance scale (Internal Locus of Control), and four were low relevance scales (Consistency of Interest, Perseverance of Effort, Quality Seeking, and External Locus of Control). Effect sizes reflecting the difference ranged from $ES = .230$ (Cooperative Teamwork) to $ES = .524$ (Consistency of Interest). Moreover, the high relevance scales as a group demonstrated a significant mean effect, indicating that in the short-term the outcomes in the highly relevant scales (as a group) were on average significantly greater for participants in the Adventure Programs than for participants in the Coaching Only Program (mean $ES = .239$, $SE = .106$).

In the Long-Term Analysis, however, there were only three scales with significant differences in effects. Two of these scales (Cooperative Teamwork and External Locus of Control) reflected significantly greater long-term effects for the participants in the Adventure Programs and were consistent with the short-term differences in effects for those scales. On the other hand, the Optimism scale evidenced significantly greater long-term effects for the Coaching Only Program participants. For the Follow-up Analysis, the Optimism and Goal Self-Regulation scales both reflected significantly greater follow-up effects for the Coaching Only Program participants, while the Physical Abilities Self-Concept scale reflected significantly greater follow-up effects for the participants in the Adventure Programs.

In summary, the Adventure Programs evidenced significantly greater effects in the short term for a number of the outcomes, particularly those outcomes that were most relevant to the THP program design and aims. This suggests that the adventure component may have an important role to play. However, over the longer term there were only a few significant differences in effects between the two programs and not all of these differences favoured the Adventure Programs.

Accordingly, there remains an open question as to the benefits of the adventure component for the THP program over the longer term.

Results of Research Question 2.6: Replication of RCT results with waitlist control group data. Research Hypothesis 2.1 predicted the experimental comparison between the participants in the intervention group who participated in a THP program and the participants in the waitlist control group (RCT analysis), would demonstrate the THP program effects to be significant and positive in the short-term. As similar post-test data was collected from the control group when they subsequently participated in a THP program, Research Question 2.6 asked whether within-subject mean comparisons of the extended baseline data (T1-T3) from the control group with that group's immediate post-test data following participation in a THP program (T4), would replicate the results from the RCT analysis. Accordingly, references in this section to waitlist control group participants are to these participants when they took part in a THP program and not in their capacity as control participants in the separate RCT analysis.

First, the extended baseline data was considered to test for stability of the measures and the control group data during this extended period prior to participation in a THP program. Of the 615 extended baseline mean comparisons of the control group data (T1-T2, T2-T3, and T1-T3) for the THP programs individually and the Adventure Programs (taken together), 572 demonstrated no significant difference (93%), suggesting that a stable baseline had been established prior to the intervention. The results from all of these mean comparisons have been included in Appendix U. As the extended baseline scores were found to be relatively stable, the pre-post analysis took an average of the extended baseline scores to represent the pre-test scores for this analysis.

All of the within-subjects pre-post analysis results are presented in Table 6.15. A significant result in the positive direction indicates that at immediate post-test the factor scores in the particular outcome variable for control group participants in the relevant THP program were significantly higher when compared to their scores in that outcome variable at pre-test. Alternatively, a significant result in the negative direction suggests that at immediate post-test there were

significantly lower scores in the particular outcome variable for control group participants in the relevant THP program when compared to their pre-test scores. A non-significant result implies that at post-test, scores on the outcome variable were not significantly different for the relevant control group participants from their pre-test scores. Separate comparisons of these pre-post results were also made between the Adventure Programs (taken together) and the Coaching Only Program, and in this case a significant positive result indicates significantly higher scores for control group participants in the Adventure Programs when compared to the Coaching Only Program, and a significant negative result indicates the reverse. Figure 6.6 illustrates the point estimates of the short-term effect sizes and 95% confidence intervals for each outcome variable across each program or group of programs. While the scales in both the table and figure are grouped by scale relevance, mean effects are not presented as the analysis was undertaken on a scale-by-scale basis.

Table 6.15

Within-Subjects Effect Sizes at T4 (from Average Extended Baseline) for Waitlist Control Group as Program Participants

Scale	Waitlist Control Group Within-Subjects Pre-Post Program Effects: ES (SE)					
	Adventure	Arctos	James Craig	Outward Bound	Coaching Only	Adv vs. CO
<i>High Relevance</i>						
Agency	.207 (.064)***	.376 (.105)***	.194 (.082)*	.051 (.104)	.567 (.251)*	-.360 (.259)
Pathways Thinking	.114 (.070)	.256 (.106)*	.111 (.116)	-.025 (.121)	.627 (.188)***	-.513 (.197)**
Goal Self-Regulation	.056 (.074)	.119 (.105)	-.032 (.130)	.080 (.098)	.150 (.183)	-.095 (.189)
General Self-Esteem/SC	.199 (.111)	.328 (.169)	.297 (.190)	-.026 (.136)	.199 (.200)	.001 (.216)
Self-Confidence	.184 (.091)*	.342 (.132)**	.234 (.130)	-.022 (.139)	.284 (.240)	-.100 (.243)
Self-Efficacy	.198 (.088)*	.211 (.134)	.213 (.120)	.169 (.129)	.574 (.242)*	-.376 (.245)
Open Thinking	.136 (.081)	.271 (.117)*	.162 (.097)	-.025 (.132)	.314 (.340)	-.178 (.342)
Cooperative Teamwork	.192 (.098)*	.182 (.147)	.178 (.155)	.216 (.179)	.476 (.207)*	-.284 (.219)
Active Involvement	.123 (.090)	.238 (.146)	.220 (.148)	-.090 (.129)	.206 (.239)	-.083 (.248)
<i>Moderate Relevance</i>						
Optimism	.071 (.074)	.261 (.122)*	.101 (.098)	-.150 (.122)	.022 (.272)	.048 (.285)
Emotion Self-Regulation	.302 (.083)***	.278 (.117)*	.351 (.139)*	.276 (.133)*	.083 (.169)	.218 (.177)
Booster Behaviours	.046 (.085)	.164 (.107)	-.034 (.134)	.008 (.171)	-.496 (.211)*	.542 (.223)*
Booster Thoughts	.028 (.082)	.239 (.123)	-.152 (.123)	-.003 (.124)	-.302 (.295)	.331 (.304)
Wellbeing	.179 (.099)	.207 (.165)	.287 (.121)*	.041 (.153)	.380 (.290)	-.201 (.296)
Same-Sex Relationships SC	.124 (.070)	.206 (.117)	.165 (.116)	.000 (.145)	.042 (.299)	.082 (.304)
Emotional Stability SC	.031 (.097)	.069 (.147)	-.001 (.166)	.023 (.166)	.040 (.195)	-.009 (.200)
Stress Management	.225 (.090)*	.325 (.129)*	.132 (.104)	.218 (.141)	.677 (.139)***	-.452 (.139)***
Social Effectiveness	.113 (.085)	.267 (.098)**	.360 (.141)*	-.287 (.202)	.436 (.177)*	-.322 (.194)
Leadership Ability	.161 (.060)**	.181 (.065)**	.331 (.089)***	-.028 (.137)	.097 (.155)	.065 (.164)
Coping with Change	.201 (.102)*	.291 (.175)	.174 (.122)	.137 (.157)	.585 (.190)**	-.385 (.201)
Internal Locus of Control	.131 (.090)	.328 (.128)*	.155 (.094)	-.090 (.154)	.312 (.271)	-.181 (.272)

Note. T4 = immediate post-test; ES = standardized effect sizes; SE = standard error. Adventure/Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; James Craig = James Craig Adventure Program; Outward Bound = Outward Bound Adventure Program; Coaching Only/CO = Coaching Only Program; SC = Self-Concept. High and Moderate Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. Significant effects are bold for ease of reference. * p < .05; ** p < .01; *** p < .001.

(continues)

Table 6.15 (continued)

Within-Subjects Effect Sizes at T4 (from Average Extended Baseline) for Waitlist Control Group as Program Participants

Scale	Waitlist Control Group Within-Subjects Pre-Post Program Effects: ES (SE)					
	Adventure	Arctos	James Craig	Outward Bound	Coaching Only	Adv vs. CO
Pessimism ^a	.094 (.088)	.199 (.149)	.241 (.126)	-.157 (.155)	-.022 (.321)	.116 (.330)
Focus	.055 (.084)	.173 (.130)	.033 (.155)	-.040 (.102)	-.006 (.155)	.061 (.164)
Consistency of Interest	.028 (.072)	.012 (.124)	-.048 (.153)	.121 (.101)	.202 (.304)	-.174 (.312)
Perseverance of Effort	.086 (.066)	.042 (.103)	.107 (.117)	.109 (.113)	.039 (.163)	.047 (.175)
Hampering ^a	.115 (.064)	.188 (.083)*	.011 (.102)	.147 (.131)	-.385 (.199)	.500 (.204)*
Life Resilience	.230 (.090)*	.396 (.118)***	.125 (.127)	.168 (.165)	.706 (.203)***	-.476 (.216)*
Academic Resilience	.148 (.088)	.233 (.111)*	.147 (.168)	.066 (.132)	.423 (.225)	-.275 (.235)
Life Satisfaction	.111 (.098)	.202 (.159)	.081 (.123)	.050 (.183)	-.199 (.216)	.310 (.236)
Physical Abilities SC	.097 (.066)	.064 (.098)	.100 (.121)	.129 (.101)	.107 (.112)	-.010 (.122)
Physical Appearance SC	.141 (.112)	.334 (.154)*	-.061 (.177)	.149 (.176)	.415 (.166)*	-.274 (.172)
Opposite-Sex Relationships SC	.218 (.065)***	.368 (.095)***	.117 (.127)	.170 (.105)	.192 (.114)	.026 (.132)
Parent Relationships SC	-.008 (.070)	-.028 (.129)	.180 (.116)	-.177 (.116)	-.060 (.201)	.052 (.208)
Honesty-Trustworthiness SC	.166 (.051)***	.167 (.082)*	.227 (.085)**	.103 (.082)	-.255 (.175)	.421 (.190)*
Math SC	.050 (.076)	.021 (.103)	.161 (.135)	-.032 (.125)	.078 (.115)	-.028 (.132)
Verbal SC	.159 (.096)	.219 (.106)*	.044 (.126)	.215 (.227)	-.022 (.217)	.181 (.237)
School SC	.109 (.083)	.185 (.143)	.212 (.097)*	-.071 (.116)	.088 (.218)	.020 (.221)
Time Efficiency	.165 (.085)	.224 (.129)	.112 (.142)	.158 (.119)	.112 (.334)	.053 (.336)
Quality Seeking	.106 (.092)	.277 (.160)	.076 (.098)	-.035 (.147)	.141 (.232)	-.035 (.239)
Overall Effectiveness	.172 (.079)*	.376 (.122)**	.118 (.137)	.021 (.119)	.334 (.235)	-.163 (.242)
External Locus of Control ^a	.181 (.063)**	.319 (.111)**	.264 (.109)*	-.040 (.112)	-.109 (.285)	.290 (.290)

Note. T4 = immediate post-test; ES = standardized effect sizes; SE = standard error; Adventure/Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; James Craig = James Craig Adventure Program; Outward Bound = Outward Bound Adventure Program; Coaching Only/CO = Coaching Only Program; SC = Self-Concept. Low Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. Significant effects are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

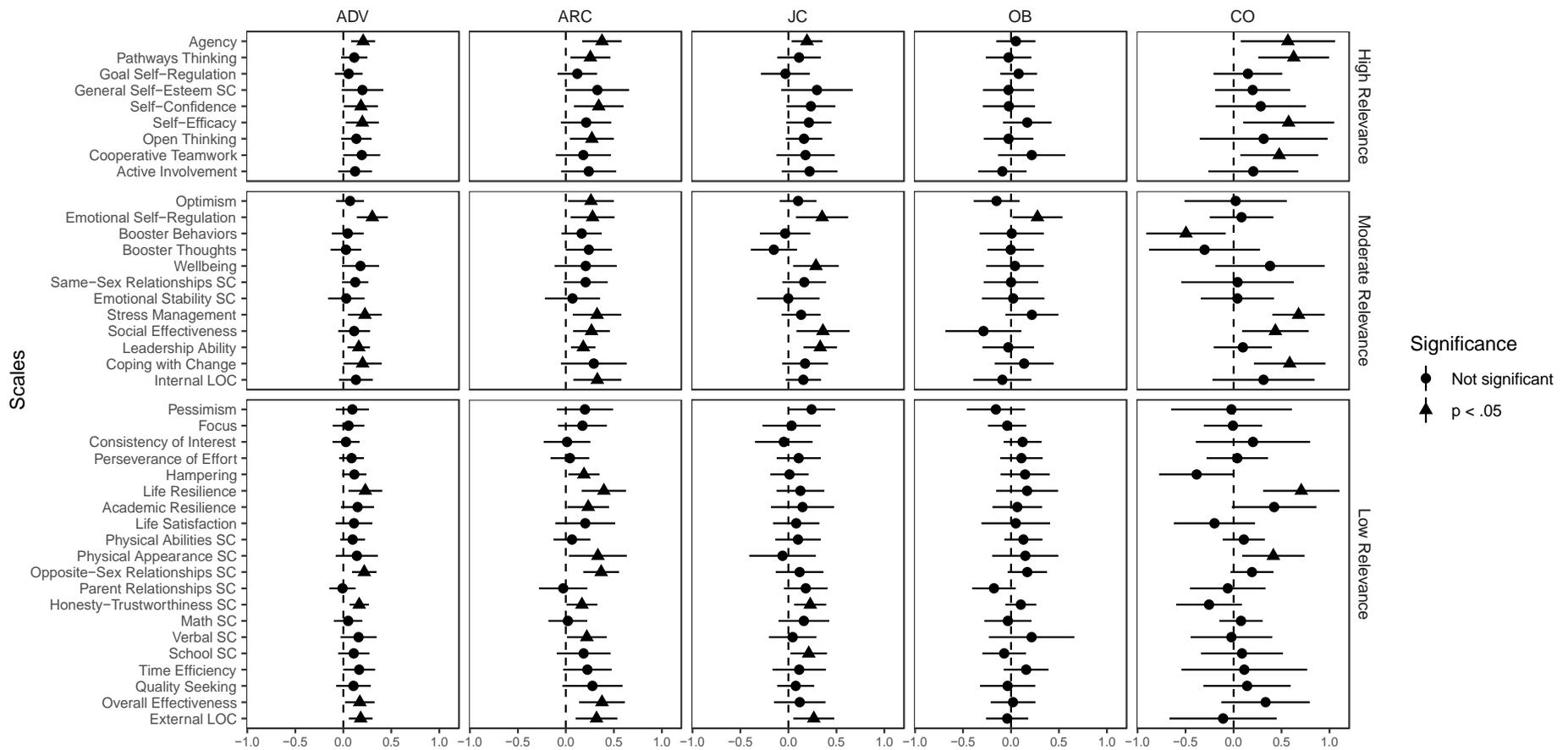


Figure 6.6. THP program effect sizes with 95% confidence intervals for control group (as program participants) within-subjects pre-post analysis (Average T1, T2, T3 to T4) with scales grouped by relevance.

Note. T1, T2, T3 = extended baseline; T4 = immediate post-test; ADV = Adventure Programs (taken together); ARC = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; SC = Self-Concept; LOC = Locus of Control. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. Effect sizes are standardised.

Considering the Adventure Programs (taken together), 13 scales evidenced significant positive pre-post effects. The effect sizes for significant effects were small, ranging from $ES = .161$ (Leadership Ability, a moderate relevance scale) to $ES = .302$ (Emotion Self-Regulation, a moderate relevance scale), with an average effect size for all significant pre-post effects in the Adventure Programs equal to $.203$. Of these 13 scales with significant effects, four of the nine high relevance scales were significant, four of the 12 moderate relevance scales were significant, and five of the 20 low relevance scales were significant. This follows a similar pattern to the results for the RCT analysis, however, there is only some overlap in the individual scales that demonstrated significant effects. The three scales with significant positive effects across both analyses are Agency, Self-Confidence, and Self-Efficacy, with each of these scales being high relevance scales.

Waitlist control group participants in the Arctos Adventure Program reported significantly higher post-test scores on 19 of the scales when compared to their pre-test scores, with small-to-moderate effect sizes ranging from $ES = .167$ (Honesty-Trustworthiness Self-Concept, a low relevance scale) to $ES = .396$ (Life Resilience, a low relevance scale), with an average effect size for all significant pre-post effects for the Arctos control group participants equal to $.289$. Of these 19 scales with significant effects, four of the nine high relevance scales were significant, six of the 12 moderate relevance scales were significant, and nine of the 20 low relevance scales were significant. These results are different to the results of the RCT analysis in which only a single scale (Physical Abilities Self-Concept) evidenced significant positive effects for the Arctos intervention participants. The James Craig Adventure Program control group participants demonstrated eight scales with significantly higher post-test scores when compared to their pre-test scores, with small-to-moderate effect sizes ranging from $ES = .194$ (Agency, a high relevance scale) to $ES = .360$ (Social Effectiveness, a moderate relevance scale), with an average effect size for all significant positive effects for the James Craig control group participants equal to $.278$. These eight scales consisted of one high relevance scale, four moderate relevance scales, and three low relevance scales. These results also are different to the RCT analysis in which 24 scales evidenced significant positive effects. However, five of these eight scales demonstrated significant positive effects

in both the within-subjects and RCT analyses: Agency, Emotion Self-Regulation, Wellbeing, Social Effectiveness, and Leadership Ability. The control group participants in the Outward Bound Adventure Program only demonstrated a single scale with significantly higher post-test scores when compared to their pre-test scores (Emotion Self-Regulation: $ES = .276$), similar to the results in the RCT analysis in which the intervention participants in that program demonstrated only a single scale (Agency) with significant positive effects. The waitlist control group participants in the Coaching Only Program demonstrated significantly higher post-test scores on nine of the scales when compared to their pre-test scores, with moderate-to-strong effect sizes ranging from $ES = .415$ (Physical Appearance Self-Concept, a low relevance scale) to $ES = .706$ (Life Resilience, a low relevance scale), with an average effect size for all significant positive effects for the Coaching Only participants equal to $.563$. These nine scales consisted of four high relevance scales, three moderate relevance scales, and two low relevance scales. These results also diverge from the RCT analysis in which two scales evidenced significant positive effects and three scales demonstrated significant negative effects. However, the Social Effectiveness scale demonstrated significant positive effects in both analyses. The control group participants in the Coaching Only Program also demonstrated significantly lower post-test scores when compared to their pre-test scores in one scale (Booster Behaviours) with $ES = -.496$, although this is different to the three scales that evidenced significant negative effects in the RCT analysis.

Comparing the scores of control group participants in the Adventure Programs with the Coaching Only Program, there were three scales in which the participants in the Adventure Programs reported significantly higher post-test scores when compared to the participants in the Coaching Only Program (Booster Behaviours, Hampering, and Honesty-Trustworthiness Self-Concept), and three scales in which the opposite effect was seen (Pathways Thinking, Life Resilience, and Stress Management). These results also diverge from the results of the RCT analysis, with no overlap between the two sets of analyses and contradictory results for the Pathways Thinking scale, which showed greater effects for the Adventure Program intervention participants in the RCT analysis and greater effects for the Coaching Only Program control group participants in the within-subjects analysis.

Moreover, the Coaching Only Program did not demonstrate any significantly greater short-term effects when compared to the Adventure Programs in the results of the RCT analysis, but it did do so in the within-subjects analysis for one high, one moderate, and one low relevance scale.

Correlating the pre-post effect sizes with scale relevance demonstrated a small correlation for the Adventure Programs (taken together) ($r = .21$), as well as for the James Craig ($r = .23$) and Arctos programs ($r = .23$). The Coaching Only program evidenced a moderate correlation ($r = .38$), while the Outward Bound program reflected almost no correlation between effect sizes and scale relevance.

Similar to the RCT analysis, control group participants in the Adventure Programs (taken together) reported significant positive outcomes in many important qualities and skills related to flourishing, including hope, self-confidence, self-efficacy, resilience, and aspects of self-regulation, self-concept, and life effectiveness. However, most of the specific scales demonstrating significant effects differed from the significant effects found in the RCT analysis. Moreover, the scales with significant effects in the within-subjects analysis ranged across the high, moderate and low relevance scales, unlike the RCT analysis in which the high relevance scales were found to dominate. Also similar to the RCT analysis, the different THP programs individually evidenced different levels of effectiveness in the within-subjects analysis. However, while the Arctos control group participants reported the greatest number of significant effects in the within-subjects analysis, the James Craig intervention participants reported the greatest number of significant effects in the RCT analysis. The Coaching Only Program control group participants also reported many more significant effects in the within-subjects analysis than the intervention Coaching Only Program participants reported in the RCT analysis. The Outward Bound program, however, did not evidence many significant effects in either of the analyses. Finally, a comparison of the Adventure Programs with the Coaching Only Program found the two programs to be more evenly matched in the within-subjects analysis than it did in the RCT analysis.

Summary

This study aimed to test the effectiveness of a number of novel OAE programs offered to students at schools in areas of socioeconomic disadvantage. To meet this aim, an RCT design was used and Survey data for the THP program participants and the waitlist control group were analysed and compared. Multiple factor score regression analysis indicated many significant positive short-term effects for the participants in the James Craig program, but almost no significant effects for the participants in the Arctos and Outward Bound programs. However, when analysed together, the participants in the Adventure Programs demonstrated a number of significant positive short-term effects. On the other hand, some significant negative short-term effects were found for the participants in the Coaching Only Program. In general, the high relevance scales demonstrated the greatest effects, particularly when considering the Adventure Programs (taken together).

Participants in the Adventure Programs (taken together), as well as the James Craig and Outward Bound programs individually, demonstrated some significant positive long-term effects, but there were no significant positive long-term effects for participants in the Arctos program. Participants in the Coaching Only Program evidenced three significant positive long-term effects and one significant negative long-term effect. In general, the effects during the follow-up period were relatively stable with fewer significant effects. Interestingly, there were a handful of sleeper effects between T2 and T3, with Opposite-Sex Relationships Self-Concept showing significant positive effects from T2 to T3 for participants in the James Craig and Coaching Only Program, as well as the Adventure Programs (taken together). There was also a sleeper effect for Optimism and participants in the Coaching Only Program.

Aptitude-treatment interactions were also considered, and these analyses evidenced some significant effects. Further investigation of these effects demonstrated several outcome variables for which low-aptitude participants reported significantly higher post-test scores than their control group counterparts, with these results being better than for the participants at other aptitude levels.

However, in some cases the effects for higher-aptitude participants on these outcome variables revealed significantly lower post-test scores than their relevant control group. There were also results for which low-aptitude participants fared worse than their high-aptitude counterparts, in some cases also demonstrating significantly lower post-test scores than their control group. Overall, however, there were a number of results in the hypothesised direction favouring low-aptitude participants, and some of these results were also in respect of outcome variables that did not otherwise demonstrate a significant positive group main effect in the initial analyses.

The effects of the Adventure Programs were also compared against the effects of the Coaching Only Program in order to evaluate the incremental benefits of the OAE element for the THP program. There were a number of significant short-term differences which favoured the Adventure Programs, particularly in relation to the high relevance scales. However, there were only three significant differences in effects between the programs over the longer term and one of these favoured the Coaching Only Program.

Data from the waitlist control group was also used to conduct a within-subjects analysis in order to test the replicability of the RCT results. Multiple-group models were used to compare post-test scores for the control group with an average of their extended baseline data. This analysis also demonstrated significant positive short-term effects for the control group participants in the Adventure Programs (taken together), as well as some of the THP programs considered separately. However, the Arctos program participants reported many more significant results in this analysis than the other THP programs. This result diverges from the RCT results, which found the James Craig program participants to have the most significant results. Moreover, any significant differences found between the Adventure Programs and Coaching Only Program in the within-subjects analysis were evenly divided between the two programs, and two of the scales had opposite results in the two sets of analyses. Notwithstanding that the results may not make a case for replicability, they do provide interesting distinctions that will be considered further in the discussion section below and the qualitative study, as well as Chapter Eight.

Discussion

Strengths

This study extends prior research on OAE by assessing a wide range of outcomes and multiple modes of adventure, as well as the incremental value of the outdoor adventure component and additional benefits of the inclusion of skilled coaching. Detailed descriptions of the various Adventure Program modalities have been provided, as well as a comprehensive comparison of the Adventure Programs with the Coaching Only Program. These details provide an opportunity to better understand what program aspects are most effective and for what outcomes. Of the 41 outcome variables, the participants in the Adventure Programs (taken together) demonstrated significant small-to-moderate short-term positive effects in 12 different areas of personal and social development: Agency, Pathways Thinking, Self-Confidence, Self-Efficacy, Open Thinking, Active Involvement, Wellbeing, Social Effectiveness, Focus, reduced Hampering, Verbal Self-Concept, and Quality Seeking. While none of these significant positive effects were found to maintain over the longer term, new significant effects were found for Cooperative Teamwork, Booster Behaviours, and Opposite-Sex Relationships Self-Concept. Sleeper effects were also found for Opposite-Sex Relationships Self-Concept across multiple programs, as well as for Optimism in the Coaching Only Program.

The short-term effect sizes for the Adventure Program effects were smaller than found in previous research (see, e.g., average ES of .35 reported across a range of meta-analytic studies by Neill, 2008). However, the methodology used in that research is variable, and research with an RCT design is known to result in smaller-sized effects (Cheung & Slavin, 2016). These results are also impacted by the diverse range of effects found for the individual THP program modalities. While the James Craig program demonstrated 24 significant positive short-term effects with moderate-to-large effect sizes, the other Adventure Programs demonstrated only one or two significant effects. However, this result was not replicated in the within-subjects analysis on the control group participants, where the Arctos program demonstrated more than twice as many significant effects in comparison to the other programs. In an effort to understand whether the difference in results

between the RCT and within-subjects analyses was a function of the distinct methods of analysis or the different program groups, a pre-post analysis of the intervention group data was undertaken. These results are included in Appendix V. This analysis was reasonably consistent with the RCT analysis, suggesting that it was more likely to be program group differences leading to the inconsistent results. Further consideration of these group differences is reserved for the discussion section in Chapter Eight.

The findings from this study also provide a direct comparison between alternative extracurricular school-based coaching programs (one with adventure experiences and one without). While prior meta-analytic research has been used to compare competing educational programs (e.g., Hattie et al., 1997; Neill, 2008), this appears to be the first study to compare outcomes between these different types of programs within a single research study. While there were a number of significant short-term differences which favoured the Adventure Programs, there were only three significant differences in effects between the programs over the longer term and one of these favoured the Coaching Only Program. Moreover, the within-subjects analysis demonstrated different results, with each program evidencing three outcomes with significant differences in effects which favoured that program. Similar to the diversity of results for the Adventure Program modalities, an open question remains as to the group differences impacting program outcomes and whether the adventure component is an important element influencing these results (refer to Chapter Eight for further discussion).

A further contribution of this study relates to the extensive aptitude-treatment interaction analysis, providing additional data on the effectiveness of the THP programs for those participants with the lowest baseline levels in the outcome variables of interest and, therefore, the greatest needs. For some of the significant aptitude-treatment interaction effects that favoured the more disadvantaged participants in terms of an outcome, there was no significant main effect for program participation. Accordingly, these significant positive effects may have been overlooked without the interaction analyses. Being a program developed particularly for disadvantaged students, these results can assist in developing a

better understanding of which program elements and focused outcomes provide the greatest impact for disadvantaged adolescents.

Finally, a primary strength of this study is found in the research design and statistical methods used, which meet existing criticisms in the OAE and coaching literature (Cason & Gillis, 1994; Ewert & McAvoy, 2000; Hattie et al., 1997; Neill, 2008; Scrutton & Beames, 2015; Sibthorp, 2000). The research design includes a randomised controlled trial, providing a stronger test of potential program effects when compared with the more usual pre-post within-subjects analysis. It should be noted that because the adventure experiences took place overnight, schools required the program groups to be single gender. Consequently, it was not possible to systematically assign participants to comparison groups. However, it is often difficult in school-based studies to achieve perfect randomisation. Despite these challenges, participation involved multiple schools and multiple cohorts, increasing sample size and opportunities for generalizability. Moreover, many well-established outcome measures were used, which withstood rigorous psychometric analysis, and factor scores were applied, thereby minimising measurement error and improving the validity and reliability of the analyses undertaken in this study. Furthermore, the research design incorporated two assessment waves of post-test data, allowing for longer-term and follow-up analyses of program effects in addition to the more commonly assessed short-term effects. An additional test of replicability was also conducted using a within-subjects extended baseline design, which provided an opportunity to assess the stability of the outcomes over the baseline period. Finally, the analysis involved advanced statistical procedures, including multiple regression analysis and more sophisticated methods for handling missing data and clustering, as well as controlling for potentially confounding variables. The measurement scales were grouped by program relevance to aid in the interpretation of the results, and standardised effect sizes also were reported for easier interpretation and comparison across research studies.

Limitations

Some limitations should be considered when interpreting the results. First, the 362 participants in this study are from a specific area of socioeconomic

disadvantage in Sydney, Australia. Therefore, it may be difficult to generalise the results to other populations. Moreover, while the multiple program modes were a strength of this study, they were also a limitation. The variety of programs reduced the effective sample size and added additional elements that could not be controlled, impacting statistical power and complicating the analyses and interpretations of the results. Measurement attrition further affected the results, particularly for the within-subject analysis.

Second, only students who completed the program application form and provided written consent were eligible to participate in a THP program and the research, either as an intervention or control group member. Consequently, this research may have excluded students who were most likely to benefit from the program, such as more disadvantaged students lacking the parental support or the confidence and skills necessary to meet the application requirements. This issue is particularly important given the aptitude-treatment interaction results evidenced stronger results for lower aptitude participants for some outcomes. Such consent bias may also mean that the study participants are not representative of all possible participants, thereby impacting the generalizability of the results.

Third, while the breadth of the outcomes included in the research are a strength of this study, they also are a limitation. Being so comprehensive resulted in a measurement instrument with over 200 items. The length of the Survey, coupled with the repeated measures design, may have caused participants to complete the Survey with less than their full attention and consideration (Davidson, Ewert, & Chang, 2016). Moreover, some of the Survey items may have been too complex or ambiguous for the participants, particularly at their developmental level, and other items have an 'all or nothing' context which do not seek incremental change of the type anticipated. Furthermore, some of the scales consisted entirely of negatively-worded items, which can be problematic, particularly for adolescents (Marsh, 1986a; Melnick & Gable, 1990). Of the five scales that comprised all negatively-worded items (Pessimism, Consistency of Interest, Hampering, Emotional Stability Self-Concept, and External Locus of Control), only Hampering demonstrated any significant positive effects in the RCT analysis and two showed significant negative effects (Consistency of Interest and External Locus of Control). Future studies

should be mindful of the amount of time and mental energy they are asking adolescent participants to expend in completing a survey instrument.

Fourth, having so many outcome measures also led to complex statistical models with a large number of parameters when compared to the sample size. Moreover, analysing and presenting the results was complicated by the sheer number of scales to consider. While we could have dropped some of the scales, transparency and knowledge accumulation were preferred in this thesis (see Wasserstein & Lazar, 2016; Wasserstein, Schirm, & Lazar, 2019). The complexity also could have been addressed by using exploratory factor analysis to consolidate the measurement items into fewer scales, however, the intention for this research was to evaluate the THP program with scales commonly used to assess OAE programs. The extent of the analyses in this thesis increases the risk of reporting false positive results. Whether and how to control for this increased risk has been the subject of much debate (see, e.g., Bender & Lange, 2001; Benjamini & Hochberg, 1995; Cabin & Mitchell, 2000; Gelman, Hill, & Yajima, 2012; Glickman, Rao, & Schultz, 2014; Nakagawa, 2004; Noble, 2009; O' Keefe, 2003; Perneger, 1998; Rothman, 1990; Schulz & Grimes, 2005; Veazie, 2006). An alpha adjustment was not made in the context of determining statistical significance given the potential reduction in statistical power to detect significant effects (Gelman et al., 2012; Glickman et al., 2014; Nakagawa, 2004).

As a way to deal with the complexity, the outcome measures were reviewed and rated for their relevance, with an eye to the ultimate design of the Adventure Programs and the explicit aims of the coaching and adventure experiences developed as part of those programs. By partitioning the outcomes to focus on those with the most relevance, the effective number of tests were reduced. In supplemental analyses, multivariate omnibus tests were conducted on the high relevance scales (being the primary focus of this study) as a control for false positive results. These test results are presented in Appendix W. The statistically significant Wald test results suggest tests of intervention effects for the individual outcome variables are appropriate. Nonetheless, future research should keep the issue of multiple tests in mind when selecting outcome variables. Particularly if selection of outcome measures can occur after program design and aims are firmly established,

then the measures selected can be more tightly aligned with the specific design and aims of the program and therefore, can be more limited. A more limited set of outcome variables, together with a larger sample size, will help to differentiate between statistical and practical significance.

Fifth, while OAE research has found the overall effects of OAE programs to be similar for male and female participants in single-gender groups (Hattie et al., 1997; cf., Ewert & McAvoy, 2000; Norton & Watt, 2014; Sibthorp, et al., 2007), given the gender imbalance across the different THP programs, there is a question as to whether the THP program effects might vary as a consequence of gender. For this purpose, gender-treatment interaction effects were tested in a similar way to aptitude-treatment interaction effects. The detailed results for the significant effects are presented in Appendix X. For the Adventure Programs taken together, of the 41 outcome variables and 82 possible interaction effects in the Short-Term and Long-Term analyses, there was only one significant interaction effect in the Long-Term Analysis (Parent Relationships Self-Concept), which evidenced a significant positive simple main effect for males ($ES = .262$, $SE = .110$, $p < .05$) and no significant effect for females. For the THP programs considered individually (noting the small numbers for many of the gender groups, including only six male intervention participants in the Arctos Adventure Program), of the 164 possible interactions at each wave, there were 46 significant gender-treatment interactions in the Short-Term Analysis (28.05%) and 37 significant gender-treatment interactions in the Long-Term Analysis (22.56%). Half of the significant gender-treatment interactions reflected significant positive simple main effects for females, with many of these occurring in the James Craig and Coaching Only programs. On the other hand, nearly 40% of the significant gender-treatment interactions reflected significant negative simple main effects for males, primarily for the Arctos and Coaching Only programs. Considering the Adventure Programs (taken together), there is little evidence that gender had an effect on program outcomes. Notwithstanding these overall results, there appears to be some evidence that females gained more than males in some outcome variables and some of the THP programs, while males may have declined more than females in some outcome variables and some of the THP programs. However, given the small sample sizes for each program grouped by

gender, it is suggested that further consideration should be given in the future to the interaction between gender and OAE program effects.

Sixth, while the within-subjects analysis provides an alternative assessment of the effectiveness of the THP programs and replicability of the RCT analysis, the outcomes of the THP programs on the control group participants may have been impacted by their interaction with the intervention group participants at their school, who participated in a THP program just prior to control group participation (generally, within the same school year). Additionally, most schools ran THP programs over multiple years. Consequently, other intervention and control group participants also may have been impacted by the stories they heard from or about prior participants in a THP program. Finally, while an attempt was made to conduct Adventure Programs and Coaching Only Programs in different schools, one school offered both an Adventure Program and a Coaching Only Program. Although these programs were implemented some years apart, it is possible that knowledge of an alternative adventure program experience had an effect on participant perception of the Coaching Only Program experience.

Finally, the data collection process required coordination with the schools and sometimes the timing of collection was close to school breaks. Therefore, some of the data was collected later or earlier than planned, and such timing differences may have impacted the results. Data collection from an intervention group and its matching control group, however, occurred at the same time. There is also the timing of pre- and post-test assessment in which there is potential for the emotions experienced by participants in the period immediately prior to or following an intervention to distort their scores (Allison, 2000; Ewert & Sibthorp, 2009; Hattie et al., 1997; Marsh et al., 1986b). However, the timing of the pre-test that occurred generally at the time a program commenced, did not correspond to the adventure experience (the first of which was at least four weeks into the program). Moreover, the immediate post-test was administered about a week after program completion, which was at least 5 weeks after completion of the final adventure experience. In each case it is suggested that the gap between the adventure components of the program and Survey assessment reduce the possibility for such feelings to bias the Survey scores. Furthermore, a comparison of pre-test data for the intervention and

control groups revealed very few pre-treatment group differences and any differences were subsequently controlled for in the analysis using statistical methods. The three waves of extended baseline data for the control group were also assessed, which did not evidence any pre-test bias and further confirmed the stability of the measurement scales. Nonetheless, any feelings associated with post-course adjustment and the experience of involvement in a Community Project may have impacted the post-test results. Further consideration will be given to these variables in Chapter Eight.

Conclusion

The aim of Study 2 was to assess the quantitative effects of the novel THP program on a number of positive outcomes for its disadvantaged, adolescent participants. Overall, the RCT analysis revealed a number of significant positive short- and long-term effects for participants in the Adventure Programs (taken together) and some of the individual THP programs, when compared to the control group. Moreover, a number of outcome variables were found to have significant positive effects for those participants with lower baseline levels of the relevant outcome variable. The within-subjects analysis also evidenced some significant positive effects, although these results were not as consistent with the short-term RCT analysis as had been expected. In Chapter Eight, the results from this study will be juxtaposed with the qualitative analysis the subject of Study 3 (see Chapter Seven), together with a general discussion and concluding thoughts on all of the results from the research for this thesis.

CHAPTER SEVEN

STUDY 3: QUALITATIVE EXAMINATION OF PARTICIPANT PROGRAM EXPERIENCES THROUGH THE LENS OF CONSTRUCTIVE-DEVELOPMENTAL THEORY AND INTERPRETATIVE PHENOMENOLOGICAL ANALYSIS

Introduction

Outdoor adventure education (OAE) provides healthy and challenging environments for exploring self-awareness and broader perspective-taking, and it is this type of environment that can provide the impetus for transformational growth (see e.g., Carpenter & Harper, 2016; Ewert & Sibthorp, 2009; Hattie et al., 1997; Martin & Leberman, 2005; McKenzie, 2000; Neill & Dias, 2001; Sheard & Golby, 2006; Sibthorp & Jostad, 2014). Research has found that engaging in challenging tasks is beneficial not only for learning, but more generally for a person's wellbeing (Csikszentmihalyi, 1992). The level of challenge, however, is a balancing act. If a task is too simple, it can lead to boredom (Csikszentmihalyi, 1992; Nakamura & Csikszentmihalyi, 2014); activities that offer no challenge are unlikely to be motivating or to bolster one's sense of efficacy upon completion (Crescioni & Baumeister, 2013). On the other hand, an overly challenging task can create anxiety (Csikszentmihalyi, 1992; Nakamura & Csikszentmihalyi, 2014), and tasks that are so difficult as to assure failure may have a negative impact on one's self-efficacy and self-esteem (Crescioni & Baumeister, 2013). Optimal experience results from the opportunity to engage in challenging activities together with possession of the capacity (through some combination of skill and support) to meet those challenges. Importantly, it is the subjective perception of the challenge and one's capacity to meet the challenge, rather than an objective one, that determines the quality of these experiences (Nakamura & Csikszentmihalyi, 2014). As a consequence, a person's capacity to make meaning of challenging experiences and their self-efficacy will influence the outcomes of OAE programs.

The Helmsman Project (THP) program integrates a series of structured developmental coaching sessions with outdoor adventure experiences for disadvantaged adolescents that aims to increase high school engagement and improve educational attainment in its participants by building hope, self-regulation, resilience, and other life effectiveness skills. At the heart of these programs is David A. Kolb's (1984, 2014) experiential learning theory. According to this theory, effective learning requires not only being immersed in a concrete experience, but also being able to observe and reflect on that experience, and then being able to analyse those reflections into abstract concepts and generalisations that are used to inform future experiences (David A. Kolb, 1984, 2014).

As a coach on one of the THP programs, I noticed that individual participants related differently to the coaching sessions, the content of the program, and the experiential learning model that framed the program. Some participants struggled with the requirements for reflective observation and abstract conceptualisation, at times leading to frustration and withdrawal from the experiences. Building self-regulation strategies was also challenging for some participants, as it required them to take ownership of their thoughts, feelings, and behaviours; to understand themselves as separate from others (Kegan, 1994).

Having studied Kegan's (1982, 1994) constructive-developmental theory as a master's student of coaching psychology, I was interested to consider the experiences of THP program participants through the lens of that theory. Constructive-developmental theory describes the qualitatively different ways in which we construct meaning around experience. For Kegan (1982), there is "no feeling, no experience, no thought, no perception, independent of a meaning-making context in which it *becomes* a feeling, an experience, a thought, a perception, because we *are* the meaning-making context" (p. 11). Berger (2002) aptly describes Kegan's theory as being concerned with the *shape* of our understanding, rather than the *content* of our minds. Constructive-developmental

theory is concerned also with the development of meaning-making capacity,²⁹ which is marked by increasing complexity. Overcoming challenging situations with appropriate support can provide the impetus for this development (Kegan, 1994), making OAE an appropriate context for consideration of a constructive-developmental framework. Chapter Two presents an overview of Kegan's constructive-developmental theory, including the five qualitatively different *Orders of Mind*, or developmental stages of meaning-making. Further details on three of these five stages is set out below in "Methodology and Procedures" (see the subsection headed "Constructive-Developmental Lens").

A person's meaning making influences not only their self-concept and self-esteem, but their interactions and relations with others, as well as their interpretations of events and ideas. If participants come to a THP program at different constructive-developmental stages, they may be experiencing the program in disparate ways. For example, their conceptions of success, expectations for their coaches, and understanding of teamwork may be differently conceived based on the ways of knowing from which they are operating. Moreover, the coaches and other program providers may have expectations of how participants will experience the program based on their own ways of knowing. Having a developmental mismatch between program providers and participants can result in ineffective programming decisions.

I was curious whether the design and delivery of the THP programs might expect a level of meaning-making that exceeded the capacities of some of its adolescent participants, resulting in what Kegan (1994) describes as a "mismatch between external epistemological demand and internal epistemological capacity" (p. 41). Much of the philosophy underlying OAE emphasises the need to be at the edge of one's "physical and psychological possibilities" in order to stimulate growth (Neill & Dias, 2001, p. 1). Kegan (1994) refers to this challenging space as the edge of a person's meaning-making capacity; their "growing edge" (p. 53). However, these

²⁹ The terms *meaning-making*, *meaning-construction*, *perspective-taking*, *way of knowing*, *constructive-development* and related terms all refer to the ways in which one constructs meaning of his or her experiences, and these terms are used interchangeably in this thesis.

gaps between demand and capacity must be met with ample support if epistemological growth is to occur. The inclusion of support systems in the THP programs demonstrates a recognition that a challenging environment without the right balance of support is likely to be ineffectual. What is required in the case of a complexity of mind mismatch is an “evolutionary bridge, a context for crossing over” (Kegan, 1994, p. 43). Importantly, an evolutionary support must begin by recognising and accepting the ways in which a person currently makes meaning of his or her experiences. It is only from this anchor at one end that circumstances can then be created which support a crossing out of and beyond that current way of making meaning (Kegan, 1994). Berger (2004) notes that by slowing down and listening for the edges of a person’s understanding, we can honour these transformational spaces and provide more thoughtful and intentional support. Consequently, what is essential to this process of support is uncovering and seeking to understand the ways in which program participants make meaning of experience.

Constructive-developmental theory provides a powerful means for systematically examining how participants are understanding their experiences in the THP programs, and whether there are different ways in which the programs could better support and extend participants’ capacities to create meaning. Developing such an understanding may also help OAE researchers and practitioners appreciate and honour not only the different ways program participants might experience aspects of an OAE program, but how their own meaning-making systems influence their interactions with, and expectations of, these participants. An appreciation of participant meaning making may also provide a deeper understanding of the interaction between OAE programs and their intended developmental outcomes, thereby offering opportunities to enhance those outcomes. Consequently, it is suggested that having an awareness and understanding of constructive-developmental theory can inform the design and implementation of OAE programming and potentially influence program outcomes for participants.

This chapter begins by outlining the broad aims and more specific research questions for this study. As this study follows an inductive approach, there are no a priori predictions for these questions. The specific methodology and procedures

used in this study are then detailed. The results section follows, in which the participant data is analysed in the context of the research questions. These results are then briefly discussed before summarising the study.

Research Aims and Questions

Research Aims

This study aims to examine the structure of the way in which 13 participants understood their experiences of a THP program, as well as developmental differences in these understandings, using interpretative phenomenological analysis (Smith, 1996) and the constructive-developmental lens of Robert Kegan (1982, 1994). A secondary aim is to gather additional qualitatively rich data of the content of participant experiences. Understanding and appreciating these developmental differences and influential moments in participants' experiences can allow the THP program providers to offer participants the most nourishing environment for growth. More specifically, this study aims to understand:

1. the constructive-developmental levels through which participants are making meaning of their experiences;
2. whether there are systematic relationships between a participant's constructive-developmental stage and the participant's understanding of experiences in a THP program;
3. whether there are identifiable changes in meaning making that occurred for participants through their participation in a THP program, and if so, whether there are explicit program elements and critical events that participants experienced as challenging their meaning making and supporting the evolution of a new meaning making; and
4. other themes that may arise in the context of the participants' own account of their experiences.

Statement of Research Questions

The qualitative methodology used in this study (see the section headed "Methodology and Procedures" below), emphasises an inductive stance which allows for unanticipated themes to emerge during analysis (Smith, 2004). Accordingly, this study is exploratory in nature and, therefore, involves open

research questions rather than clear hypotheses. These questions are outlined below and numbered sequentially beginning with the number of this study.

Research Question 3.1: Assessment of constructive-developmental stage. What are participants' constructive-developmental stages at the time of the interview?

Research Question 3.2: Constructive-developmental stage and program experience. Are there parallels in understanding of THP program experiences for participants making meaning at similar constructive-developmental levels and divergences in understanding of those experiences for participants making meaning across different constructive-developmental levels?

Research Question 3.3: Constructive-developmental stage at THP program commencement and growth through program experience. Is there evidence that participants were at a different constructive-developmental stage at the time of participation in a THP program? Is there any evidence that participants experienced constructive-developmental growth stimulated through participation in a THP program? If so, are there explicit program elements and critical events that participants experiences as challenging their meaning making and supporting the evolution of a new meaning making?

Research Question 3.4: Additional themes arising from participants' accounts of their program experiences. Are there other themes that arise out of participants' own accounts of their experiences of the THP program?

Summary

While quantitative research is important for establishing an evidentiary basis for the effectiveness of the THP programs, it is suggested that incorporating a qualitative examination of participant program experience will provide a more holistic perspective of the THP program and its effects. The qualitative method applies interpretative phenomenological analysis and constructive-developmental theory to systematically explore the meaning-making capacities of program participants and how those capacities may interact with key aspects of the THP program to construct participants' experiences of the program and stimulate

development, including growth in their meaning making. This appears to be the first time a developmental coaching intervention has been assessed with a constructive-developmental lens in the context of real-world challenges precipitated through OAE. Developing an awareness and understanding of constructive-developmental theory and the different ways in which individuals make meaning of experience can inform the design and implementation of OAE programming and potentially influence program outcomes for participants. The next section of this chapter outlines the methodology and procedures used in this study, with results presented in the section that follows.

Methodology and Procedures

Participants

My goal was to interview participants across both genders, each THP program mode, and a variety of the schools that participated in the research. I hoped to have enough participants to be able to see multiple constructive-developmental stages and some patterns in the participant experiences. However, I was also aware that the type of interview I wanted to conduct was lengthy and therefore needed to be limited in numbers. Accordingly, the aim was to interview between 10 and 20 past THP program participants. Participating schools were advised of the desire to conduct participant interviews and requested to ask past participants whether they would be willing to volunteer for such an interview. Aside from the length of the interview, nothing about the nature or structure of the interview was disclosed.

The sample for this study consisted of 13 students each of whom had been a participant in a THP program, whether as a member of the intervention or waitlist control group. Table 7.1 provides a summary overview of the various THP programs, including the different adventure education experiences. These programs and associated experiences are described in detail in Chapter Four. The study participants were from five of the 11 high schools that took part in the THP programs for this research. None of the study participants and interviewers were known to each other prior to the interview. Additional information on the gender and THP program mode for these participants is set out in Table 7.2.

Table 7.1
THP Program Overview by Program Mode

	Adventure Programs			Coaching Only Program
	Arctos	James Craig	Outward Bound	
Adventure Type	Small yacht	Tall ship	Hiking	No adventure
Experiences	Sail boat independently & plan/ prepare meals	Assist to sail boat, group tasks, non-sailing activities	Hiking, rock climbing, abseiling, meal prep, survival skills	Presentation skills seminar, individual project
Program Group Size	8-10	8-10	8-10	8-10
Experience Group Size	8-10	40	16	8-10
Experience Inter-School	No	Yes	Yes	No
Framework	Experiential learning cycle			
Support	Developmental coaching			
Other Opportunities	Community Project			

Note. THP = The Helmsman Project. Adventure Programs refer to those THP programs with an adventure component and Coaching Only Program refers to the THP program without an adventure component. Arctos = Arctos Adventure Program; James Craig = James Craig Adventure Program; Outward Bound = Outward Bound Adventure Program.

Table 7.2
Interview Participant Details by THP Program and Gender (N=13)

THP Program	Gender		Total
	Female	Male	
Arctos	5	0	5
James Craig	1	3	4
Outward Bound	0	1	1
Coaching Only	2	1	3
Total	8	5	13

Note. THP = The Helmsman Project; Arctos = Arctos Adventure Program; James Craig = James Craig Adventure Program; Outward Bound = Outward Bound Adventure Program; Coaching Only = Coaching Only Program.

At the time of the interview, participants ranged in age from 14 to 18 years. However, the participants were between the ages of 14 and 15 years when they

participated in a THP program. Consequently, some of the study participants had experienced a THP program as early as three months prior to the interview, while one had experienced a THP program almost three years prior to the interview, and others were somewhere in between. A graph is included in Figure 7.1 indicating at the time of the interview the number of years which had elapsed since the participants completed their THP program. All participants were close enough to the program to be able to recall details of their experience. However, I was curious whether those participants who were a year or two away from their experience of a program, might have a broader perspective of that experience.

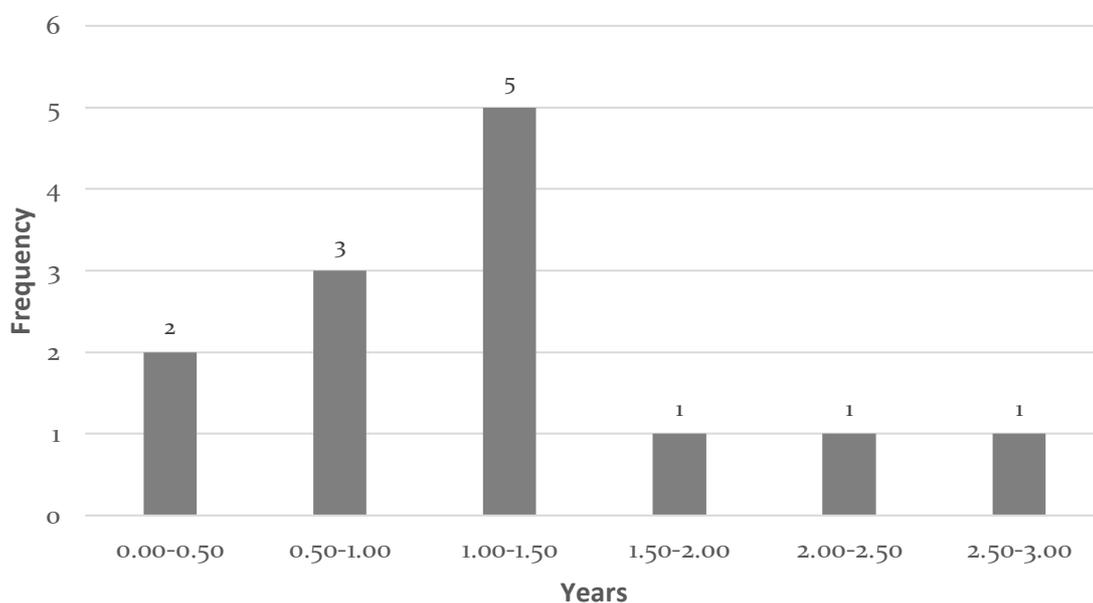


Figure 7.1. Histogram reflecting distribution of interview participants' time (years) from completion of THP program to interview date (N = 13).

Research Design

Spinelli (2005) has argued that in order for psychology to understand human beings, it must begin from the study of lived experience. In OAE research, a specific call has been made for qualitative data collection in order to obtain more detailed information on the relationship among program participants, the various program elements, and the numerous outcome measures (Barrett & Greenaway, 1995; Ewert & McAvoy, 2000; Harper, 2010; Klint, 1990; Martin & Leberman, 2005; McKenzie, 2000; Rowley, 1987). Scholars in coaching psychology have also noted that notwithstanding the importance of quantitative research, qualitative research

provides unique insights and learnings fundamental to developing a deeper understanding of coaching processes and relationships (Grant, 2016a). Given the complex and subjective nature of participants' experiences in both OAE and coaching, it is suggested that a mixed-method research approach can provide a more nuanced and complete picture of the effects of the THP programs.

The overarching qualitative approach applied in this study is that of interpretative phenomenological analysis (IPA; Smith, 1996). Although IPA is a relatively new methodology in qualitative analysis, it has been widely applied across a range of sub-disciplines within psychology, including applied psychology (Reid, Flowers, & Larkin, 2005). IPA is closely aligned with the original conception of cognitive psychology as the science of meaning and meaning making (Smith, 2004), which makes it an appropriate approach for the aims of this study.

At the core of IPA is a dual emphasis on the detailed examination of individual lived experience and how individuals make sense of that experience (Eatough & Smith, 2008; Smith, 2004). The interpretative component of IPA allows for the analysis to be informed by theoretical constructs (Larkin, Watts, & Clifton, 2006; Smith, 2004). Kegan's (1982, 1994) constructive-developmental theory provides a theoretical framework for this study. Constructive-development is a theory of the development in the way one constructs meaning of experience. Therefore, it provides a useful framework for considering the structure with which the THP program participants make meaning of their experiences in the THP programs, as well as any growth in their meaning making through participation in a program. Therefore, Kegan's constructive-developmental theory is foregrounded in the analysis for Research Questions 3.1 to 3.3. However, in the analysis for Research Question 3.4, which focuses on the content of individual participant program experience, it is the participants' accounts of their program experiences rather than any theoretical framework, that is central.

Interpretative Phenomenological Analysis (IPA)

IPA uses first-person accounts as data and a style of analysis that is idiographic, phenomenological, and interpretative (Larkin et al., 2006). Being strongly idiographic, IPA expects a detailed examination of an individual case

before moving on to another case, and only when each case has been reviewed independently is there an attempt to cross-analyse the cases for themes (Smith, 2004). IPA is phenomenological in its concern with the individual's explicit perceptions (Smith, 2004). However, IPA has joint underpinnings in phenomenology and hermeneutics and therefore, places emphasis on the process of interpretation, which is both subjective and reflective (Eatough & Smith, 2008). IPA research is said to involve a *double* hermeneutic; while the participant is trying to make sense of their experiences, the researcher is also trying to make sense of the participant trying to make sense of their experiences (Smith, 2004). Consequently, although the objective is to understand and describe the participant's experiences, it is important to recognise that any account of participant experience is co-constructed by participant and researcher (Larkin et al., 2006; Mishler, 1991; Smith, 1996). The researcher is at once accepting of what the participant has said at a summary level, while also reflecting on these words in a more probing manner (Eatough & Smith, 2008). In particular, as previously mentioned, the reflective process has been guided in part by Kegan's constructive-developmental theory.

Constructive-Developmental Framework

In constructive-developmental theory, there are five stages (or systems) of qualitatively different ways of constructing meaning, referred to as natural epistemologies, with each stage having a Subject/Object relationship where what is *Subject* (unconscious to us) in one stage becomes *Object* (within our perspective and control) in the next (Kegan, 1994). Growth, for Kegan, involves both an emergence from embeddedness and a new relating to that which was previously embedded. Intrinsic to this growth are the contexts in which we are embedded, referred to as our *holding environments* (Kegan, 1982). For Kegan (1982),

There is never just a you; and at this very moment your own buoyancy or lack of it, your own sense of wholeness or lack of it, is in large part a function of how your own current embeddedness culture is holding you (p. 116).

A holding environment has three primary functions: holding on, letting go, and remaining (Kegan, 1982; see also Popp & Portnow, 2001). A holding environment holds onto a person by *acknowledging* how that person thinks and feels and *joining*

the way that person makes meaning of their world. A holding environment lets go of a person by gently pushing on the edges of their meaning-making system and *challenging* their current way of knowing. This requires experiences and ideas that the person's current way of understanding the world cannot adequately make sense of in order to promote the creation of a new way of making meaning. Finally, the holding environment remains, so as to *scaffold* the new meaning-making system through the integration of new experiences, interactions, thoughts, and feelings. Although this kind of transformational growth is desirable, it can be uncomfortable and, therefore, is often resisted. Accordingly, what is essential for growth is an environment that both challenges and supports and, more importantly, strikes the right balance between that challenge and support.

Each of the five stages of constructive development and its related Subject/Object relationships are outlined in Table 2.1 located in Chapter Two. These five distinct ways of making meaning will be referred to as the *Impulsive* way of knowing (stage one), the *Instrumental* way of knowing (stage two), the *Socializing* way of knowing (stage three), the *Self-Authoring* way of knowing (stage four), and the *Self-Transforming* way of knowing (stage five; adapted from Kegan, 2000; Popp & Portnow, 2001). The three stages of meaning-making most relevant to this study are the Instrumental, Socializing, and Self-Authoring stages, as adolescents are likely to be somewhere in, or transitioning between, these ways of knowing. Each of these stages is outlined below. These outlines are drawn from the work of Kegan (1982, 1994, 2000), supplemented by the work of Berger (2002; 2003), and members of The Adult Development Research Group of the Harvard University Graduate School of Education (2001).

The Instrumental knower. Kegan's second stage of meaning-making, the Instrumental way of knowing, is typified by a concrete orientation to the world. At this stage, making abstractions is difficult and thinking is much more dualistic in nature: good versus bad, right versus wrong, etc. People at this stage construct meaning through the filter of their own needs, wants, and interests; their single point of view. As such, they can appear self-centred. While they are aware that other people have opinions and beliefs distinct from their own, they cannot hold

their own perspective and another person's perspective at the same time. These meaning makers tend to view other people's perspectives in terms of the implications for their own interests. Instrumental knowers consider other people as either pathways or obstacles to having their concrete needs met, and relationships are much more transactional in nature, based on a kind of tit-for-tat mentality.

Instrumental knowers tend to be focused on the rules and directions for doing things the right way, as well as the concrete consequences of their actions. Self-esteem is derived through the achievement of concrete goals and doing things the right way. They value authority figures, such as teachers, and feel supported when these people provide specific advice and explicit procedures so that they can accomplish their goals. Students at this stage will adopt the teacher's point of view but will not be able to reflect on it or consider it in relation to their own perspective. They view knowledge as a possession, an accumulation of facts and skills, and focus on finding the right answers and the correct way of doing things.

The Socializing knower. When people reach Kegan's third stage of meaning making, the Socializing way of knowing, they have the capacity to think abstractly, to hold multiple perspectives simultaneously, and to self-reflect. They can be devoted to something that is greater than their own needs. They also no longer see others as simply a means to an end. Relationships and mutuality are at the heart of this meaning making system. Meaning makers at this stage may function from a sense of loyalty to a larger group and are able to subordinate their needs and wants to those of the group. Socializing knowers internalise the ideas, values, and feelings of the institutions, cultures, and people that are most important to them. Knowledge is seen as something that comes from authority and experts who hand down the truth. This relationship to public authority is at odds with self-regulation which requires a sense of personal authority.

While people at this stage are no longer Subject to their needs and interests, they are Subject to the perspectives of valued others. As a consequence, these knowers can find it difficult to express their own views or to combine the best parts of several ideas into their own new one, and this lack of autonomy can make them seem robotic. Socializing knowers feel responsible for the feelings of others.

Equally, they depend on authority figures and important others for acceptance, belonging, and a sense of identity. They have a need to connect with valued others around shared beliefs and a common sense of identity or purpose. They find comfort in similarity, while difference, criticism, and conflict can be threatening. True self-esteem is difficult for these knowers because they have not yet developed an internal source for feeling good about themselves; rather, their esteem is derived from the opinions of others. Consequently, the concerns of Socializing knowers revolve around understanding other people's feelings and judgements about them, and they may be more inclined to take things personally. Moreover, they may need external validation in order to feel successful.

The Self-Authoring knower. At stage four, meaning-makers are autonomous, self-authoring, and self-regulating. These Self-Authoring knowers have thoughts, feelings, and beliefs that are independent from the ideas, values, and feelings of the institutions, cultures, and people that are most important to them. People at this stage are no longer Subject to their relationships or the internalised perspectives of others. Knowers at this stage can generate and evaluate various standards, values, and ideas, and can mediate among them, using their own self-governing system. They evaluate experiences by reference to their own self-constructed goals. For Self-Authoring knowers, knowledge is understood as constructed and tenuous, rather than given. Students at this stage want to create and explain their own complex ideas, and they are comfortable holding ideas or opinions that differ from those of their teachers. They are able also to self-direct their own learning.

Self-Authoring knowers are not defined by others and can distinguish the opinions of others from their own opinions. Development of this independent thinking is supported, for example, by teachers who value their ideas. While an Instrumental learner prefers a teacher who gives them the information that they need to be successful, a Self-Authoring student is more interested in being part of the learning process and, therefore, would be frustrated by such an approach. These knowers can question the expectations and values of others, take stands, and solve problems with their own independent frames of mind. Equally, a person operating

at this stage can consider the opinions of others in order to enhance their own ideas, values, and understanding. Self-Authored knowers can appreciate the differences between themselves and others rather than needing to find the similarities, and they can see conflict and contradiction as ways to learn, provided that such differences are not too great. At this stage, however, people are Subject to their own self-authored system of meaning making, rendering it difficult for them to question that system.

The sub-phases of development between stages. The outline above establishes the broad base of the three key constructive-developmental stages of meaning making most relevant to the THP program participants. Growth in meaning making is toward greater complexity; the more people can take as Object, the more complex their perspective becomes because they can examine and act on more. Development from one stage to the next is a gradual and active process of “increasingly organising the relationship of the self to the environment” (Kegan, 1982, p. 113). Consequently, people are rarely at a distinct stage. Rather, they are generally somewhere between stages and, therefore, are Subject often to the conflicting interests of two stages. In order to account for these in-between stages, Kegan’s constructive-developmental theory proposes four sub-phases between each of the five primary stages of meaning making. For example, between stages 2 and 3 are the following sub-phases: 2(3), 2/3, 3/2, 3(2). The first number indicates the dominant primary stage, and this number reverses midway through the sub-phases. These sub-phases can also be expressed in words: Instrumental(+), Instrumental/Socializing (Instrumental-dominant), Socializing/Instrumental (or Socializing-dominant), Socializing(-). A brief description of the sub-phases between the Instrumental and Socializing ways of knowing is presented in Table 7.3 by way of example.

Table 7.3
Constructive-Developmental Sub-phases between Kegan's Instrumental and Socializing Ways of Knowing

Sub-phase	Description
Instrumental(+) 2(3)	At this sub-phase there is an emergence of dissatisfaction with one's current Instrumental way of knowing. The person begins to see elements of a Socializing way of knowing, but is still entirely Subject to Instrumental embeddedness (e.g., I can bring another person's point of view inside me or see how they might be taking a point of view on me, but only as a source of information for my own efforts to meet my needs and interests).
Instrumental/Socializing (i.e., Instrumental- dominant) 2/3	At this sub-phase, a full Socializing system is operating in conjunction with a full Instrumental system (e.g., I can bring another's point of view inside and see how they might be taking a view on me, and I can also derive my thinking or feeling as a consequence of my seeing that point of view), but an Instrumental way of knowing is still dominant. The person has to work for a Socializing way of knowing and can talk about a time when they didn't understand things this way.
Socializing/Instrumental (i.e., Socializing- dominant) 3/2	At this sub-phase, an Instrumental system is still operating in full, but a Socializing way of knowing is dominant (e.g., I feel obligated to follow my group's views, but sometimes I just want them to listen to what I have to say).
Socializing(-) 3(2)	At this sub-phase, a Socializing way of knowing is more matter-of-fact, but the person still has to work to avoid slipping back to an Instrumental way of knowing (e.g., I can feel uneasy or confused when there is no "right" answer or way of doing something).

It is important to note that the actual evolution from one full constructive-developmental stage to another can take years (Kegan, 1994), and its process varies from individual to individual (Helsing, Broderick, & Hammerman, 2001). More importantly, while growth is always in the direction of greater complexity, each stage has its own internal consistency and shouldn't be judged by any other stage. The purpose of using a constructive-developmental lens to assess the meaning-making complexity of the THP program participants is not to rank the participants; rather the primary aim is to better understand how the participants interacted with their THP program (including differences in those capacities across program participants) and how the program might be more supportive of their meaning-making capacities and development of those capacities.

Assessing constructive-developmental stage. Constructive-developmental stage is assessed by conducting a *subject-object interview* (SOI; Lahey et al., 2011). The form of this interview and scoring process is described in more detail in the sections that follow. There are very few studies found in which the SOI has been applied to evaluate the meaning-making capacity of adolescents (see McCann, 2005; Villegas-Reimers, 1996). However, college-aged students and adults have been assessed using the SOI (in relation to college-aged students, see e.g., Gabb, Tinberg, & Weisberger, 2011; Lewis et al., 2005). Kegan (1994) brought together a number of studies with highly educated participants ranging in age from 19 to 55, and found that 13% of those participants were making meaning at the Instrumental stage or transitioning between an Instrumental and Socializing way of knowing; 46% were making meaning at the Socializing stage or transitioning between a Socializing and Self-Authoring way of knowing; and 34% were fully making meaning from a Self-Authoring system (Kegan, 1994). When Kegan (1994) considered only those studies with a participant demographic more representative of the general population by social class and level of education, he found an increase in the percentage of people at the earlier stages of meaning making. While Self-Authorship generally is considered most relevant after secondary school, meaning making at this stage has been found in individuals facing marginalisation or challenging environments, including at-risk adolescents (Baxter Magolda et al.,

2010; Pizzolato, 2003). It has been suggested that these youth may be stimulated into Self-Authorship earlier as a consequence of facing and overcoming difficult life experiences that challenge their ways of knowing (McGowan, 2016). These findings are relevant to our study which is focused on adolescents from a more disadvantaged segment of the population. While age has an effect on stage of constructive development, given the range of stages found at various ages, a person's age is not conclusive (Kegan, 1994). Therefore, an assessment of constructive-developmental stage cannot be made based on age alone. In relation to gender, Kegan suggests that the meaning-making structure measured by the SOI does not differ for males and females (Kegan, 1994). Villegas-Reimers' (1996) research in connection with adolescents supports these propositions.

It is preferable to assess constructive-developmental stage at a particular point in time with an SOI conducted at that time. Equally, pre/post SOIs are the ideal method for assessing any change in constructive-developmental capacity. However, the conduct of pre/post SOIs was not feasible given the resources available to conduct the study. Nevertheless, because the SOIs in this study asked participants to reflect back on their experiences in a THP program, some participants specifically indicated an earlier way in which they made meaning. For example, a person might speak about an earlier time when they felt dissatisfied with their way of knowing or a time when they didn't understand things the way they do now. This type of reflection is not uncommon in an SOI and can provide evidence for assessing current meaning making (Lahey et al., 2011). Consequently, it is suggested that using a post-program SOI to investigate the constructive-developmental stage of participants at the time of participation in a THP program, as well as any growth in constructive-developmental stage stimulated by the program, is a reasonable approach.

Interview Process

An SOI was conducted with each participant for this study. The SOI provides a method for assessing a person's meaning-making stage based on Kegan's (1982, 1994) constructive-developmental theory. I conducted 12 of the interviews and Dr. Michael Cavanagh conducted one of the interviews, in order to ensure that each

THP program mode was represented. Both of us have been formally trained to administer, and certified to score, the SOI.³⁰

The semi-structured format of the SOI uses prompts (e.g., proud, moved, anxious, uncomfortable, torn, angry, change) to direct the discussion. For this study, the prompts were directly connected to a participant's experience of a THP program. In this way the interviews could serve the dual purpose of providing structural evidence of a participant's meaning-making stage, as well as other content-driven insights into aspects of the participant's experiences of a THP program. The prompts were grouped into three categories: "A" for positive experiences, "B" for negative or more challenging experiences, and "C" for a single prompt related to the element of the program experience that was most important to the participant. The aim was to discuss at least one prompt from each category during the course of the interview, if that was acceptable to the participant.

The interviews generally lasted for 90 minutes in total. During the first 20-30 minutes, the interviewer provided an introduction. First, the interviewer let the participant know the approximate length of the interview, that the interview would be recorded, and that their name and any person or place they mentioned would be sanitised when the recording was transcribed. Moreover, the interviewer advised the participant that the discussion would be guided by them and that the aim was for the interviewer to understand their experiences of the THP program. The interviewer explained that the interviewer might ask questions of the participant about experiences the participant described until the interviewer felt like they understood what the participant was saying as best as they could. The interviewer further explained that these questions might bring the participant to the edge of their understanding of the experience and that if they did not understand the

³⁰ In 2016, I attended the Subject Object Interview Course (<http://www.subjectobjectchange.com/>), a three-day workshop, designed by Subject Object Change in partnership with Minds At Work, to train practitioners and researchers in how to conduct and analyse the Subject Object Interview. Dr. Cavanagh received his training in 2006 through a four-day workshop, conducted by Dr. Jennifer Garvey Berger. Competency for both of us was assessed via the submission of five accurately scored (within one standard deviation) SOI protocols following the workshop.

question or did not have a response, that was okay. The interviewer made sure the participant was aware that they did not have to talk about anything they did not want to discuss, and they could terminate the interview or change topics at any time. The participant then signed an additional consent form in order for the interview to proceed.

After the consent was signed, the interviewer handed the participant a set of index cards with a prompt on the front of each card. The interviewer then took the participant through the series of prompts and for each prompt asked them to think about a time in the program when they felt a particular way related to the prompt (see Appendix Y for a complete list of the prompts). An explanation of the prompt *Proud/Successful* follows by way of example.

If you were to think back over your time in the program, and you had to think about times you felt proud or successful, for example, because you had achieved something that was difficult for you, are there one or two things that come to mind?

The participant was given time to reflect and make some notes on the index card for each prompt. Participants were able to choose the prompts they wanted to discuss, with a suggestion that they try to choose one prompt from each category.

The recorded portion of the interview began with the participant being asked to a card (and therefore, a prompt) and to describe an experience that expressed that prompt. For example, a participant might describe a time when they felt proud or sad or angry. The interviewer would actively listen to the narration of the experience and use questions to probe for the structure of the participant's meaning making and the edges of their understanding of that experience. These questions focused on uncovering a number of structural aspects of a participant's meaning making: e.g., what is most at stake for the participant in a given situation; what can the participant take responsibility for, reflect upon, and exercise control over (i.e., what is Object for the participant); what aspects of the participant's life are directed more externally (i.e., what is the participant Subject to; Kegan, 1994, 2000). Examples of questions include "What is most important to you about x?", "What is the best/worst thing about x?", "How do you know x?" (for further information on administration of the SOI, see Lahey et al., 2011). Oftentimes interviewees

experience the SOI as a helpful tool for discovering aspects of their understanding of themselves, their relationships, and the world around them (Berger, 2010). Therefore, it is important to recognise that the interview itself can act as an intervention. Further detail on the administration of the SOI is included in Appendix Z, and a complete set of the prompts is set out in Appendix Y.

Ethics

In addition to the parental consent obtained for student participation in a THP program and the associated research, each interview participant signed an additional consent specifically related to the interview (a copy of this consent is included in Appendix J). Additional details on the consent process is described in the section above, “Interview Process.”

Data Analysis

Recording and transcription. All of the interviews were digitally recorded with the permission of the participants, and subsequently transcribed using HyperTRANSCRIBE software. I completed all of the transcriptions and then uploaded them into NVivo for further analysis.

Scoring constructive-developmental stage. As mentioned above, I have been certified to score the SOI (see footnote 30). Andrea Brownlow was also engaged as a second certified scorer, to score each of the interviews. The interviews were scored using the principles and techniques described in the *Guide to Scoring the Subject-Object Interview* (Lahey et al., 2011). Four tests of inter-rater reliability have found inter-rater agreement of constructive-developmental stage within one discrimination (of a possible 21 stages, including sub-phases) to be between 82% and 100%, with an inter-rater agreement in the 70-80% range suggested as reasonable (Lahey et al., 2011). We demonstrated inter-rater agreement within one discrimination to be 61.5% and within two discriminations to be 77%. While these numbers are somewhat lower than expected, these were the first SOI interviews either Andrea or I had scored. Lahey et al. (2011) suggest that scoring the SOI is a learnable skill that improves with practice, and this was our experience as well. It is also important to note that these interviews are more generally conducted with adults. As the subjects for these interviews were disadvantaged adolescents, some of

whom described challenging life experiences, it is suggested that these circumstances may have made the interviews more difficult to score. Moreover, the gap between program completion and interview added an additional level of complexity to the scoring. Each interview with any scoring discrimination was discussed at length between us until a final score was agreed.

Coding. Coding assists with interpretation. I used open coding to uncover patterns of meaning, first within an interview and then across interviews (Creswell, 2013). Because I intended to score the interviews for meaning-making capacity, I intentionally did not make any notes of my initial impressions of a participant, so as to reduce the opportunities for bias in my scoring. On my first read-through of each interview for the purpose of scoring the participant's constructive-developmental stage, I made notes of common themes and concepts that were arising. Moreover, I was interested to consider participants at different constructive-developmental stages and their corresponding views of aspects of the program, such as teamwork, leadership, and any program challenges and support. I also explored the data for participant views on specific aspects of the THP programs considered to be relevant to program outcomes in the literature, including the one-on-one and group coaching, the Community Project, and interaction with the coaches and other program facilitators, as well as peers. I used coding within NVivo to categorise the interview data into these emerging themes and classifications. As the interpretative process requires a sustained immersion in the data, I reviewed the transcripts multiple times both individually and across individuals within themes in order to identify any sub-themes or contexts in which those themes occurred.

Potential Bias

As a past coach on a THP program, I came to this research with some pre-existing beliefs in relation to the interaction between participants' constructive-developmental level and their experience of the program. Moreover, having interviewed the participants, I was conscious that I might form a view about them which could create an expectancy bias that might impact my scoring of their meaning-making capacity. I took a number of steps in order to minimise the potential impact of these biases: I waited approximately three months before

scoring the interviews; I engaged a second scorer; and I discussed my interpretation of the participants' structural data with one of my supervisors, Dr. Michael Cavanagh, who was experienced in conducting SOI interviews.

Quotations from Interviews

Excerpts from the interviews are quoted in the results section. Some quotations have been lightly edited, for example, to remove interjections, filler words, and repetition. In addition, for the sake of coherence and brevity, occasionally part of a quotation has been excluded or comments about the same subject from different parts of an interview have been placed together as a single quotation. In each case where this has occurred, an ellipsis indicates the place of departure from the original quotation (note that an ellipsis may be used also to indicate a pause within a sentence). Edits were made only for clarity, conciseness, and readability, and with the intention of not altering the sense or tone of the original quotation.

Summary

This section described the methodology and procedures specific to the qualitative analyses carried out in this study. Within qualitative research, reliability is concerned with the dependability of the data and validity relates to the trustworthiness of the findings (Kirk, 1986). The methodology and procedures outlined in this section, together with the details provided in Chapter Four, demonstrate the considered and thorough approach that was taken to establish and answer the research questions for this study. The following section presents the results of the analyses undertaken in connection with those research questions.

Results

Introduction

This section presents the findings from the analyses of the participant data in response to the research questions for this study. It was anticipated that participants might approach their experience of THP programs from different systems of meaning making. The aim was not only to assess these different ways of knowing, but also to understand how these different meaning systems might impact program experiences for participants and whether the program provides the most

effective and nourishing holding environment for its participants. The results of the analyses are presented below in the order of the research questions outlined above. All participant names referred to in this chapter are pseudonyms.

Results of Research Questions

Results of Research Question 3.1: Assessment of constructive-developmental stage. Research Question 3.1 arose out of a curiosity about the fit between the meaning-making capacity of the THP program participants and the expectations of the THP program process and delivery. Accordingly, the research began by seeking to assess participants' current constructive-developmental stage.

Constructive-Developmental stage at interview. At the time of their interview, the 13 interview participants were assessed at six different constructive-developmental phases ranging between 2(3) (i.e., Instrumental(+)) and 3/4 (i.e., Socializing/Self-Authoring). For more information on the various subphases of development, refer to Table 7.3 above. Nine of the 13 participants (69.23%) were transitioning somewhere between the Instrumental and Socializing ways of knowing. Complete results are set out in Table 7.4. Pearson's correlation coefficient was used to test for associations of scored constructive-developmental stage with age, gender, program mode, and time since program completion. Correlations ranged from $-.04$ for program mode to $.25$ for time since program completion and were not significant. However, the sample size is small and not every program was fully represented by each gender, making it difficult to assess significance or draw any firm conclusions. Some passages from the participant interviews are included below to highlight some of the salient features of those ways of knowing expressed by the participants.

Table 7.4
Interview Participants and their Constructive-Developmental Stage

Participant	Gender	THP Program Mode	Age (years) at Interview	Years Since Program Completion	Scored C-D Stage at Interview	Possible C-D Stage at Program Commencement	Sub-Stages from Program to Interview
Amy	Female	Arctos	15	0.67	2/3	2	2
Beth	Female	Arctos	15	0.67	2(3)	2	1
Cathy	Female	Arctos	NR	0.67	3(4)	3	1
Emily	Female	Arctos	16	1.17	3/2	2/3	1
Daisy	Female	Arctos	16	1.83	3/4	3	2
Fran	Female	James Craig	18	2.25	3(4)	3/2	3
Alex	Male	James Craig	17	1.50	3/2	2/3	1
Ben	Male	James Craig	17	1.50	2(3)	2	1
Charlie	Male	James Craig	17	2.83	3/2	2/3	1
Eric	Male	Outward Bound	16	1.25	3/4	3	2
Grace	Female	Coaching Only	15	0.25	2/3	2	2
Holly	Female	Coaching Only	14	0.25	3(2)	2(3)	3
Dan	Male	Coaching Only	15	1.25	3/2	2/3	1

Note. Names are pseudonyms. C-D = Constructive-Developmental. Scored C-D Stage at Interview indicates the assessed current constructive-developmental stage at the time of the interview. Possible C-D Stage at Program Commencement is provided where the participant revealed an earlier constructive-developmental stage at the time of THP Program participation. Sub-Stages from Program to Interview indicates the number of sub-stages between Possible C-D Stage at Program Commencement and Scored C-D Stage at Interview. NR = self-reported age not reported by participant.

Instrumental-Socializing systems. On average, participants were assessed as being in the transition between the Instrumental and Socializing ways of knowing at the time of the interview. For example, Beth (who was assessed as fully making meaning from an Instrumental system), spoke about her teachers in the following way:

I like following instructions. I don't really like giving them. Like with teachers and students, teachers give you something and you do it, and you're happy with it if you agree with it, so I like that person.

Consistent with an Instrumental way of knowing, Beth reflects a concrete assessment of the role of teachers and students, and the desire for a teacher who will tell her how to do things the right way. Another participant, Grace (also making meaning from an Instrumental-dominant system), noted how "it's good to be known as a good student and stuff, not like a naughty one," reflecting the dualistic

pattern of thinking common in an Instrumental world. When asked why it would be important to be known as a good student, she replied:

Because it can help you within the school. So if you want to do a program like [the THP program], you get chosen. So you want to have those opportunities instead of being a disturbance kind of thing.

What seems most at stake for Grace is meeting her needs and satisfying her goals. Another participant, Emily, spoke about the challenge of having a difference of opinion with her teachers:

With teachers it's hard 'cause you have to make sure you don't cross the line of making them pissed off 'cause then they'll have a bad opinion of you and then that affects your marks and they can get you on detention, and I don't want that.

Although Emily was assessed as having a Socializing-dominant system of meaning making, she still makes some meaning at an Instrumental stage. Her focus here on the concrete consequences of her actions (bad opinion, bad marks, detention) is more reflective of an Instrumental way of knowing. If she were understanding this situation from a Socializing way of knowing, she would be more focused on her relationship with the teacher and the teacher's "bad opinion" might make her feel bad about herself.

Charlie was also assessed as making meaning in the transition between an Instrumental and Socializing stage. At an Instrumental level, he was focused on getting things right and people were useful to help him meet this goal. When asked what it would feel like if he did something wrong and got negative feedback, Charlie said:

that would make me feel upset to know because I know I stuffed up, and I don't want you telling me 'You stuffed up' because I already know I stuffed up ... and to know that they're not going to help me fix my problem that would also hurt as well. Like 'You're not gonna show me how to fix it, but you're gonna give me this negativity?' - like 'What are you doing?'

Charlie's concern was not with the other person's judgement of him, but whether they would help him fix his mistake. Making meaning of the situation in an

Instrumental way, Charlie responded to the situation by finding someone else who could help him.

Emily exhibited a more Socializing way of knowing when talking about the importance of her friendship group.

[I]t means there's someone outside of family and closer relations that I can still talk to and I can trust to not tell everyone else sort of thing, but someone who gets it as well 'cause we have similar lives and similar families, so they're able to also relate, but I can also tell them about stuff and not be judged and just be underlyingly like understood.

When asked what the hardest part was about being judged, she replied:

[F]eeling like you're not normal, I guess, or that there's something wrong with you. Um, I don't know. Every time I have been judged in the past, I don't necessarily like beat myself over it and try to change it immediately, but it does lower your confidence quite a bit and ... it kind of stops you from talking about it and talking about how you actually feel about things because you're worried that people will think 'Oh, you're weird' because of it.

In a Socializing world, how others perceive you becomes how you perceive yourself. Alex, an immigrant to Australia (and assessed as Socializing-dominant), spoke about the close relationships he has with his friends overseas and the lack of that closeness with his peers in Australia:

I don't feel like I can fully trust anyone because ... I just don't know why. People here are just different. They don't match with what I want in a best friend. [Where I'm from] people are honest to each other. They can keep a secret. You can trust everyone. You can rely on everyone to do stuff, and I just do lots of things the way that matches with my friends over there. ... things like embarrassment and awkwardness and isolation won't come from people that you have total trust with ... and I have it with those people.

Alex also expressed his concern about what it would be like with his old friends back home when he visits them in the future:

I'm pretty sure that part of me will be nervous if they will accept me; if they have their ... like if I have to talk to new people that they've made friends with A small part of me will just feel nervous because what if those people

don't wanna do all the stuff I want to do. What if they have other stuff to do with their own lives. What if they're just like 'Ok, yeah hi. You're here, but there's other people that we're friends with and they don't know you, and you're just a random from Australia.'

The emphasis in these passages on belonging versus isolation and similarity versus difference are consistent with a Socializing way of knowing.

Self-Authoring system. While we did not interview any participants who were making meaning from a more dominant Self-Authoring system, there were four participants who demonstrated a beginning movement from a Socializing to a Self-Authoring system, with two of these participants making some meaning through a Self-Authoring lens. Daisy reflected a Self-Authoring way of knowing when asked how she felt about situations in which someone might take a different perspective to her own:

[With] those people that [have a] different opinion from you but are still mature enough to have a conversation and hopefully hear you out, I get excited about hopefully getting to understand perhaps something different. But there's people who just don't want to listen and then there's no point in talking to them. So, I think for those people who can hopefully, you know, open up the avenues for another way to view things, that's what I'm excited about. Hopefully trying to convince them that there's another way to look at things.

When asked whether the best thing in this situation would be to convince the other person of her view, Daisy responded:

... not necessarily 'cause funny enough, when I engage in those type of conversations, my perspective gets changed, so ... people [who] have the mentality that 'x' opinion is right or wrong ... I think having a right or wrong opinion - I think that's the wrong way to look at things. ... My teacher was like 'I don't care what your opinion is. If you can convince me because your arguments are strong, then that's what a good opinion is.' Because if we end up talking and we end up, you know, like swapping ideas and then I realise 'Wait. Oh, that's really interesting. Oh wait. That's a really strong point,' then maybe my perspective is changed, and I'm always open for that as well. ...

[But] it doesn't have to be the end goal of like the person agreeing with the other person's opinion or seeing it like that, but I think just having that conversation and perhaps seeing that your opinion is not the only one. I think that's what my end goal is when I have conversations like that.

Rather than feeling threatened by difference, as a Socializing knower might, Daisy reflects an understanding of difference as an opportunity to learn “something different” and “another way to view things.” While she is strong enough in her own views to try to convince someone else about her perspective, she also is open to reframing her perspective in a considered way. Daisy reflects how for someone making meaning at this stage, a teacher who values and encourages alternative views can stimulate meaning-making growth.

Conclusion. The 13 interview participants varied in their constructive-developmental stages. No association was found between participant's stage of meaning making and their age, gender, program mode or time since program participation. On average, participants were somewhere transitioning between the Instrumental and Socializing ways of knowing at the time of their interview.

Results of Research Question 3.2: Constructive-developmental stage and program experience. Given the different constructive-developmental levels from which individuals can understand their experiences, Research Question 3.2 sought to explore whether there was a relationship between a participant's way of knowing and their experience of a THP program. The interview data was analysed for similarities of experience among participants making meaning from similar constructive-developmental stages, as well as differences across diverse meaning-making capacities. Particular attention was paid to aspects of the program such as program structure, coaching sessions, and the challenges in the program that afford participants the opportunity to experiment and make mistakes, together with the support that assisted in managing those challenges. There is also more of a focus on the Instrumental and Socializing ways of knowing since these were the systems from which most participants were making meaning.

Rules, boundaries, and structure. The THP programs seemed to be especially challenging for participants who were Instrumental-dominant knowers,

possibly because there is less structure to the THP programs than is found in a classroom setting. Moreover, the coaches explicitly do not take on an authoritarian role, operating in a manner quite different to that of a teacher. Beth, an Instrumental-dominant knower, had difficulty making sense of her Arctos Adventure Program before the program even began.

So being told that you're gonna go sailing ... on a boat like out at sea, not even on land, that's something I couldn't understand because I'm always on land walking. So being on water in the middle of the ocean was a bit like worried and scared because I didn't know what to expect. ... [I]f I'm not sure what will happen, I'll always be confused, quiet, and I'll just be frozen like a statue.

Beth's concerns about concrete aspects of the adventure, such as being on water, are consistent with an Instrumental way of making meaning. Having a clearer sense of the program structure prior to commencement may have allayed some of her fears.

The lack of rules and boundaries in the group coaching was also frustrating for some participants with an Instrumental way of knowing, particularly when the coaches did not manage the group as a teacher would manage the classroom. Holly, who likely began the program making meaning from a more Instrumental system (see below the results of Research Question 3.3 for more on her constructive-developmental growth), reflected on the anger she experienced during the program because of the failure of the coaches to handle rude, disruptive behaviour by some of the participants in the group:

Instead of just letting it be, [the coaches] should tell them and show them like you're being disrespectful, like you shouldn't do that and then getting that person to know 'Oh, I'm doing the wrong thing. I should stop doing that.' It's just like seeing if the coaches did do something about it, it would open [the girls] up and realise 'Oh I'm not doing the right thing. I should stop doing that otherwise people are gonna think I'm a bad person.'

Holly exhibits a dualistic and concrete orientation to this situation. For her, the girls can be either good or bad, and the person in authority should establish and enforce the rules. Participants at an Instrumental stage of meaning making found

greater comfort in groups where coaches took more control. When asked what made her group such a good team, Grace said:

I guess that we all knew each other and understood the boundaries. ... So if someone might take over so much, you could tell people like 'come on ... let me say something,' but our coaches kept us quiet, so that each individual could talk and put an idea down.

Similar to Holly, knowing the rules and boundaries for the group was important to Grace, as well as having a coach manage those boundaries and ensure the process was fair.

For Dan, who was moving toward a Socializing-dominant system, having coaches who were less prescriptive was a benefit rather than a hindrance.

I thought when we started this process [that] it's just gonna be a bunch of talking and, like most projects you see these days, someone comes out front and talks. And what I really liked is the Helmsman Project allowed you to express yourself, and work at sort of your own pace, to work at the way that suited you because you see programs that force you to work at a certain pace to get things done over a certain period of time, and even though we were given the 13-week timeframe, we never felt we were at a point where we had to rush. We were never at a point where we felt we had to overly stress about something, and that was really good.

The flexibility in the sessions seemed to give these students more autonomy. While participants making meaning from a Socializing system could lean into this level of challenge, concerns for these participants might arise from differences among group members and any feelings of judgement they might experience as a consequence. Such experiences in a group context are discussed further below.

Coaching sessions. In general, the Instrumental-dominant knowers seem to have preferred the one-on-one coaching sessions to the group coaching sessions. Of the four participants assessed as Instrumental-dominant at the time of the interview, all four found the one-on-one coaching beneficial. Amy preferred the one-on-one coaching sessions because,

[i]t's just one-on-one; not like other people, so [the coaches] were able to just concentrate and understand how I felt. ... If I say in front of my friends ... they say something about it. Whereas just coaches, since they're older, they're able to understand.

The opportunity to talk to the coach without anyone else there was also described by Ben and Grace as being helpful. For participants making meaning from an Instrumental system, the coaching relationship likely was most valued when it served their interests. Dan (transitioning between an Instrumental and Socializing system) noted how,

[y]ou had the time that it was just your coach and yourself, and you'd talk about whatever part of the process you were up to and that's what I thought was good was you weren't necessarily caught in the middle of something, if that makes sense. You didn't have others piping in on top of you. You had the chance to talk about what you wanted to talk about and explain things your way without the sense of other people jumping in on top of you if, if that makes sense.

Beth, an Instrumental knower, also preferred the one-on-one sessions because,

[i]f I told them [what was going on for me], they can help me ... I would get advice, and they can assist me now they know what's going on, and they can look out for me and watch over me if I needed it.

Like Beth, Charlie valued his coach because she was “a coach who was understanding, who knew what you wanted as a career, who would help you achieve your goals, [and] give you confidence in things.” Consistent with an Instrumental way of knowing, both Beth and Charlie experienced support as the provision of advice and direction that help them achieve their goals. For participants making meaning in an Instrumental way, the coaching relationship seemed more transactional and was regarded as supportive in a more concrete way, focusing on things the coach could do for, and give to, the participant.

Socializing knowers would likely view the coaching relationship as more mutual, and support would be found in the relationship itself and the ways in which the coach expressed care for the participant. However, some participants with a Socializing-dominant way of knowing did not seem to experience the one-on-one

coaching sessions as an element of the THP program that was particularly valuable for them. Daisy and Emily could barely recall those sessions, Holly expressly preferred the group sessions, and other participants assessed at this level did not discuss this element of the program in any detail in their interviews. At this stage of meaning making, it might be difficult for a coach and participant who only met independently once every two weeks, to develop the type of relationship such a meaning maker would understand as supportive. Additionally, it is difficult for Socializing knowers, who are subject to external perspectives, to have their own ideas and values separate to those of their group, family, or culture. This orientation could make the one-on-one coaching sessions feel threatening, not only because a participant at this meaning-making stage may have to find their own voice, but also because of a fear of judgement or criticism. Emily found that the hardest thing about being judged would be “feeling like you’re not normal or that there’s something wrong with you.” People at this stage experience themselves as a function of how others experience them, making the judgement and criticism of others potentially destructive to their self-identity. The consequence of being judged, Emily noted, is that it can “lower your confidence quite a bit and kind of stops you from talking about it and talking about how you actually feel about things.” Cathy, who was solidly making meaning from a Socializing system, described a situation in which one of the coaches suggested she was “bossy.”

I was sort of a little bit taken aback because I got all these people saying that I didn't feel bossy, ... but then the coach said that, so that sort of got me thinking in my head ‘so is one of, like my friends, are they lying or like what just happened?’ Like that was very um, what's the word, conflicting and very, yeah, I sort of like, I didn't know how to interpret it. Like if I was just being a little bit bossy, if I was being really bossy, and I sort of, like when I looked back on the situation in my head after she had said that ... I sort of was like ‘Um ... ok ...,’ and we sort of moved on, but that sort of like - it didn't hurt me, but it sort of bothered me a little bit and it sort of became awkward after that. It became really, I felt really awkward and like really uneasy about taking control of a situation as in like guiding the team.

Feedback that is unfavourable can be threatening to Socializing knowers because such feedback contributes to their self-image. In this instance, Cathy needed to canvass the views of the other girls in her group to determine if she was being bossy, rather than being able to evaluate the situation for herself, as a Self-Authoring knower might do. The conflicting perspectives of her teammates and the coach caused her to feel “torn.” Although the support of Cathy’s peers helped to alleviate the impact of the coach’s comment, Cathy still experienced a residual uneasiness about continuing to act in a leadership capacity for the group.

For someone with a more Instrumental way of knowing, such as Grace, the worst thing about being judged is that they might form an opinion about you and “it might not be true,” and the worst thing about that would be that “they don’t know the truth.” The concern here is with concrete consequences. Ben, also an Instrumental knower, reiterated this concrete concern with people spreading false information. If people hear your things, he noted, “they might spread them,” and “it might get altered in the process of spreading,” and people might “get the wrong idea about stuff.” The worst thing about this for Ben is that “they’re not getting the right information.” For Instrumental knowers there is only one truth and the concern with the judgement of others is the concrete consequence of people having false information rather than the truth. Socializing knowers, on the other hand, are at risk of internalising these judgements irrespective of their veracity, which can impact their sense of confidence and optimism about themselves.

For Instrumental knowers, teamwork is generally seen as beneficial only when it is in service of their self-interest. When asked whether it was important to work well as a team, Ben said, “Yeah, ‘cause then we wouldn’t really get any work done if we didn’t work well as a team, and that’s really the whole point of the thing.” Making meaning in an Instrumental way, Ben views his group as serving a functional purpose to meet his concrete goals. While Socializing knowers would be concerned more with the social dynamics of the group, Instrumental knowers view people as either helpers or hindrances to meeting their needs. As a result, they may have difficulty incorporating others into their activities. Alex demonstrated this way of understanding group work when talking about how he used to work within a team:

I would have either done work where I tried to force them to do something or leave something to them and if they don't do it, I would do it. I would try and fill everything that was teamwork oriented.

Now Alex says that he approaches group work differently. He doesn't force people to do what he wants, however, his need to meet expectations is still what is ultimately at stake for Alex.

Now I make sure to choose to cooperate with people that I might not necessarily like, but I know that they're reliable and they can meet the expectations we need to meet ... and if someone just doesn't wanna listen to me, I don't make them listen to me. I just say if you don't wanna listen to me then that's fine. You can just do your own thing, and I'll do my own thing.

Here, Alex continues to exhibit an Instrumental way of understanding the group dynamic. He needs to control his world in order to get his needs met, and people are either helpful to that process or not.

Amy, another participant making meaning from an Instrumental-dominant system, commented on how "it was hard to communicate" in the group coaching sessions "because they might have different perspectives to what I have." This experience is consistent with an Instrumental system in which one cannot hold their own perspective and another's at the same time. Amy also noted the difficulties she had sharing in the group sessions:

I felt uncomfortable when we had discussions about how we felt 'cause sometimes people would say this and other people would say something else, and I would have a different perspective on how I feel, which I thought it would be weird for me to say because people might disagree or go against how I felt ... I thought that by saying something different it might seem like a bad thing 'cause I thought there was only one way to do it and people all have to do it like that perspective and not like another perspective.

Instrumental knowers believe there is only one right way to do things and are afraid of getting it wrong.

For participants operating from a Socializing system, the team and the group coaching sessions were often the most memorable part of the THP program. A sense of belonging and mutuality are core to how a Socializing knower makes

meaning, and perceptions of interpersonal connectedness and acceptance can be self-esteem boosters. For these participants, the group dynamics in the program could be powerful. In her interview, Cathy named “teamwork” as being the most important part of the THP experience for her, especially because she hadn’t found that within her family. For her, teamwork meant that “everyone’s opinion mattered and everyone had input as much as they could or if they wanted to put their input in.” In a Socializing system, equality means that everyone’s needs deserve to be heard, but it does not necessarily involve an equal exchange (as would be important for an Instrumental knower) because some people need more than others. At the heart of teamwork for Cathy, was the feeling of “unity.” To her, this meant,

We're all together. We're all on the same page. We're all working towards goals that are similar or the same goal, and it's just very important to me to see that we're all together and happy and all working.

While Instrumental knowers focus on satisfying their needs and interests, Socializing knowers attend to their relationships. One aspect of the program that Fran valued was her peer group and how “kind,” “caring,” “comforting,” and “supportive” they were. Daisy spoke of how “closely bonded” she was with her group. Socializing knowers are in relationships and find themselves defined by those relationships. Fran noted how she learned from the other girls in her team and “adopted” some of their ways of being. On the other hand, difference and conflict can be threatening. Holly, a Socializing-dominant meaning-maker, described how her group resolved any differences:

If we had different opinions, we would think of a strategy to put it together. We would think of how we would like to do it. If one said, ‘Ok we wanna do this,’ we would think about it and see the disadvantages and the good advantages and then we would all vote for what we wanted most, and then all of us surprisingly voted for the same project, and it was really good.

Socializing knowers prefer consensus. Dan, who was transitioning between an Instrumental and Socializing system, acknowledged that since the program, when I am forced into a group scenario, I feel that I have improved that little bit thanks to the Helmsman Project because they taught ways to listen to

people, the ways to put yourself forward, the ways to present ideas, and the ways to sort of put ideas together and come to a consensus within the group. Eliminating differences in an equitable way can be important to Socializing meaning makers because it maintains group cohesion and happiness.

As Socializing knowers develop Self-Authoring capacities, the need for similarity fades and difference can be interpreted as an opportunity for growth. Talking about her team, Daisy said, “I see values in myself that are in [my team members] and I see values that they have in myself or I see ‘Oh, I want to be more like something that you have’ within my group.” Making meaning with a Self-Authoring system, Daisy can now appreciate the differences among her peers rather than seeing those differences as a threat to her self-identity. Being open to diversity of opinion allows Self-Authoring knowers to form more complex and inclusive solutions. For Instrumental knowers, as we saw with Alex, difference had a more all-or-nothing feel. For these knowers, it was either my way or your way, but not both, with the resolution of differences resulting either in a fusing with the other person’s perspective (to the exclusion of their own) or cutting them off.

For participants at all stages of meaning making, both the one-on-one and group coaching sessions could be experienced as positive and transformational, as well as challenging and scary. What seemed essential to the success of the coaching element of the program for participants at all stages was for the coaches to understand and meet them where they were at in order to scaffold opportunities to question their current way of knowing and experiment with new ways of knowing. The peer group could also provide important elements of both challenge and support for the participants.

Challenges and supports. A core element of the THP programs is to build resilience by creating challenging opportunities for participants to make mistakes from which they can recover. Participants at different constructive-developmental stages, however, made sense of mistakes and managed challenges in varying ways. Instrumental knowers were focused on the right way to do things and the bad things that can happen when you get it wrong. Accordingly, those participants making meaning in this way could be very fearful of making a mistake, particularly

when there was a negative concrete consequence to the mistake, such as “tipping the boat over” (Emily) or “the ship goes completely opposite than what it’s meant to” (Charlie). Moreover, sometimes the challenges could be “nerve racking” (Charlie) and sometimes “everyone was so far out of their comfort zone, people would just get frustrated and angry for no reason” (Ben). While such feelings could reflect an imbalance between challenge and support, many of the interview participants described being able to meet the challenges set by the program with the support of their peers, coaches, and other program facilitators. However, support looked different for different meaning makers. For Instrumental knowers, their fears could be alleviated when they were assisted to do something the right way. What was important to Emily when learning to sail the small yacht, *Arctos*, was how the coaches and crew,

[were] making sure I was doing the right thing, and they reaffirm even if you weren’t doing something wrong. They’d be like ‘Ok good job. You’re doing well’ all throughout it to make sure you knew that everything was good, which was good because it was reaffirming that I wasn’t messing up and that they were actually paying attention and caring.

For Emily, the most important thing was to get it right, and she felt supported by not only being told how to do a task but also by having her doing of the task reaffirmed,

[be]cause sometimes when someone tells me to do something, I may understand it, but I second guess myself and I think ‘Oh, did I hear that right? Is this what they meant?’ So being able to reaffirm it, then you know for certain I’m doing the right thing, so I’m able to keep going and not make a mistake and second guess myself.

This support and the perception of care Emily felt were in service of her need to do the right thing, reflecting an Instrumental way of knowing. Other participants could also provide this support to an Instrumental knower. For Beth, support came from another participant in their group who took on the role of leader.

...our leader was the oldest member in the group and she had more experience because she actually did sailing and that, so I had someone to look up to, someone to be a role model, so I wasn’t lost or anything. ...

Usually people say leaders are bossy and they tell you what to do and you don't like it, but to me I felt like I knew something to do, like I had a procedure to follow and took it off my head, so it wasn't bad for me.

Charlie also experienced the support of a peer. Charlie had expressed being worried about doing things the wrong way and the fear of someone pointing the finger at him saying "you did this" and being blamed for the consequences. He gave an example of tying a knot on the ship incorrectly. When he realised he needed to correct the mistake, he went to one of the other participants. Charlie expected "negativity," but his actual experience was a positive one. His peer said, "Ok, just undo it. We'll do it together," and that response made Charlie "feel amazing."

Like it made me feel as though if I stuff something up, I guess people are gonna support me around the ship. People are going to show me how it's gonna be done. And if you stuff up, you stuff up, and people will help you at the end of the day, and that's what I realised and that's amazing I guess, to know.

When asked why having that support was so important, Charlie said it was "because I would know no one's blaming me about anything, no one's telling me 'Hey you did this wrong, you're not gonna be able to do it again.' No one's telling me that."

Ben also experienced the support of his peers when facing the challenge of climbing the mast on the James Craig tall ship. Ben said, "I was kind of second-guessing myself 'cause it was very high, and I didn't think I could do it." When asked how he got the courage to make the climb, Ben noted that the other students kept encouraging him with words like "keep going - you can do it." Moreover, Ben found value in having the opportunity to watch other students make the climb: "It helped a lot 'cause then I could see what they did and then kind of follow it." Amy had a similar experience of being able to meet her fear of heights when climbing the mast with the support and encouragement of her peers and their "positive thinking." In each of these situations, participants making meaning from an Instrumental system found support in peers who helped them to find the right way to do things, whether through encouragement or modelling. Beyond the concrete help that these peers provided in connection with assimilative learning (as more knowledgeable others), the relationships and other emotional support created an

environment that allowed participants to be at their edge of meaning making, opening up the possibility to see a new way of understanding. For more on this transition, refer below to the results for Research Question 3.3.

While Socializing knowers also want to know the right way to do things, they are focused not only on what those in authority know and can tell them, but also on the qualities of those people and the connections they make. Charlie, who was making a shift between Instrumental and Socializing ways of knowing, spoke about how he used to not try things because of a fear of getting it wrong and getting the blame for the mistake. However, during the program, he had opportunities to make mistakes and experience support through that process.

[E]ven the captain one time came to me, because we were doing a little ropes course. He goes, 'You did that wrong, but I can help you fix it,' and I'm like 'Ok cool.' So, he showed me how to do it, and he showed me like a few times like probably 5 or 6 times - undid it, redid it - and then he gave it to me, and I did it wrong obviously for the first few times again, and then I got it to the point where I did it right, and I guess that's a memorable moment for me, and that makes the whole program important to me to know that there was so much support, not just from crew, not just from peers and teachers, but from my coaches who were sailing aboard, and from the company.

While these people seem important to Charlie in helping him to meet his own specific concrete needs, which is consistent with an Instrumental way of knowing, the support here is beginning to take on a more abstract character, reflecting a more Socializing system. Fran also spoke to her experience of climbing the mast and her intense fear of heights. For her the support of the group and her coach was instrumental to her success, however, what she valued most from the support was not the concrete direction or modelling, but "knowing that they did care about me and knowing that they did genuinely want to see me succeed and they did genuinely want me to just try it." It was the care and emotional support that buoyed this Socializing knower to meet her fears.

Community Project. Given the value that Instrumental knowers place on having someone to show them the right way, an expectation for self-direction could

be experienced as a relinquishment of support. For this reason, work on the Community Project,³¹ which occurred after the coaching sessions had finished, could be difficult for Instrumental participants. Without the coaches, Beth was unsure how she would get through this part of the program.

So the coaches have been with us through thick and thin. We've talked to them about our problems. They helped us, so when they had to go and not be there with us in our Community Project, we felt a bit sad because, like I said, I depended on someone, so they weren't there and I was like 'Oh, they're not here, so what do I now,' you know. Because I was sad [that] they were gone, I was a bit like 'I don't know what to do,' but then the leader she was like 'Ah, don't worry. You can do this,' so I still had a way to get through all my negative thoughts, and she solved them all for me cause the coaches were always there to solve things.

For Instrumental knowers, coaches provided support by solving her problems. In this instance, the group leader was able to provide the support that was needed when the coaching concluded. However, this may not have been the case in all of the program groups, potentially leaving some Instrumental participants feeling abandoned and lost. Ben, an Instrumental knower whose group did not complete the Community Project, felt this was because they "didn't have a teacher encouraging [them] to do it and making sure that [they] were on the right track the whole time." For Dan (making meaning in part from an Instrumental system), his group "got to the point where [they] had kept asking staff to do things and it got brushed aside," so they figured "why should we bother and waste the time and the effort into wanting to do this if it's not gonna happen." Moreover, the Community Project was considered to be "a big part of the program" (Ben), and failure to complete the Community Project had consequences, including not "graduating" from the program. Therefore, not getting this "right" could affect the program

³¹ The Community Project was an element of the THP program that was introduced about midway through the research. Two of the interview participants (Fran and Charlie) took part in programs where a Community Project was not formally offered.

outcomes experienced by the participant. Failure to complete the Community Project “annoyed” Dan because,

we had gone through the process of thinking about it and making and pitching it and practicing the pitch. We were taken out of class, missed work, had to catch up on work, and nothing's come as a result of it.

Dan would have liked to have “felt the sense of accomplishment that it had actually been completed.” Alex also noted that the “experience just didn't feel that complete” without finishing the Community Project, and “[he didn't] get the sense of achievement of completing something or achieving something.” From an Instrumental perspective, the focus is on the concrete status of completion and the consequences of non-completion. Even more irritating to Dan was the fact that the girls' group that did the program after them completed their Community Project.

What really made us angry is we put in the effort to help them to get theirs done and theirs got off the ground and happened and looked wonderful and their project was amazing, and nothing happened to ours.

Although he felt his group did what they were supposed to do, they did not get the support or the result he felt they deserved. There is a strong sense of the unfairness to it all. The consequence for Dan is that he is now “a bit sketchy” when it comes to projects like the Community Project, and “apprehensive when it comes to new programs similar to what [THP] does.” When Ben was asked whether he learned something from the experience, he said, “Yeh, I guess to not rely entirely on the fact that people will be there,” reinforcing a more Instrumental way of knowing. For more Instrumental knowers, the combination of challenge and support during this phase of the program may have lacked the right balance between challenge and support.

For participants making meaning within a Socializing system, it was not the simple goal of completing the Community Project that was most important, but rather the feelings and judgements of others around that completion. When Holly was asked what would have been the worst thing for her if the Community Project had not gone well, she said “just seeing the girls upset about it and the coaches seeing us fail. It just feels like we just let them down and let the community down and like that big failure and like guilt.” Many of these Socializing participants

wanted to complete the Community Project for the coaches, their school, and the other members of their team. They derived their self-identity through the happiness, praise, and appreciation of these people. Alex wished that he could see his coaches again because,

I don't wanna leave them with an image of someone that didn't finish his project or was quiet and had a lot of anxiety about things in life ... I wanna tell them about all the things I've done since, and I feel they would be very happy and they would see that I've gone past their expectations of the quiet kid and I've achieved something, and then they'll be like proud and happy and stuff.

Charlie's group did not have the option to do a Community Project, but he was equally concerned with not being able to show the coach what he had achieved since the program.

I would love to have said to my coach a few years back 'I did it. I finally have the confidence to do the things I've wanted to do and had the fear of doing, and hey I can do it.' And my coach, I guess, would be so proud of me, like being able to do these things, getting out of these terrible situations, and doing what I love.

When asked what it would mean to Charlie for his coach to know about these achievements, he said that it would make him "really happy because in a way that's giving her feedback that she did her job right and knew what she was doing." He not only wants his coach to be proud of him, but also wants her to know that he valued the job she did. This reflects the mutuality inherent in Socializing relationships.

Socializing knowers also found the Community Project challenging, but they seemed to be better able to find more support through peers and their emerging self-regulatory abilities. Holly initially was worried about her team's ability to complete the Community Project.

I just felt like the time was too short, and I was overthinking a lot of things like 'We're not gonna get this done. We're not gonna buy the things in enough time and nothing's gonna get finished,' and I just kept thinking that in my head ... [but the other girls would say] 'You can't just let yourself down.

Just try and think positive and say you know you're gonna do this,' and the girls just kept boosting me up, and that boosted my energy towards the project.

By internalising the positive thoughts of her teammates, Holly was able to persist in the project. Eric also mentioned how “internal conflict” in his group could have derailed successful completion of their Community Project. He suggested that it was the school’s belief in them “as the group that could get stuff done” that was a “driving force” in their ability to achieve project completion. Internalising this belief could provide the necessary boost to a Socializing knower’s self-confidence in this situation. Eric also observed that the project “was also a learning curve in the sense where getting it done involved listening to people,” something he had not been good at doing before. Completing the project, therefore, also required Eric to call on developing self-regulatory skills.

For Socializing knowers, completing their Community Project was an impactful experience. Being able to put some of his newfound skills into action was powerful for Eric.

It was like discovering Australia for the first time, and you've realised you're in new territory. You've been trained to do this. You've been set to do this. You know this is possible, but you don't know what to do ... I felt proud of the group and I felt proud of myself that we had some drama – well, not some – we had a lot of drama, but we still managed to finish it. So that was new territory for me, but it was also a very proud achievement.

It may be this kind of experience that assists with transfer of program learning back to the school context. For Daisy, the Community Project was important because,

I don't think our school ever experienced something that was so student led, and I think it was such a great opportunity to show people what we can do; that we can make a difference and set a precedent.

Implementing the Community Project made Cathy “feel happy because I’ve taken an experience that I’ve learnt from and I’m helping others with it, and it just made me feel really happy and joyful.” Holly had a similar reaction to her group’s project providing help to a homeless shelter.

The thing that stood out to me the most was seeing [the homeless people] happy and seeing the staff actually having a break and seeing [us] kids stepping up and changing the community around and seeing us kids happy and wanting to serve others to make them [happy] ... It just brings joyness and I reckon it would make a change in the future for other kids that are in kindy now gonna grow up and do the same as we're doing and just make a change.

Socializing knowers are apt to subordinate their needs to the needs of others and to derive their own happiness from the happiness of others. This way of knowing may make the Community Project more suited to participants making meaning from a Socializing system.

Conclusion. The interview data demonstrates how participants at a similar stage of meaning making understood aspects of their program experience in a similar way and how those at different junctures may have understood and interacted with those experiences in different ways. For those participants making meaning in an Instrumental way, what was most at stake were the person's own needs, interests, and desires. Beyond their immediate self-interest, such Instrumental meaning makers were concerned about knowing the rules and the concrete consequences of their actions and wanted to ensure they were doing things the right way. On the other hand, participants at a Socializing-dominant stage were focused on their relationships and sought out acceptance and approval in order to feel successful.

The interview data revealed that the Instrumental knowers often preferred the one-on-one coaching where the coach could focus on their needs and they did not have to work to incorporate others into their world. In contrast, the Socializing knowers recalled the group sessions as the most memorable, with a sense of belonging and acceptance being core to their self-esteem. The coaching sessions could also be challenging environments, with the particular challenges dependent on one's way of knowing. For Instrumental knowers this was often related to concerns about doing something the wrong way, and for Socializing knowers the

judgement of others could create unease. Support was found in the coaches but also in the other program participants.

Beyond the coaching, there was also evidence that the outdoor adventure experiences and Community Project created a rich environment for growth. These experiences provided opportunities to try new things and to make mistakes from which they could recover. Again, participants at different constructive-developmental levels made sense of mistakes in varying ways. Making meaning from an Instrumental system, Beth was fearful of getting things wrong and met challenges better when she was assisted to do something the right way. Charlie, who was moving toward a Socializing way of making meaning, also sought support for how to do things, but he was beginning to focus on the relationship that arose out of that support, with a particular need for approval from those providing support.

The Community Project provided a particular challenge for many participants. For the more Instrumental knowers, the lack of direction was often felt as frustrating, and participants, like Dan, seemed to give up without the necessary support. For the Socializing knowers, like Cathy, the concern was more with doing a project they felt proud of and being recognised for what they had achieved. Participants with the benefit of group support seemed to be able to manage the self-direction that was expected. Beth, an Instrumental participant, was buoyed by another group member with a Socializing way of knowing. Alternatively, a group with participants operating from a Socializing-dominant system may have been better able to self-direct the process as a group. Irrespective of meaning-making stage, the Community Project outcome (satisfactory completion versus failure to complete) seemed to contribute to the overall effect of the THP program on a participant.

Results of Research Question 3.3: Constructive-developmental stage at THP program commencement and growth through program experience.

Research Question 3.3 sought to consider whether participant interview data evidenced participants who were at a different constructive-developmental stage at the time of participation in a THP program and who experienced shifts in meaning-

making capacity through participation in the program. To the extent constructive-developmental growth was evidenced in the data, Research Question 3.3 also sought to isolate any particular program elements or critical events that participants experienced as challenging their meaning making and supporting the evolution of a new way of understanding.

Constructive-developmental stage at THP program commencement.

Although the interviews were conducted with participants after completion of a THP program, most of the participants spoke about how they made meaning in a different way before or during the THP program, and this can provide evidence of an earlier way of knowing. For example, someone in the initial sub-phase moving from one full constructive-developmental stage to another, can be seen as wondering about their way of being. Moreover, when a person begins to make meaning from a new constructive-developmental system, they may speak of a time when they did not see or do things in the new way. Furthermore, at the last sub-phase prior to making meaning fully from a new system, a person may have to work to avoid the prior way of knowing and can sometimes protest that earlier understanding. Where available, this kind of data was used to hypothesise about the constructive-developmental stage of the participant at the commencement of their THP program. For example, Holly described a time in her life when she didn't care about anyone else but herself:

I was lazy and not wanting to hear others. I didn't really care about it that much. ... I didn't want to meet new people that I didn't really care what they thought or what they are going through. [Just more focused] on me.

As an Instrumental knower, Holly was focused on herself and had little interest in others. However, Holly also described how at the beginning of the program she felt an unease with that way of being:

[W]henever I just shut them off or rolled my eyes, I'd feel this big anger inside of me I'm holding back, and the anger just doesn't come from the person, it just comes from like 'Why aren't you meeting them or getting to know them instead of just shutting them off?' and that was always in the back of my head, but I just didn't quite get the question. Like why would I

want to meet someone else when I want to focus on myself a bit more? But that kept replaying in my head, and then I just realised that's a bit selfish just focusing on you and not caring about others.

Here she is questioning her way of experiencing the world, but she is torn about being another way. At this time, Holly is not yet making meaning in a Socializing way, but she is on the precipice.

Based on the available data, it is hypothesised that at the time of their participation in a THP program, all participants were at an earlier developmental stage than at the time of the interview, with an average developmental shift of 1.6 sub-phases from program to interview. Based on this assessment, potentially nine of the 13 participants were making meaning from an Instrumental-dominant system at the time of their participation in a THP program (see Table 7.4 for complete results).

Constructive-developmental growth through program experience.

There is also evidence from the interview data that experiences in a THP program stimulated constructive development for some participants, demonstrated through a broadening of participants' meaning-making capacities which enabled them to understand their world differently. As someone moving toward a Self-Authoring way of knowing, Eric was able to reflect in a more complex manner on his experience of a THP program, describing the program as something that "doesn't force you to change" but rather "forces you to see a part where we don't think we know ourselves." Eric aptly characterises here the particular internal shift from Subject to Object that is an essential element of constructive-developmental growth. However, Eric is the only participant to have provided evidence in the interview of an explicit awareness of this change in his way of thinking.

Instrumental to Socializing. By continuing with Holly's description of the earlier time in her life when she had a more solid Instrumental way of knowing, it is possible to see the kind of internal shift Eric speaks about. At the beginning of her THP program, Holly was not yet making meaning in a Socializing way, but she was dissatisfied with her current way of knowing. This point at the juncture of a new meaning-making system is experienced as challenging, and it created for Holly

feelings of uncertainty, confusion, and anger. During the THP program, Holly said that she was challenged by her coaches to try to interact more with other people. Initially she resisted this task and lied to the coaches about it, but her lying caused her to “feel for like a couple of minutes a lot of guilt, and then afterwards I would have been like ‘Oh, whatever. [The coach] doesn’t even know. Who cares?’”

Instrumental knowers worry about the consequences of not following the rules or they try to figure out how to get past the rule if it is in their way. Her lying seems to be an attempt to get past the “homework” that has been set for her. The flickers of guilt, however, are indicative of more than just a wondering about her way of doing things. Her feelings of guilt reflect the beginnings of a movement toward a Socializing way of knowing. The Socializing stage is where conscience is born and the potential for guilt and shame arises, as well as the potential for empathy. Holly said that she eventually engaged with the challenge the coaches had set, and she described the process:

...step-by-step I got better at starting conversations and making the conversations last a bit longer than before. ... made me happier and a better person and I felt like I’m a better person than I was before and just like seeing them, knowing that you’re willing to listen to them ... it just makes me happy to see other people feel comfortable and willing to talk to you about their stories ... it’s good to do that cause then you gain more friends and a lot of trust.

She is both happy to be doing what she considers to be the right thing and also deriving her happiness from the happiness of others and from her developing relationships. When asked what the best thing for her was about gaining trust, she said:

[i]t’s like, for that person, to know you can tell them everything and they wouldn’t just walk away and forget it. They would mean like ‘Oh, I wouldn’t tell anyone about this’ and just feel more safe, and like someone knows about it and is aware that it’s happening and it’s not just kept inside of you.

While trust involves a relationship, the relationship here still seems to be in service of her own needs, reflecting an Instrumental system. However, when asked what

the best thing was about having a real story to tell her coach about the challenge they had set, she said:

It's seeing them know you've actually done it and tried to do it and is willing to take the challenge instead of just letting it slide and not wanting to change you. I guess seeing them know that you're trying to change you is a good thing as something they really wanted to see.

As Holly moves toward a more Socializing way of knowing, she is able to bring the coach's point of view inside of her and derive her own feelings as a consequence of seeing that point of view. What seems most important to Holly here is meeting her coach's expectations and getting the coach's approval for what she is doing, and this is consistent with a more Socializing way of knowing.

During the program, the coach met Holly where she was at, but also challenged her to put herself into situations that necessitated relating to others and supported her through the process. This undertaking extended into the Community Project where she spent time at a homeless shelter, an experience that seemed to impact her development further. She felt "lucky" to meet and get to know the people in the shelter "instead of just judging them and not caring." What was most important to Holly about having "good conversation and more knowledge" about the people in the shelter was that "knowing you're making someone happy, makes you happy." Here she can take a perspective on someone else taking a perspective on her, and she can derive her feelings based on how the other person is feeling. For Holly, supporting people at the homeless shelter was not about concrete consequences but about "building up love."

Dan provides another example of how the THP program influenced a broadened perspective-taking capacity from an Instrumental toward a Socializing way of knowing. In the interview, Dan describes a situation in which a teacher he and his friends had "given trouble," left the school. When the teacher left, his primary concern was not with how the teacher felt, but on the "opportunity we'd given up." He didn't realise "how good a teacher, how good a knowledge giver" they had until he left. At the time, Dan didn't consider the situation from the teacher's perspective because at the Instrumental stage the distance between his mind and his teacher's mind is too great. Dan was less concerned with how his teacher felt

and more concerned about the loss of his “knowledge giver”; the person who has the right answers and knows the right way to do things. However, the one-on-one coaching sessions in the THP program seem to have supported and encouraged in Dan a new way of making meaning about the situation. He described these sessions as having allowed him to make himself “a more understanding person.” These sessions gave him the opportunity to develop his perspective-taking capacity by putting himself “in the shoes of others and imagin[ing] how others see [him].” Whereas Instrumental knowers are all about their own immediate self-interest, Socializing knowers are more oriented to understanding other people’s feelings and judgement. Reflecting on his teacher’s departure and the way in which his behaviour may have contributed to that departure, Dan noted in the interview that if the situation were to occur now he would put himself “in the shoes of the other students to see what they felt and to even put [himself] in the shoes of the teacher.” He also reflected more broadly on a shift he noticed in himself away from self-interest to a focus that was more relational in its orientation:

Going through those [group] sessions I realised that maybe I need to think about the others as well because you hear people ‘Oh think about yourself, think about yourself.’ If you don’t think about others, sometimes there is no yourself because sometimes you’re put in an environment, a group environment specifically, when you have to work with others and sometimes you get the feeling that you can cut the tension with a knife and that’s why I felt that maybe I need to put myself in the shoes of others, realise what I’m doing that they don’t like, and find ways to correct that.

Dan shows here how he is able to consider how his actions might affect someone else beyond the implications for his own interests. By taking an external view of himself, Dan is now more fully experiencing himself as a function of how others experience him.

Working in groups also supported developmental growth for Instrumental knowers where these participants needed to interact and share their views. During the interview, Amy mentioned how before the THP program she “didn’t like talking to other people” or working in a group of people she did not know because it “was hard to communicate cause they might have different perspectives to what I have.”

However, she experienced times during the program where the group was “able to accept how I felt, which made me more happy and confident to talk about my feelings.” When asked what it was about the group acceptance that was most important, she said “it shows that it’s ok to have a different perspective, and we were able to talk about it.” Another Instrumental meaning-maker, Beth, also spoke about the value she felt in giving her opinions in the group sessions and having them accepted:

I felt good like [my opinions] were actually in use ‘cause they didn’t say it was a bad or good thing ... It felt really good ‘cause my opinion was accepted by everyone, so everyone agreed with it, and it was a good idea.

Although Beth’s meaning making is still anchored in the dualistic and concrete understanding of good versus bad, Beth is beginning to bring the views of others inside of her, consistent with a more Socializing way of knowing. However, these views of others function only as a source of information for Beth to meet her needs of having a “good idea” that “everyone agreed with.” Nevertheless, this represents the initial point of transition from an Instrumental to a Socializing system.

Developmental change for an Instrumental knower is also supported by listening to the views of others and having to consider multiple perspectives. When asked what was important about listening and understanding, an Instrumental knower, Ben, said, “I guess it makes you feel better in a way ‘cause you’re not always in control and wanting everything your way.” In each of these situations, the group experience provides a holding environment that challenges the participant’s current way of knowing and promotes a new way of understanding.

Socializing to Self-Authoring. While the shift from Instrumental knower to Socializing knower reflects a growing need for belonging and acceptance, the move from Socializing knower to Self-Authoring knower exhibits a movement toward autonomy and independence. Daisy reflected on this evolution in her development through her participation in a THP program:

[A]fter we finished the adventure excursions and even after the Community Project, I realised how self-independent I could be or become. I don't know, I was just kind of surprised by my own, how can I describe it, my capabilities to do things on my own. ... But I don't know, is that counterproductive since

the thing was kind of about teamwork? But, of course, I think I learnt what it meant to work in a team and what you really needed in order to have an efficient, cohesive unit, but also I think I learnt how much I could do just by myself as an individual, so I think that helped my self-esteem and my viewing of how I could do a lot of things. ... I think not only is it important for your self-esteem, knowing that you're capable of doing things, but especially in a unit, if you're unsure about your talents, then you can't really provide your part for the team. ... I think by knowing that you can do it, that doesn't just help you but helps the other [people in your] team because then they can rely on you, and you're confident in your abilities, and hopefully that will show in your work later on which will be better than when you are unsure.

Daisy still recognised and valued the importance of being part of a team, but she was no longer embedded in the team. She expressed here her feelings of a newfound sense of autonomy and self-direction, independent from her environment. She also valued being part of a team but recognised that she could make a greater contribution to that team by being more certain of who she is as an individual. For these participants, it is suggested that opportunities for independent contribution, including encouragement to develop and express their own ideas, would be supportive of constructive-developmental growth.

Conclusion. The interview data provides evidence of constructive-developmental growth for at least some of the participants during the course of their THP program. Those participants who were further away from the program and more complex in their meaning making were more self-reflective and better able to express their broadened capacities. Only one participant demonstrated an explicit awareness of this change in his way of thinking. While other participants may not have been as conscious of these shifts or only became so through the interview, there is still evidence of growth through their own words.

Coaching and group activities that involved appropriately supported challenges seemed to help some participants begin to make Object that to which they were previously Subject, assisting them to broaden their perspective and for

some, to see bigger possibilities for themselves than before. Such growth became apparent to some of the participants through the one-on-one and group coaching, which possibly challenged their current way of knowing and afforded them an opportunity to try out ways of understanding that were at their growing edge. Reflective conversations with coaches could often help participants notice their current ways of making meaning and consider ways of expanding their perspective. Furthermore, the group work in a THP program asked for mutuality and often led to dialogues with multiple perspectives and alternative possibilities that went beyond a single “right” way of doing something. As we heard from Amy, Alex, Ben, Beth, and Dan, having to interact and share within a group setting, to consider the thoughts and feelings of others, and to possibly subordinate their own needs and interests, were tasks at the growing edge for Instrumental knowers, and with appropriate support from their peers and coaches, these challenges could stimulate growth. For Socializing knowers, such as Eric, Cathy, and Daisy, group conflict without threat, leadership opportunities, and coaching support in establishing self-generated values, provided opportunities for increasing the breadth of their perspective. Eric sums it up well when talking about disputes he has had with his friend:

[W]hen we have a fight, at first it's a bit disheartening, but at the end of the day I think [the program] has given me that kind of perspective to kind of step aside from my position and his position and see it as a bystander, in a sense where I made a mistake and a sense where he made it, so once everyone's calmed down we'll talk about it then and everything's fine. So, I think ... it's given me a new perspective to see a situation. I think that's what it really is.

The ability to appreciate an experience as a “bystander” is a step in the development of one’s own Self-Authoring system.

Results of Research Question 3.4: Additional themes arising from participants’ accounts of their program experiences. Some themes arising from participants’ accounts of their experiences of the THP program were explored through the analysis of meaning-making capacity, including the experience of the

one-on-one and group coaching sessions, challenging program elements, and the Community Project. However, reanalysing the data more for its content than its structure allowed other themes to arise or reinforced existing themes. In particular, I was drawn to the emphasis participants placed on having the opportunity for novel and challenging experiences, and the impact these experiences had on a number of positive outcomes, particularly increased self-confidence, self-esteem, and resilience. Participants also stressed two elements that seemed to be necessary to the positive influence of challenging experiences: both support and multiple opportunities to meet a challenge. I will briefly touch on these themes and provide some passages from the interview data in order to lend colour to the themes.

Novelty, challenge and support. Many of the interview participants who participated in a THP program with an adventure component spoke about how they were doing things that “were new” (Amy) and that they “usually wouldn’t do every day” (Beth). They were “facing new challenges” (Charlie) that made them “step out of their comfort zones” (Fran; similar language was used by Amy, Beth, and Daniel). Some participants emphasised the novelty and enjoyment of the experience (e.g., “sailing out on sea that was something I never thought I would ever do in my life,” Charlie), while others stressed the value of being away from their normal lives (e.g., “I couldn’t even put it into words ... like even when I think about it, how simultaneously detached we were from the outside world ... like no internet, no service; it was just us,” Daisy), and for others the best thing was that it was “fun” (Alex) and how “it reminded [you] of childhood like just playing games in the backyard and stuff with your friends” (Cathy). For disadvantaged youth experiencing challenging life circumstances, just being away from day-to-day life and able to have fun seemed beneficial. Similar comments were not found in the interviews with the three participants who took part in a Coaching Only Program (without an adventure component).

Besides the novelty and enjoyment of the experiences, the participants spoke of the challenges they faced. Irrespective of experience level, Fran felt that there was a challenge for everyone.

Like you could be the most confident person, the most adventurous person.

There's gonna be one thing on this ship that you're gonna be like 'Oh hell no, I'm not going near that kind of thing.' Like it's really everywhere.

The Coaching Only participants also spoke of “personal challenges” (Holly) and “personal goals” (Grace) set through the coaching. Therefore, the adventure experience may not be the only way in which to introduce an element of challenge to the program.

Stepping out of their comfort zones was intimidating, but persisting through the program’s challenges enabled the participants to experience achievement.

The Helmsman Project created activities for us to step out of our comfort zone and do things that we usually don't do ... At first I felt really scared, and later when I got up [the mast], I felt relieved and happy that actually I was able to accomplish and face my fears. (Amy)

Emily also spoke about a challenge that was “scary” and how she was “able to get over [her fear].” For Emily, the best thing about that experience was “the fear going away ... and now it's like it looks scary, but it's a lot less scary than you think sort of thing, and I would just do it.”

Interview participants reflected on the multiple opportunities they had to meet a challenge. For some of the participants, the first time they were faced with a challenge, they were afraid. However, with time, they were able to conquer that fear. Initially, Amy said that she didn't want to do the program activities “cause I got shy,” but “then I got used to it, and then I kind of just did it and I enjoyed it.” Fran noted how the THP program “encompassed all [her] fears” and how she had to “overcome them in order to just do basic things.” Having two different adventure trips afforded her the opportunity to observe on the first trip, reflect, and then attempt to meet a challenge, despite her fears, on the second trip.

On the two-day [trip] I didn't even get off the deck. I was too afraid. But then on the five-day [trip] I managed to get up onto the second rung [of the mast], and I think I even tried the third one, but I was like 'no we're staying here' [laughs]. But knowing I could do that and knowing that ... this is going to sound really cheesy, but knowing that I could do it, like knowing that I have got the confidence and all that kind of stuff to actually do it ... yeah that just

made me feel so much better and so confident knowing that I could even get up to like the second rung for someone who's got such an intense fear of heights.

Emily had a similar experience with her fear of helming the smaller sailing yacht.

[O]n the five-day trip we were going out past the rocks, so like where the ocean's like really rough and stuff, and I remember the first time we were taking turns steering and I didn't know what I was doing, so I was really scared and I didn't want to do it, and so I didn't the first time. But when we were going our way back, I was able to do it, and it was actually a lot easier than I thought it would be and a lot more fun, so that's probably one of the things I'm most proud of is being able to do that.

Charlie also mentioned small things, like being able to put up his own hammock each night and packing it away in the morning. For the first few nights he asked the teacher to do it for him. He took the time to observe his teacher and other students doing the task, and then “on the last two nights I was able to do my own hammock, and it was awesome!” Daisy summarised these opportunities to make mistakes, noting “I guess in order to do something right, you have to do something wrong because then if you do something wrong, then you know how to do it right.”

As evidenced in the results for the constructive-developmental analysis, the role that the coaches played in the program and the ways in which they managed that role seemed to be an essential element in the balance of challenge and support. They were “relaxed and stuff” and “spoke to you as if you were just talking to them, not as if they were like a teacher or something like you had to learn” (Grace). While not acting in the role of authority figure could be challenging for some participants (as described in the constructive-developmental section above), doing so gave coaches the opportunity to develop a different kind of relationship with participants, one which may have allowed them to provide challenge and support in a new way. For Charlie, one of his “biggest fears” was public speaking and conquering that was also “one of my biggest achievements.” He attributed this achievement to his coaches who “helped me come to that hope and resilience to know that I can do it, and I can go out there and I can achieve that goal.” Another challenge for Charlie was meeting new people (a common fear expressed by other

participants). The worst thing, for him, about meeting new people was not knowing what to say. Charlie described how the coaches taught him about closed and open-ended questions and other strategies for making conversation. They also set him small challenges between coaching sessions that encouraged him to practice having these new conversations. Charlie summed up his experience, when he said, “it makes the whole program important to me to know that there was so much support, not just from crew, not just from peers and teachers, but from my coaches who were sailing aboard, and from the company itself.”

The support of the teachers, adventure facilitators, and other participants was also important. In response to the prompt “Moved and Touched,” Ben said, “I think the main thing was how everybody just supported each other on the ship ... like when somebody wasn’t feeling well or that kind of thing ... like people would make sure that they weren’t alone; someone would always be there.” Ben received this support, but he also experienced giving it and that “felt really good.” There was also support through modelling. Without this combination of support from peers, teachers, coaches, and other program facilitators, participants may not have been able to step outside their comfort zone and challenge their own perceptions of themselves and their capabilities. Beyond meeting challenges, this support afforded participants the opportunity to experience the care and understanding of others and, equally, to provide that emotional support to others. For Emily, “that was probably the first time I realised ‘Oh, like I can see that they are there for me,’ and that kind of stuck.” Moreover, because these were intact groups, this support could carry over into the day-to-day lives of participants. Charlie expressed how before the THP program, he and his teacher interacted “as teacher/student relationships do.” However, the “strong bond of friendship” they developed during the program enabled Charlie to ask the teacher for help when he was experiencing a difficult life situation, something he recognised he wouldn’t have been able to do before the program: “I would still be in the same situation if I didn’t reach out. I probably, honestly, probably wouldn’t even be here. So something as little as this has reached out to make a huge impact in my life.”

In their interviews many participants expressed being pushed outside their comfort zone, but with ample support and multiple opportunities to succeed, these

experiences enabled them to accomplish things they would not have thought possible before the program. It also appears that the fun, novelty, and excitement of the activities engaged the participants in a way that made stepping outside their comfort zones an intrinsically motivated process. It seems to have been this combination of elements that created a favourable holding environment for learning and growth.

Program outcomes. By being in unfamiliar territory and doing things they did not expect of themselves, many of the participants expressed developing greater positive self-beliefs, a primary outcome focus of the THP programs. For Ben, the THP program gave him “the confidence to do everything and not second-guess” himself. Eight of the interview participants expressed increased confidence as an important outcome of their participation in a THP program, using the word “confidence” positively in relation to their growth, a total of 31 times. Alex not only felt his “confidence started going up,” but he also recognised that this increased confidence allowed him “to take more initiative for other things ... after the Helmsman Project I’ve done lots more things that I don’t think I would have done if it wasn’t for the Helmsman Project.” Beth also was surprised by her newfound initiative.

You achieve something you didn't expect to achieve – ‘cause I did not expect to be SRC [Student Representative Council]. I always thought that I'll just be the student that's quiet and just finish high school this way, go after what happens after high school. But then I actually went from the quiet student to being a leader from like so unexpectedly, and it all happened after Helmsman too, so I'm very grateful to be in the Helmsman Project.

Emily came out of the program,

able to do more different things I was scared of previously that I didn't want to do, and I can just do them now and I'm not ... I'm still afraid, obviously, but not as afraid as I might have been, and I'm not so afraid that I'm unable to do it ... yeah, and just having more confidence and being able to get stuff done faster and not feeling as awkward about things and just being able to let bad things go.

Emily acknowledges how she has been able to take the opportunities she had in the THP program to meet and overcome her fears and apply these skills in her day-to-day life. Emily described not only self-confidence but also indicated how she could now “get stuff done faster” and “get out of a bad situation if something does go wrong,” perhaps reflecting a development in her self-efficacy and self-regulatory skills. She also spoke of her resilience now in “being able to let bad things go.” Self-efficacy, self-regulation, and resilience are equally key outcomes targeted by the THP programs.

When asked what the most important thing about the THP program was, Daisy said it was “realising your self-worth and capability as an individual and within a group.” Here Daisy seems to be making reference to the growth in her self-esteem and self-efficacy. Other participants also expressed a growing awareness of their own self-efficacy and resilience when they made mistakes or were unable to do something initially but ultimately achieved a positive outcome. For Ben, these experiences gave him the skills “to handle bad things that happen and not just get bogged down.” When asked how he might respond to a difficult situation now, he said, “I’d just think of ways to solve it and make sure those get executed.” The program was the “first big experience” that taught Alex “about how to deal with something [he] didn’t expect,” and now he doesn’t “judge [himself] so harshly anymore.” After overcoming her fears to climb the mast, Amy said that now in the face of new fears, “I would just try to face it and overcome because I know that by doing that I’m able to accomplish something.” Charlie’s experiences in the program taught him that when faced with something new, it is not simply a matter of assessing whether you can or cannot do it.

[P]eople either can do it or cannot do it or can give it a go ... and me, I don't want to be like a cannot person, I wanna be a definite yes, but there are times where I'm like 'I can't do it,' but I can give it a go, and if I give it a proper go, I guess it would push me more towards the can, like you can do it. For Charlie, the fear of not being able to do something right may no longer hold him back because he has seen that in the attempt he creates the possibility for success. For a number of participants, these increases in their own confidence,

efficacy, and resilience appear to have generalised beyond the program and into their day-to-day lives.

In addition to growth in general self-confidence, participants also mentioned being more confident specifically when meeting or having to talk to new people, and this was found to be helpful in forming new relationships. For Fran, the THP program really helped her,

become a lot more confident with just people in general where I can just go up to anyone whether it's at a party or just on the street or at a restaurant. I can just go up to someone and just start to talk to them ... and I can thoroughly thank the Helmsman Project for that and all the girls on the ship for giving me that confidence 'cause I didn't really have that before.

These outcomes reflect not only increased confidence but development in social competence as well. Charlie experienced a similar outcome from his time in the program.

Like my confidence 1 out of 10 would have to have been before the Helmsman Project, probably a good 4 or 3, and now it's like 9 and it's amazing like just to go up to someone and say 'Hi, I'm this person, nice to meet you' and getting that smile of a reaction back and saying 'Oh hi, I'm this person.' That's just amazing to know I can go up to someone and start a conversation and tell them a bit about me and talk to them about themselves.

Amy also found her confidence and ability to communicate improved from her experiences in the program.

During the Helmsman Project I've learnt to build more confidence, to speak up and talk to others ... by stepping out of my comfort zone, I [was] able to gain more confidence to be able to communicate with other people ... usually in class I will be shy, and now if I need help, I'm able to talk to the teacher and talk to other students ... I'm able to step up so I can help myself too ... I'm able to contribute with class discussions and that, and I'm not afraid to ask for help anymore.

These participants reflect an increased ability to not only handle new situations and new people, but to also communicate better in existing relationships with teachers

and peers at school and in a way that is helpful for their learning. Eight participants explicitly spoke of improved communication as a valued outcome from the program. When asked why communication was so important to her, Daisy responded:

Because I think if you understand each other and if you know how someone's feeling then you can be like more empathetic, and you can understand yow your group works and how you can make it better to get things done more efficiently.

Through increased confidence and improved communication, participants developed other important interpersonal skills, including the capacity to “listen and understand” others (Ben; similar comments made by Amy, Cathy, Eric, Grace, Dan, and Holly), put yourself in the “shoes of others” (Dan), “speak out and say [your] opinion” (Amy), “present ideas,” “put yourself forward,” and “come to a consensus within the group” (Dan).

Other valued outcomes participants spoke of included being more “open-minded” (Ben and Eric) and “flexible with other people's ideas” (Holly), improving their “self-control” (Fran), “goal-setting and achieving” (Alex), and there was the broader growth in perspective-taking that was highlighted in the constructive-developmental analysis above (see the results for Research Question 3.3). There also appeared to be a shift, for some participants, in how they felt about themselves or life more generally. The THP program on the whole made Emily “a lot happier as a person.” Fran also expressed a general sense of contentment derived from the program:

There's a song by the Beatles called *Octopus's Garden* where it makes me feel so happy, and I kind of made a place in my head called ‘Octopus's Garden’ where there's no wars, there's no famine, there's really no religion, no one argues over anything. It's just this really happy place, and if Octopus's Garden was real, the Helmsman Project with all these girls, that's the place to go.

General wellbeing is a secondary, but important, focus of the THP programs.

Moreover, Alex noted that the “Helmsman Project was one of the first big things I did in high school that really changed my character,” and he said, “I recognised that

I have stuff to offer.” Charlie’s increased confidence allowed him to envision a future for himself.

[T]o feel that the Helmsman Project has pushed me this far, got me out of my shoes, it’s amazing, and that’s helped me with my future goals, I guess, ‘cause a few years back I would have never thought about my future the way I do now. The way I think about my future now is like ‘I want to do this, this, this, this.’ I’ve got so many open doors. Back then, I was like ‘What am I gonna be doing with my future?’ because we would have these career advisors come in our classrooms and talk to us and I’d be like, ‘Yeh, I don’t know what to do.’

Seeing multiple possibilities for the future reflects increased hope, and this outcome is a strong focus of the THP program.

There was also an indication that these outcomes may not have been immediately obvious to the participants, with time for further experiences and opportunities for reflection possibly being an important element. Alex (whose interview occurred a year and a half after completing the program) noted that for him it was “not an immediate change” and was something he could only come to see “over the years, like reflecting on it now.” The interview also generated “interesting” reflection for Daisy who found the interview had “a lot of value” and was “really useful.”

Conclusion. Re-exploring the interview data for its content reaffirmed some themes that arose in the constructive-developmental analysis and allowed additional themes to arise, particularly around the beneficial effects of program participation and those elements of the program that seemed essential to the positive outcomes. In particular, novel and challenging activities, with multiple opportunities for success, came across as an important component of the program. Moreover, participants in both the Coaching Only Program and the Adventure Programs spoke of having these important opportunities for challenge. However, challenge needed to be balanced with ample support. As seen in the constructive-developmental section, understanding how participants make meaning of their experiences could inform the challenge/support dynamic. The coaches, teachers,

other program facilitators, and other participants provided different elements that in combination were important to achieving the right balance of challenge and support.

The outcomes from the THP program most often mentioned by interview participants included increased self-confidence, self-efficacy, and resilience, as well as improvement in self-regulation, communication skills, relationships and one's understanding of others, open-mindedness and perspective-taking, goal-setting and achievement, and general wellbeing.

Summary

The 13 interview participants for this study were assessed at six different constructive-developmental phases, demonstrating a range of ways of knowing. However, on average, participants were somewhere transitioning between the Instrumental and Socializing ways of knowing, with an Instrumental-dominant system hypothesised as being most prevalent during the program.

The interview data reflected how participants at similar constructive-developmental stages made meaning of their program experiences in similar ways, which was distinct from the meaning made by participants at different constructive-developmental stages. For participants making meaning in an Instrumental way, the focus was on their own needs, interests, and desires. Group work was at the growing edge for these participants. Beyond their immediate self-interest, such Instrumental knowers were concerned with understanding the rules and concrete consequences of their actions. Greater structure, including clear procedures and expectations, was supportive for Instrumental participants. On the other hand, participants at a Socializing stage were focused on their relationships and sought out acceptance and approval. For these participants, the group was central to their THP program experience, and structure and facilitation did not seem as essential other than to manage group conflict where difference felt threatening. However, the autonomy needed for self-regulation was at their growing edge. Understanding difference as non-threatening was also important for their continued development.

Both the one-on-one and group coaching sessions seemed often instrumental to participant development, and having both types of sessions could

be helpful in a group of knowers with different capacities for meaning making. However, the coaching sessions could also be challenging environments, with the particular challenges dependent on one's way of knowing. Support was found in the coaches but also in the other program facilitators and participants. Beyond the coaching, there was also evidence that the outdoor adventure experiences created a rich environment for growth. These experiences provided opportunities to try new things and to make mistakes from which they could recover. Again, participants at different constructive-developmental levels made sense of mistakes in varying ways. For Instrumental knowers this was often related to concerns about doing it wrong and the concrete consequences of their mistake, and for Socializing knowers the judgement of others could create unease.

The Community Project provided a particular challenge for many participants and may have been more suited to Socializing participants who were better able to self-direct the process. Given the value that Instrumental knowers place on having someone to show them the right way, the expectation for self-direction implicit in the Community Project could be experienced as a relinquishment of support. Consequently, the combination of challenge and support during this phase of the program may have lacked the right balance for Instrumental knowers. Although based on different understandings, failure to complete the Community Project could be disheartening for participants at any stage.

There was also evidence that the THP program effected a broadening of perspective-taking capacities for some of the participants, resulting in developmental growth. This growth was particularly demonstrated when the program met participants where they were at but also challenged them to bring into view some aspect of their way of knowing to which they had been Subject, while supporting them to a new way of understanding. While the fun and novelty of the program could create the intrinsic motivation to fuel growth, the coaches could provide essential scaffolding for the developmental process. Moreover, the peer group formed an important element of the holding environment.

Re-exploring the interview data for its content reinforced the importance of challenge to program outcomes, as well as appropriate support. While the

adventure experiences provided many challenging opportunities, Coaching Only participants also spoke of being challenged. In terms of program effects, participants revealed self-confidence, self-efficacy, and resilience as being primary outcomes they experienced for themselves, as well as effects on their relationships with others, and their overall wellbeing.

Discussion

Strengths

The depth of the data obtained from the 13, hour-long interviews conducted for this study provides valuable insight into participant experience of the THP program that is distinct from the evidence obtained through the quantitative methodologies of Study 2. The beneficial outcomes reported by participants support prior qualitative research findings in OAE, particularly in relation to improvements in self-confidence, resilience, and relationships with others (e.g., Cooley, Holland, Cumming, Novakovic, & Burns, 2014; Davidson, 2001; Duerden, Taniguchi, & Widmer, 2012; Ewert & Yoshino, 2011; Martin & Leberman, 2005; Sibthorp et al., 2008). Moreover, this appears to be the first time a developmental coaching intervention has been assessed with a constructive-developmental lens in the context of real-world challenges precipitated through OAE, thereby extending prior research in the fields of OAE and coaching psychology. There is only limited prior research that considers the OAE experiential learning process from this type of framework (Collins et al., 2012; Davidson, 2001; McGowan, 2016). Finally, while research in the areas of leadership and adult learning have used constructive-developmental theory, there appears to be little research that applies this theory to better understand adolescent meaning making.

The findings from this study demonstrate how the stage at which a participant makes meaning has an impact on the way that participant engages with the THP program. The range of different ways of knowing in these findings also supports the relevance of meaning-making capacity as an important form of individual difference among the adolescent participants in the THP program. Not only do they experience the same activities differently, the level of complexity with which participants understand their world can influence what from their program

they will experience as challenging and what they will find supportive. For example, participants at the Instrumental stage enjoyed the opportunities for one-on-one coaching, particularly where these sessions helped them to meet their needs. However, a participant making meaning predominantly at the Instrumental stage may find an expectation for self-reflection in one-on-one coaching, without appropriate scaffolding, to be stressful. On the other hand, participants making meaning toward the Self-Authoring stage might find a program that lays out precisely the right way to do things, lacks sufficient challenge. For participants making meaning in a Socializing way, the group coaching sessions were most enjoyable provided the differences among group members were not too great. For these participants, the one-on-one coaching could create an environment with too much risk of judgement and criticism. These findings suggest that participants in OAE programs that work from the level of the individual and are flexible enough to adapt the program experience, particularly to afford challenge and support that meets each participant where they are, may experience more beneficial program outcomes. More notably, there may be implications in these findings for the way in which OAE programs are designed and implemented.

Specifically, these findings demonstrate how program facilitation through developmentally trained coaches can provide an effective scaffold both to manage program challenges and to stimulate constructive-developmental growth. However, coaches need to be able to adapt their coaching method and style to accommodate these different ways of knowing, especially as they are likely to be distinct from the coach's own way of understanding. For example, the coaching literature suggests that effective coaching is about asking the right questions rather than telling someone what to do (see Stober & Grant 2006). Some of these coaching methods are suggested to emanate from principles of adult learning and an understanding of the coachee as someone who is autonomous, has a readiness to engage in reflective practice, and comes to coaching with a foundation of life experiences and knowledge from which they are able to generalise (Grant & Stober, 2006). As has been demonstrated in the findings from this study, adolescents are unlikely to possess these qualities (as, it is suggested, are some adults). Beyond any meaning-making gap, the interview data demonstrates the challenges inherent in

constructive-developmental growth, which may necessitate being able to not only see the world in a new way but to articulate it. There is a role for coaches in helping participants explicitly construct these new ways of seeing and talking about themselves and their experiences. However, doing so may be at odds with some coaching theory, such as the “ask, don’t tell” principle. Consequently, coaches may need to coach in a different way to how they have been trained. Importantly, they need to approach the coaching process from the understanding of the participant.

While flexible programming that meets individual participants at their capacity for meaning making is beneficial, doing so can be difficult. Developing an understanding of constructive-developmental theory and being in a position to establish the structure of an individual’s meaning making is a time-consuming and demanding process. However, simply appreciating the theory may create an awareness of individual difference in experience that can qualify program expectations and help participants to feel understood and supported for the ways in which they make meaning of their experiences. Moreover, given the similarities in how individuals at particular constructive-developmental stages make meaning, understanding those meaning-making systems can assist program providers in noticing what a participant is experiencing in the moment in order to rebalance the challenge and support in an OAE program, as well as to scaffold new ways of making meaning. A constructive-developmental lens, therefore, provides a valuable and novel frame of reference for OAE researchers and practitioners.

Finally, all interviews were conducted after the follow-up quantitative data was collected, with some interviews occurring one or two years later. The interview data, therefore, provides additional insight into the longer-term impacts of participation in a THP program and OAE more broadly. In particular, there was an indication that some changes may not be experienced immediately and may only come with further experiences that provide the opportunity to apply program learnings, as well as time and capacity for reflection. For some participants, the interview itself may have provided the catalyst for the participant to recognise the benefits they had derived from program participation. Consequently, it is suggested that future studies should continue to pursue longitudinal analysis that considers program outcomes a year or more after program completion. Furthermore,

incorporating a qualitative element to the methodology provides not only additional data on the program elements influencing program outcomes, but can also serve as an intervention in its own right by scaffolding participants in their meaning making development.

Limitations

Limitations of this study must also be taken into account. First, it is important to note that the schools were responsible for finding THP program participants who were willing to participate in the study. In addition, these participants were required to provide additional consent documentation and only those students who returned the documentation were able to participate in the study. Accordingly, although the study participants represent every THP program modality, some selection bias may have been introduced such that the study participants are not representative of the overall population of THP program participants. While such bias should not impact the constructive-developmental analysis (i.e. evidence as to the structure of participants' meaning-making), it may have affected the general content analysis.

Second, study design was dictated to some extent by financial and other logistical issues. Consequently, the data for this study did not include a comparison group, which potentially impacts the internal validity of some of the inferences drawn in this study, particularly in relation to program influence on constructive-developmental growth. Moreover, while the students were at a similar age at the time of participation in a THP program, they varied in age by up to three years at the time of interview. While age is not suggested to be a determinant of constructive-developmental stage, it can have an influence and may confound the results. However, no association was found between age and constructive-developmental stage. The lack of a pre-program SOI provides a further limitation on the assessment of the constructive-developmental stage of a participant at the time of participation in a THP program, as well as any growth in participant meaning-making stimulated through program participation. Nevertheless, it is suggested that these aspects of a participant's meaning making can be considered through the interview data.

Finally, when responding to the interview questions, participants may have been orienting to the particular demands of the interview, for example, the need to justify their, or their school's, participation in the THP program, the requirement to play the role of interviewee, or to respond to what they thought the interviewer wanted to hear (see Smith, 1996). However, it is suggested that the depth of the questioning that is a basic element of an SOI interview should distract from any such orientation. In fact, I was struck by how willing the participants were to share their stories, both positive and negative, freely and with great candour. It also seemed to surprise participants how the process of talking about their personal experiences through the interview was helpful for their own understanding of their program experience, including their growth and the other benefits they derived from that experience.

Conclusion

This study adopted a qualitative perspective using IPA and a constructive-developmental lens to examine how a subset of the THP program participants made sense of their program experiences. It is suggested that by having a deeper understanding of these participant experiences and the meaning-making systems from which participants approach the program, those who design and implement the THP programs will be able to provide a more effective holding environment for the growth of its participants. Beyond the THP programs, this method of applying a meaning-making perspective to OAE may enable OAE researchers and practitioners to consider matching program processes and expectations more closely to the developing capacities of program participants. Furthermore, insight from the interview data into the process of constructive-developmental growth may inform the role of coaching in scaffolding that process. Consequently, it is hoped that this study provides a useful resource for the fields of OAE and coaching psychology more broadly.

CHAPTER EIGHT

DISCUSSION AND CONCLUSION

Introduction

Both outdoor adventure education (OAE) and coaching have been used independently with adolescents to foster personal and social growth and the development of important life skills (see e.g., Green, Grant, et al., 2007; Green et al., 2006; Hattie et al., 1997; Marsh et al., 1986a). These outcomes are said to be important to educational achievement and life potential (Garcia, 2015; Gutman & Schoon, 2013; Heckman, 2006).

The Helmsman Project (THP) brings a novel extracurricular program to high school students located in areas of socioeconomic disadvantage. Integrating a series of structured developmental coaching sessions with outdoor adventure experiences, this program aims to positively foster participants' personal and social development by cultivating a range of qualities and skills. A further program objective is for participants to develop broader perspective-taking capacities, enabling them to perceive different and bigger possibilities than before. Through these outcomes, the ultimate goals of the program are to improve participants' educational engagement and wellbeing, thereby assisting participants to flourish and reach their full potential.

Both the coaching and outdoor adventure components of the THP program seek to integrate existing evidence-based research and practice from the fields of OAE and coaching psychology. While the research findings on the separate application of OAE and coaching have been largely positive, academics and practitioners in both areas have suggested that more robust research design and analysis is required, as well as a better understanding of the relationship between program variables and outcomes (e.g., Grant, 2012a, 2016b; Grant & Cavanagh, 2007; Grant & Cavanagh, 2011; Hans, 2000; Neill, 2008; Sheard & Golby, 2006).

The aim of this thesis was to investigate the effectiveness of the THP program on participants' personal and social development through a range of outcomes, using a methodologically robust design and statistically rigorous analysis. Three studies were undertaken to meet this aim:

- Study 1 sought to undertake a comprehensive psychometric evaluation of the outcome measures.
- Study 2 used a longitudinal, randomised controlled trial (RCT) to evaluate the short-term, long-term, and follow-up effects of the THP program. Study 2 also examined the differing effects of the three modalities through which the OAE element was provided. Moreover, an alternative program without the OAE component (referred to as the Coaching Only Program) was implemented for Study 2 in order to consider the incremental benefits of the OAE element for the THP program (the program with the OAE element referred to as the Adventure Program).
- Study 3 applied interpretative phenomenological analysis and a constructive-developmental lens to examine how 13 THP program participants understood their experiences of a THP program, as well as developmental differences in those understandings.

After briefly summarising the findings from the three studies, the findings of Study 2 and Study 3 are juxtaposed within the broader context of OAE and coaching research and practice. Next, some general strengths and limitations of the overall research are addressed, including directions for future studies. Finally, implications of the research findings for educational policy, as well as OAE design and practice, are elaborated on before providing some concluding remarks.

Summary of Findings: Studies 1, 2, and 3

Study 1

Study 1 was concerned with the rigorous investigation of the psychometric properties of the scales that were used to measure the outcomes in this research, in order to establish the robustness of those scales and the integrity of the data. Study 1 tested the internal consistency reliability, factor structure, measurement

invariance, and construct validity of each outcome scale. Study 1 produced the following results (see Table 5.3 for a high-level overview).

- **Internal consistency reliability:** Of the 41 measurement scales, 31 scales demonstrated internal consistency reliability with omega estimates greater than .70. Omega estimates between .60 and .70 were found for five of the remaining scales, which is not uncommon with psychological constructs such as those used in this research (Field et al., 2012). Moreover, all but three scales evidenced internal consistency reliability through acceptable inter-item correlations. Those scales that did not meet the thresholds for internal consistency reliability (Agency, Consistency of Interest, and Hampering) were nevertheless accepted based on their factor analytic properties.
- **Factor structure:** The a priori factor structure for each measurement instrument used in the measurement survey (Survey) had an appropriate fit to the data when modelled as either a confirmatory factor analysis (CFA) model or exploratory structural equation model (ESEM), with three exceptions:
 - The CFA for the Children's Hope Scale (CHS; Snyder et al., 1997) was modified to allow one Agency item to cross-load onto the Pathways Thinking factor, consistent with other research (cf. Edwards, Ong, et al., 2007), and this model was preferred to the ESEM.
 - Of the 36 items in the original Adolescent Self-Regulatory Scale (ASRI; Moilanen, 2007), 15 items were selected for a revised scale more targeted to the THP program goals (ASRI-R), and a three-factor ESEM fit this data well.
 - A three-factor model provided a better fit to the 11-item adapted Motivation and Engagement Scale (MES-S; see Martin, 2007; 2009) than the four-factor model originally proposed.
- **Measurement invariance:** Scalar measurement invariance over time was supported for all measurement instruments (with the CHS and SWLS achieving only partial scalar invariance).

- **Construct validity:** Multitrait-multimethod (MTMM) analysis evidenced general support for the convergent and discriminant validities, as well as test-retest stabilities, of the measurement scales.

Overall, the measurement scales demonstrated sound psychometric properties. Where measures did not initially evidence satisfactory results, theoretically supported modifications were made, and these adjusted scales showed acceptable psychometric properties. The results from Study 1 add to and extend existing validation research on the measurement instruments and provide a strong foundation for the further analyses in Study 2.

Study 2

Study 2 used a longitudinal, RCT to evaluate the short-term, long-term, and follow-up effects of the THP program on 41 different outcome measures. The analysis considered both the combined effects of the different Adventure Programs (taken together), as well as each Adventure Program mode individually. In addition, the effects of the Adventure Programs (taken together) were compared with the effects of the Coaching Only Program. Each THP program, irrespective of modality, was designed to foster a range of positive outcomes related to participants' personal and social development through the following key mechanisms: (a) an experiential learning framework that provides opportunities for experimentation and reflection; (b) experimentation through a structure of focused goal setting and goal striving; (c) developmental coaching designed to work from, and stimulate growth in, participants' perspective-taking capacities; and (d) an environment providing an effective balance of challenge and support. In broad terms, Study 2 found the following results:

- **Adventure Programs:** When compared to the control group, participants in the Adventure Programs experienced small-to-moderate significant positive short-term effects for 12 outcomes related to hope, positive global self-beliefs, open thinking, wellbeing, social effectiveness, aspects of self-regulation, certain life effectiveness skills, and reduction in negative behaviours and beliefs associated with motivation and engagement. In general, those scales considered more relevant to the program design and

aims were more affected. While these short-term effects were not found to maintain in the long-term analysis, the Adventure Program participants reported additional significant positive long-term effects for outcomes related to cooperative teamwork, positive behaviours associated with motivation and engagement, and positive self-concept regarding relationships with the opposite sex. Specific effect sizes and scale names can be found in Chapter Six. A comparison of these effects with reference to existing research is discussed later in this chapter (see the section headed “Overall Quantitative Program Outcomes” below), as are some of the most consistent positive outcomes (see the section headed “Consistent Program Outcomes Across Studies” below).

- **Individual Adventure Program modes:** Considering the individual program modes separately, the participants in the James Craig Adventure Program evidenced moderate-to-large significant positive short-term effects for more than half of the 41 outcome variables when compared to the control group. These significant effects were maintained longer term for outcomes related to optimism, wellbeing, social effectiveness, thoughts related to motivation and engagement, and certain life effectiveness skills. Moreover, there were additional significant positive long-term effects for outcomes related to cooperative teamwork, life satisfaction, and positive self-concept regarding relationships with the opposite sex. These results stand in contrast to the other program modes that evidenced only one or two significant short- and long-term effects. Moreover, these results are also distinct from the within-subjects analysis, in which the Arctos Adventure Program reported many more significant results than the other Adventure Programs. These diverse outcomes are address later in this chapter (see the section headed “Diverse Program Outcomes” below).
- **Coaching Only Program:** Participants in the Coaching Only Program, when compared to the control group, reported significant positive short-term effects in social effectiveness and academic self-concept related to English. However, they also reported significant negative short-term effects in grit

and locus of control (being more external). Consequently, the Adventure Program participants demonstrated significantly greater effects than the Coaching Only Program participants for a number of outcomes, including hope, cooperative teamwork, grit, locus of control, open thinking, and certain other life effectiveness skills. This result suggested that the adventure component may have an important role to play. The significant short-term positive outcome for academic self-concept and negative outcome for locus of control for the Coaching Only Program participants were maintained longer term, and new significant positive long-term outcomes were found for optimism and self-concept related to relationships with the opposite sex. The long-term effect for optimism for the Coaching Only Program participants was significantly greater than for the Adventure Program participants. Moreover, in the within-subjects analysis any significant differences found between the Adventure Programs and Coaching Only Program were evenly divided between the two programs, and two of the scales had opposite results in the RCT and within-subjects analyses. These results do not lend themselves to a clear answer on the question of the incremental benefits of the adventure component for the THP program.

- **Aptitude-treatment interaction effects:** Adventure Program participants with lower pre-test scores in outcomes related to hope, optimism, cooperative teamwork, and academic self-concept in English reported significantly higher scores at immediate post-test (when compared to their respective control group), than participants with higher pre-test scores. More particularly, the Outward Bound participants with lower pre-test scores in outcomes related to cooperative teamwork, optimism, wellbeing, and academic self-concept in math, and the Arctos participants with lower pre-test scores in outcomes related to cooperative teamwork and academic self-concept in English, reported significantly higher scores at immediate post-test. Relevantly, the Outward Bound and Arctos Adventure Programs did not evidence significant positive short-term effects for these outcomes. While there were some interaction effects that were not in the hypothesised

direction, almost half of the effects found that participants with lower baseline levels of an outcome showed more positive interactions.

Overall, Study 2 found a number of significant positive effects for participants in the Adventure Programs as a whole and some of the individual Adventure Programs, as well as the Coaching Only Program, although the results were not as consistent as had been expected. These results are considered further below in the context of the findings from Study 3.

Study 3

Study 3 aimed to examine the structure of the way in which 13 THP program participants understood their experiences of a THP program, as well as developmental differences in these understandings, using interpretative phenomenological analysis and a constructive-developmental lens. A secondary aim was to gather additional qualitatively rich data of the content of participants' experiences. The results from Study 3 can be summarised as follows:

- **Constructive-developmental assessment:** The interview participants were assessed at six different constructive-developmental phases, demonstrating a range of ways of knowing. Most participants were somewhere transitioning between the Instrumental and Socializing ways of knowing, with an Instrumental-dominant system hypothesised as being most prevalent during the program. For more information on constructive-developmental theory and these ways of knowing, refer to Chapter Seven.
- **Distinctions in meaning-making:** The interview data reflected how participants at similar constructive-developmental phases made meaning of their program experiences in similar ways, which was distinct from the meaning made by participants at other constructive-developmental phases. More specifically, participants making meaning in an Instrumental way demonstrated:
 - concrete and dualistic orientations to their experiences (e.g., good versus bad, right versus wrong);
 - an emphasis on fairness and equality;
 - a focus on their own needs, interests, and desires;

- value for other people who helped them meet their goals, with relationships having a more transactional nature;
- concern with knowing the one right way to do things and understanding the rules and concrete consequences of their actions; and
- difficulty holding more than one perspective at a time and reacting to different perspectives in an all-or-nothing way.

On the other hand, participants making meaning in a Socializing way demonstrated:

- an emphasis on their relationships;
 - feelings of support through a sense of acceptance and belonging;
 - a need for similarity with others, working to reduce differences;
 - internalisation of ideas, standards, and values from their peers and other valued people, while struggling to generate these things for themselves; and
 - self-esteem derived from the opinions of others, making criticism and judgement threatening, while praise and positive feedback could bolster their self-image.
- **Challenge and support for different constructive-developmental capacities:**
 - *Outdoor adventure experiences:* Beyond the coaching, there was also evidence that the outdoor adventure experiences created a rich environment for growth. These experiences provided opportunities to try new things and to make mistakes from which participants could recover. However, participants at different constructive-developmental levels made sense of mistakes in varying ways. For Instrumental knowers this was often related to concerns about doing things the wrong way and the concrete consequences of their mistake, and for Socializing knowers the judgement of others could create unease.

- *Self-reflection and abstraction:* Some participants were challenged by the self-reflection and abstraction inherent in the experiential learning process, making support essential for this element of the program.
- *Coaching:* Both the one-on-one and group coaching sessions seemed often influential in participant development and having both types of sessions could be helpful in a group of knowers with different capacities for meaning making. Instrumental knowers preferred the one-on-one coaching over the group session as it provided greater opportunity for meeting their needs. The Socializing knowers, however, preferred the group session as it provided opportunities for acceptance and belonging, while one-on-one coaching may have created possibilities for judgement that could feel threatening.
- *Other support:* Support was found in the coaches but also in the other program facilitators and participants. Instrumental knowers found support in greater structure, including clear procedures and expectations, as well as reinforcement that they were doing things in the right way. These meaning-makers understood success through the achievement of concrete tasks and goals (e.g., skill gain), and felt supported in this process by instruction, modelling, and repeated opportunities for practice. Socializing knowers valued care and emotional support over other kinds of support, and often required external validation to feel successful.
- *Community Project:* The Community Project (see Chapter Four) provided a particular challenge for many participants and may have been more suited to Socializing participants who were better able to self-direct the process. Given the value that Instrumental knowers place on having someone to show them the right way, the expectation for self-direction implicit in the Community Project often was experienced as a relinquishment of support. Consequently, the combination of challenge and support during this project may have lacked the right balance for Instrumental knowers. Although based on

different understandings, failure to complete the Community Project could be disheartening for participants at any stage, while success further bolstered participants' positive self-beliefs and other outcomes from the program.

- **Growth in constructive-developmental capacity:** There was also evidence that the THP program effected a broadening of meaning-making capacities for some of the participants, resulting in developmental growth. For example, while Instrumental knowers found working with others to be at their growing edge, facilitated group work could stimulate development by providing opportunities to interact and share within a group setting, consider the thoughts and feelings of others, and possibly subordinate their own needs to the needs of others. On the other hand, Socializing knowers found autonomy and self-regulation to be at their growing edge. Growth for these knowers could be supported by opportunities for non-threatening group conflict, leadership, self-generated values, and independent thinking. Constructive-developmental growth was particularly demonstrated when the program met participants where they were at but also challenged them to bring into view some aspect of their way of knowing to which they had been Subject, while supporting them to a new way of understanding.

The qualitative method in this study provided an analysis of the way in which a subset of participants described their experience of a THP program, and the meaning-making systems from which they approached the program. The insights derived from this study are reflected on further in the discussion below.

Summary

This section provided a high-level summary of the separate findings from the three research studies conducted for this thesis. In the next section, the findings from Study 2 and Study 3 are drawn together in order to provide a deeper and more holistic perspective of the effectiveness of the THP programs. Following that analysis, all three studies are considered in order to outline their strengths, limitations, and directions for future research, as well as implications of the findings for policy and practice.

Juxtaposing the Quantitative and Qualitative Findings

Introduction

The qualitative research in this thesis provides valuable data that can inform the quantitative findings. In this section, the findings from Study 2 and Study 3 are juxtaposed. First, the quantitative program outcomes are compared with the existing research. Next outstanding questions about the distinct and varied program outcomes from the quantitative analysis are explored through the qualitative findings. In particular, additional variables that potentially influence the THP program outcomes are discussed, with a focus on situational events, program delivery, group dynamics, post-program experiences, and individual constructive-developmental differences. Consideration is then given to those positive program outcomes that pervaded both the quantitative and qualitative findings: positive global self-beliefs, social competence, and perspective-taking capacity. Finally, insight from the qualitative data is examined in the context of certain key program elements to elucidate the relationships between those program elements, participants' meaning making, and program outcomes. Particular attention is paid to the effective balance of challenge and support as an essential element of program success, and the important role that meaning-making capacity plays in that balancing process.

Overall Quantitative Program Outcomes

Adventure Programs. Notwithstanding the diversity in some of the research results, this research suggests that the Adventure Programs result in some beneficial outcomes in the development of both intrapersonal and interpersonal skills and positive self-perceptions. The quantitative findings of a number of significant, small-to-moderate positive effects across a range of outcomes fits within the research and literature indicating OAE has a positive influence on various aspects of the personal and social development of adolescents (Bowen & Neill, 2013; Cason & Gillis, 1994; Ewert, 1987; Ewert et al., 2007; Hattie et al., 1997; Marsh et al., 1986a; Neill, 2008), including disadvantaged adolescents (Norton & Watt, 2014). Focusing on the Adventure Programs (taken together), there were significant short-term effects ranging from .169 (Verbal Self-Concept) to .337 (Wellbeing), with a

mean effect size for significant effects of .226 and an overall mean for all effects of .125. Although some longitudinal OAE research has found program outcomes to maintain over the longer term (Hattie et al., 1997; Marsh et al., 1986a; Norton & Watt, 2014), the quantitative results from this research did not demonstrate maintenance of these effects over the three month follow-up period. However, the longitudinal results from this research indicated some new significant effects. Moreover, the qualitative data suggests that participants did experience some long-lasting effects, which are discussed below (see the section headed “Juxtaposing the Quantitative and Qualitative Findings,” subsection “Consistent Program Outcomes across Studies”). It is possible that some of the post-program experiences and timing of data collection may have influenced the longitudinal results. Moreover, there is OAE and coaching research to suggest that some effects may take more time (Davidson, 2001; Duerden, Witt, et al., 2012), particularly given the deeply reflective nature of coaching (Spence et al., 2019).

Meta-analyses in OAE research have reported average short-term effect sizes that are somewhat larger than the significant short-term effects found in this research: .31 (Cason & Gillis, 1994) and .34 (Hattie et al., 1997). However, there are a number of distinctions that are relevant to any comparison between the effects from this research and the existing OAE meta-analytic research:

- While the Cason and Gillis (1994) meta-analysis involved studies conducted on an adolescent population, the studies included both normal adolescents and those from a clinical population. The clinical scales in this meta-analysis evidenced much greater effect sizes. This result is consistent with research that has found adventure therapy programs to have larger effects (e.g., mean $ES = .47$; Bowen & Neill, 2013). Cason and Gillis acknowledged that combining the effects from such diverse programs could be misleading and may have been the reason for the large variance in effect sizes (reflected in the standard deviation of .62). Although the population of study for the research in this thesis consisted of disadvantaged students, this classification refers to socioeconomic disadvantage rather than the clinical psychological disadvantage or delinquency that is more common in adventure therapy programs.

- The meta-analysis undertaken by Hattie et al. (1997) involved primarily adults and college-aged students. When the mean effect sizes from that study were broken down by age, they found the effects to be greater for adults (.38) than for college students (.21), which was similar to secondary school students (.18). The effect sizes from this research are more consistent with these mean effect sizes found for students.
- Those studies that were the subject of the meta-analyses primarily involved pre-post within-subjects analysis without a control group. An RCT design, which is the most rigorous form of measurement and the form used in the present investigation, is known to result in smaller-sized effects (Cheung & Slavin, 2016).
- These meta-analyses do not involve school-based OAE programs, with such programs considered to have a lower quality of programming and research (Hattie et al., 1997). In general, there is little empirical research on school-based OAE programs. This makes direct comparisons somewhat more tenuous. The effects of other types of educational interventions are varied. A meta-analytic review of school-based programs for developing students' social and emotional learning provides a point of comparison (noting that the studies primarily involved younger students). That synthesis of 213 programs involving 270,034 students found overall significant effects for social and emotional skills (.57), attitudes toward self and others (.23), and positive social behaviour (.24). It is suggested that the results from this research are comparable to the attitude and social behaviour outcomes from these alternative interventions. The skills results are not necessarily comparable as they were the primary focus of these programs and appear to have been expressly taught.
- Finally, it is important to recognise that the overall effects of the Adventure Programs (taken together) are impacted by the variation in findings across the individual Adventure Programs.

Coaching Only Program. The participants in the Coaching Only Program evidenced more significant negative, than positive, short-term effects. However, the

Coaching Only Program participants reported three outcomes with significant small-to-moderate long-term effects, and only the significant negative short-term effect for External Locus of Control was maintained longer term. Despite these results, the within-subjects analysis found nine significant positive moderate-to-large pre-post effects across a range of outcomes for the Coaching Only Program, suggesting that it may have a positive influence on the personal and social development of disadvantaged adolescents. There is limited coaching research with which to compare these results. However, the significant small positive long-term effect for Optimism and the significant within-subjects strong effects for both elements of hope (i.e., Agency and Pathways Thinking), are consistent with the results of an RCT study by Green, Grant, and Rynsaardt (2007) on a life coaching intervention for high school students. That study evidenced significant effects for both elements of hope, as well as cognitive hardiness, which were not evidenced for the control group.

Interestingly, the aptitude-treatment interaction analysis suggests that there may be some outcomes for which low baseline levels are challenging in a skills-based program such as the Coaching Only Program. In particular, participants in the Coaching Only Program who had low baseline levels of Grit and Focus and were more externally oriented in their locus of control, reported significant large negative effects in those outcome variables in the short- and long-term. On the other hand, Coaching Only Program participants with a low baseline level in Verbal Self-Concept reported significant large positive effects in that outcome variable in both the short- and long-term. Further research is needed to understand these outcomes.

The variation in results for the Adventure Programs and Coaching Only Program resulted in further discrepant results when comparing the effects of the Adventure Programs with the Coaching Only Program. While some of the positive program outcomes from the Adventure Programs were significantly stronger than those outcomes for the Coaching Only Program, most of these results did not maintain over the longer term. Additionally, the outcome for Optimism for the Coaching Only Program was significantly stronger than for the Adventure Programs in the long-term analysis. In the within-subjects analysis, each of the Adventure and Coaching Only Programs had three outcomes for which the program effects were

significantly stronger than the other program. Given the variability in these results, it is difficult to draw any firm conclusions on the incremental benefit of the adventure experience from these findings.

Summary. Although the THP programs evidenced some promising positive outcomes, there was variability in the findings. These results are consistent with the findings of Hattie et al. (1997) that only some adventure programs are effective, and then only with some participants and some instructors, and only on some outcomes, and probably only parts of the programs are influencing these outcomes. In the next section these findings are juxtaposed with the qualitative findings in an attempt to elucidate some of the variables that may be influencing the diversity in these outcomes. Consideration is then given to those outcomes for which the THP programs were consistently found most effective based on both the quantitative and qualitative results.

Diverse Program Outcomes

The research for this thesis involved rigorous research design and analysis that sought to meet the criticisms of existing research in the fields of OAE and coaching psychology. The THP program design was based on evidence-based research and best practice in those fields. Although the THP programs were offered through a number of distinct modes, each of the programs shared common goals and outcomes, as well as a similar program duration and design, including a small-group format, challenging experiences, goal-setting and goal-striving techniques, and an experiential learning framework. Moreover, each THP program was implemented by experienced coaches who underwent specific training for the program. The primary difference among the THP programs was in the existence or type of adventure experience provided and the size of the group that undertook the adventure component of the program together (i.e., Adventure Programs versus Coaching Only Program; Arctos Adventure versus James Craig Adventure versus Outward Bound Adventure). These program modes constituted the primary independent variables the subject of the quantitative analysis. A broad range of previously validated outcome measures were selected as indicators of the personal and social development of program participants. The program participants were of

similar age from schools located in a limited area of Sydney, Australia particularly affected by socioeconomic disadvantage, thus providing a natural control for some of the individual-level variables that might influence program outcomes. A number of other individual-level variables considered important to program outcomes were measured and controlled for in the analysis: gender, prior experience, socioeconomic status, and pre-existing flourishing (for more information on these covariates, refer to Chapter Six). An RCT was implemented with a waitlist control group in order to minimise other independent variables that might obscure the program effects (e.g., maturation, pre-program anxiety, etc.). Finally, the waitlist control group data was analysed using a pre-post within-subjects design as a test of replication of the results from the RCT analysis.

While the RCT analysis in Study 2 evidenced a number of significant positive outcomes for participants in the Adventure Programs, there were varying results among the three Adventure Program modes. The within-subjects analysis also reflected varying results among the three Adventure Program modes, however, these results diverged from the results in the RCT analysis. Similarly, the comparison of the Adventure Programs with the Coaching Only Program produced differing results between the RCT and within-subjects analyses. In an effort to understand whether these divergent results between the RCT and within-subjects analyses were a function of the distinct methods of analysis or the different program groups, a pre-post within-subjects analysis of the intervention group data also was undertaken. This analysis was reasonably consistent with the RCT analysis, suggesting that it was more likely to be group differences leading to the incongruous results.

The quantitative findings reinforce the complex nature of OAE and the challenges this environment provides for researchers (Ewert & McAvoy, 2000; Ewert & Sibthorp, 2009; Rowley, 1987; Scrutton & Beames, 2015). Understanding the various elements inherent in OAE and their impact, is necessary to the development of an evidence-base in OAE. Accordingly, it is important to recognise additional aspects of the THP program not already accounted for, that potentially make it more difficult to clearly establish the positive program effects on the outcome variables. The findings from the qualitative analysis highlight some of

these additional elements, providing context for the quantitative results. In particular, distinct situational events, program delivery, group dynamics, and post-program experiences, as well as individual differences in constructive-developmental capacity, may have influenced the quantitative findings on program outcomes.

Situational events. Unanticipated events that occur during an OAE program can have a beneficial or detrimental effect on program participants and, therefore, can influence program outcomes (Ewert & Sibthorp, 2009). Some examples of such events that occurred in a THP program (inferred through the qualitative interview data and THP feedback) included inclement weather experienced at sea on the Arctos, inclement weather that prevented the James Craig from leaving the dock, an Outward Bound facilitator getting lost on a hiking trip, one or more students failing to show for an adventure experience, and a change in coaching staff during the course of a program. It is also relevant to note that as there could be five times as many participants on an adventure experience for the James Craig program as there were for the Arctos program, such an event (whether positive or negative) might be reflected in the data more widely for a program like the James Craig program than for the Arctos program. In Study 2, we had not expressly gathered information on such events and, therefore, were unable to consider this element in the analysis. However, these group size differences applied only to the adventure components of the program and not to the program as a whole. Moreover, the adventure components were not temporally close to data collection, being embedded in the middle of the three-month program. Irrespective, the qualitative data found that some of these experiences remained firmly in participants' memories. Furthermore, as evidenced in Study 3, participants' ways of knowing influenced how participants experienced these events. Depending on the timing and influence of such events, they may have influenced program outcomes for participants.

Program delivery. Research has found that the level of implementation in prevention and promotion programs affects program outcomes (Durlak & Dupre, 2008). However, Durlak and Dupre (2008) also acknowledge the benefits of

adaptation for outcomes, indicating that what is important is to be aware of the adaptations. There are a number of aspects of program delivery for the THP programs that may have affected program outcomes, including possible improvement in program design and coach training over time, the different times of year programs were offered, and potential differences in levels of coach experience.

Timing of delivery. Program delivery improved over time, however, data were being collected from the first implementation of the THP programs. The Arctos Adventure Program was the only initial THP program offering, and it was run more predominantly in the initial year of the THP program than more recently. Moreover, participants in later offerings of the Arctos Adventure Program were control group participants undergoing a THP program. This may explain why the Arctos Adventure Program demonstrated many more significant results in the within-subjects analysis than in the RCT analysis. For specific details on the dates and modalities of the THP programs, see Appendix K. It is also relevant to note that the THP programs were run twice a year, with the first round commencing near the beginning of the school year (summertime in Australia) and the second round commencing in the middle of the school year (winter in Australia). Accordingly, programs were experienced by participants at different times of the year. These timing differences may have influenced variability in program outcomes.

Coach training. The THP program coaches engaged in supervision meetings, and these meetings provided data on the coaching process. Consistent with my own experience, other coaches noticed some variation in how participants were making meaning of, and interacting with, the various components of the THP program. The coach training that preceded each THP program, was enhanced through the feedback from these supervision sessions. In particular, later training incorporated additional education in constructive-developmental theory and was able to use examples from previous programs as a way to explore the theory in more detail. This may have resulted in a more effective coaching process and program experience for participants in the THP programs delivered later in the research.

Coach experience. Related to program delivery and coach training, is the level of experience of the coaches. Implementation of the program and coach

effectiveness is likely to be increased as experience is gained and experienced coaches retained. The coaching in the program was largely a voluntary position. Moreover, the coaching role involved a commitment of approximately 200 hours for a single THP program. Consequently, it was difficult to retain a consistent coaching staff. The low level of coach retention is one factor to be considered in interpreting these results. Additionally, general variance in coaching experience may have created differences in coaching effectiveness across programs. Coaching effectiveness is not a variable that was measured in this research. Moreover, given the number of coaches and nature of the coaching role (joint and individual), the particular coach assigned to a participant was not a cluster variable that could be accounted for in the statistical analysis. However, given the important role that program facilitators are said to play in OAE (e.g., Ewert & McAvoy, 2000; Gass & Gillis, 1995; Hattie et al., 1997; Luckner & Nadler, 1997; Martin & Leberman, 2005; Martin & Legg, 2002; Sibthorp et al., 2011; Sibthorp et al., 2007), coach effectiveness may have been a relevant variable in participant outcomes. A number of interview participants made strong connections between the coaching and their positive outcomes. For example, one participant said that the most important thing to him from his experience of the program was “building that network with [my] coaches, being so comfortable to talk with my coaches, and having those [coaching sessions] ... and still is important to me” (Charlie). This result is also consistent with existing OAE research that emphasises the important role that positive adult relationships and support play in adolescents’ developmental endeavours (Norton & Watt, 2014). Consequently, the manner of program delivery and effectiveness of coaches across programs may have influenced program outcomes for participants.

Group dynamics. The literature suggests that group dynamics and group cohesion are an important element of influence on participant outcomes in OAE programs (e.g., Carpenter & Harper, 2016; Davidson et al., 2016; Ewert & McAvoy, 2000; Jostad, Sibthorp, Pohja, & Gookin, 2015; McKenzie, 2003; Sibthorp et al., 2011; Sibthorp et al., 2007). Positive social experiences in OAE can promote positive adolescent development, while negative social experiences can lead to feelings of isolation and abandonment with potentially devastating effects (Jostad et al., 2015).

For example, Neil and Dias (2001) found a positive relationship between perceived social support and growth in resilience during an Outward Bound course. Having found that the perceived support from the least supportive group member was the best predictor of growth in resilience, they warned that negative group members could potentially retard the growth of other group members (Neill & Dias, 2001). The OAE group context, however, is a complex area to account for in research (Jostad et al., 2015). Although it has been suggested that groups that are more homogeneous may have greater group cohesion (Jostad, Paisley, Sibthorp, & Gookin, 2013), little is known about the social context that emerges in OAE environments (Mirkin & Middleton, 2014).

Initial grouping. The THP programs were offered through schools and, therefore, brought together intact groups from the same year at the same school (referred to as program *cohorts*). However, the qualitative interview data suggests that the group context varied: some program cohorts consisted of individuals who knew each other well and were friends outside of school; other cohorts consisted of one or more cliques with one or more individuals who were outside the clique; and still other cohorts consisted of individuals who knew each other but were outside of each other's friendship groups in their day-to-day lives.

Group experience. Beyond the initial grouping, each program cohort experienced different group dynamics. The interview data indicates that some groups going into a program as friends came out with even stronger friendships, while other groups experienced a falling out between some pre-existing friends. On the other hand, one participant who went into the program having felt "alone" most of her life, found the group experience a positive one that made her "feel less of an outcast" (Fran). The interview data also reflected how some groups worked together seamlessly, while other groups had one or two members that did not seem to integrate with the rest of the group, and yet other groups appeared to have difficulty sharing the workload among them, with some members described as "lazy" (Ben) or "not up to my standards" (Alex). There were also descriptions of group members who experienced homesickness or who were perceived by other group members as receiving special treatment.

Moreover, the qualitative data also demonstrated how participants within a group could make meaning of the group dynamics in different ways. Therefore, there seems to be no single way to ensure a positive group experience as some research may suggest (e.g., the program facilitator as social engineer; Mirkin & Middleton, 2014). It is also relevant to note that while the THP program involves group work, there is also an individual component to the program. Nevertheless, it is suggested that group dynamics, and how they were managed, had the ability to influence program outcomes for participants.

Post-program experiences. While some adolescents who experience an OAE program can extend the learning from that experience into their lives, others face challenges when the program ends, including difficulty processing the experience, feelings of loss, and a sense of isolation (Allison, Davis-Berman, & Berman, 2012). Unlike stand-alone OAE programs (e.g., Outward Bound, National Outdoor Leadership School), the THP program was provided through schools using intact school groups with the expectation that experiencing an OAE program with students from the same school would assist with transfer of learning and create important connections within and to the school that would be beneficial post-program (see Richmond et al., 2018; Sibthorp & Jostad, 2014). However, the qualitative data reflected THP program participants with different experiences of post-program cohort relationships, as well as post-program opportunities for learning transfer.

Post-program relationships. The interview data showed that some groups maintained their program group relationships upon returning to school, but in other groups these relationships dissipated. Readapting to school life after the program could be emotional (e.g., “I felt sad that [our group] couldn’t capture this kind of [bonded] feeling back when we were on land,” Daisy). The interview data also demonstrated beneficial effects of program connections that were maintained when the participants returned to school. One interview participant mentioned the important role the relationship with his cohort teacher played in his future (see Chapter Seven, “Results of Research Question 3.4: Additional themes arising from participants’ accounts of their program experiences”). However, there did not

appear to be any formal structure within the school context that assisted these groups to sustain their relationships and shared experiences. Richmond et al. (2018) have suggested that shared narratives of challenge and achievement allow schools to reinforce program outcomes and build upon OAE experiences.

Post-program project. The desire to connect the THP program back to the school context and the participants' broader community was promoted by THP through the introduction of a Community Project. However, this project was not available to some of the earlier THP programs, creating a difference in program experience. Moreover, as the Community Project took place after the THP program had concluded, the school played a central role in ensuring the challenge of the Community Project was met with ample support. The interview data suggests that where the scaffolding provided within the THP program was lacking in the school environment, some program participants struggled. On the other hand, other program participants were bolstered by the value their school placed on the Community Project and the support provided to assist project completion (whether by the school or, in some cases, through the group itself).

Other post-program opportunities. Beyond the Community Project, THP has continued to implement additional program elements in order to provide program participants with ongoing opportunities to reinforce program learnings, maintain relationships, engage in further development, and contribute to the community. In particular, THP coordinates an alumni program and has recently established a youth board comprised of a select group of program graduates. The youth board will shadow THP's board of directors and will also take a lead role in developing and coordinating initiatives to engage THP program alumni and other stakeholders. Some participants may have been engaged in these programs, although the youth board was established after most of the data for this research had been collected.

These post-program experiences, whether positive or negative, may have influenced longer term program outcomes for participants. In particular, this may have been more likely in cases where the immediate post-program or follow-up data collection occurred in close temporal proximity to these experiences.

Individual developmental differences. Traditionally, the focus on individual difference in OAE research has been on variables such as age, gender, and other demographic factors (McKenzie, 2000; Neill, 2007; cf., Collins, Paisley, Sibthorp, & Gookin, 2012). Nevertheless, there may be other individual differences that influence program outcomes (McKenzie, 2000). One such individual difference relates to the developmental characteristics of program participants (Collins et al., 2012). Participants come to a THP program with distinct differences in their cognitive, physical, social, and emotional development (Gilbertson, Bates, McLaughlin, & Ewert, 2006). These differences may result in participants experiencing the same program in diverse ways. Consequently, it has been suggested that individual program experience deserves greater attention for its explanatory power in the variability of OAE outcomes (Ewert & Sibthorp, 2009; McKenzie, 2000; Neill, 2008).

One aspect of this developmental difference relates to the ways in which participants construct meaning of experience and the development of those meaning-making capacities. This individual-level variable was the subject of Study 3, and the results of that study provide insight into the influence of meaning-making capacity on participant experience of a THP program and program outcomes. In particular, the interview data evidenced a relationship between a participant's constructive-developmental level and his or her understanding of various elements of the THP program. More specifically, individual meaning-making capacity appeared to influence the ways in which participants interacted with the THP program coaches, their peer group, aspects of the program delivery, situational events, and their lives post-program, resulting in distinct program experiences and outcomes. Accordingly, this individual developmental difference may be a critical element in the relationship between program participation and program outcomes. Implications of this individual difference for OAE program design and practice are discussed below (see the section headed "Implications for Educational Policy and OAE Design and Practice").

Summary. The qualitative data revealed aspects of the THP programs not accounted for in the quantitative analysis, which may explain some of the distinct

and varied program outcomes more than inherent differences in the program modes. This section highlighted how unexpected situational events, as well as potential differences in program delivery, group dynamics, post-program experiences, and individual constructive-developmental capacities may have influenced program outcomes, providing additional context for the quantitative findings. Understanding these additional program elements is important for the development of an evidence-base in OAE. The next section considers the overall positive program effects in comparison to prior research findings, following which consideration is given to those outcomes that were most consistently found across both the quantitative and qualitative analyses.

Consistent Program Outcomes across Studies

While not wanting to detract from the full breadth of positive outcomes this research has found to be associated with participation in a THP program, this section highlights a few of the key outcomes that demonstrated some positive, consistent results across the quantitative and qualitative analyses, and, therefore, deserve special attention. These outcomes broadly relate to increases in certain global positive self-beliefs, social effectiveness, and perspective-taking capacity. Results across the quantitative and qualitative analyses for these outcomes are discussed below. Following this review of some of the consistent positive program outcomes, the lack of significant outcomes for Goal Self-Regulation, a scale rated as highly relevant to the THP program, is discussed.

Global positive self-beliefs. Positive self-beliefs is an umbrella term used to capture a range of positive psychological constructs focused on competency self-perceptions and related self-beliefs (Marsh et al., 2017). Such self-beliefs are held to fundamentally influence a person's success or failure in school (Pajares & Schunk, 2001; Stankov et al., 2013). While many positive self-beliefs are constructed as domain- and task-specific, more generalised concepts do exist (e.g., self-efficacy and self-esteem/self-concept: see Chapter Three for more detail). Outcomes included in this thesis that fall under the umbrella of global positive self-beliefs include self-esteem, self-efficacy, self-confidence, hope, optimism, locus of control, and resilience. Self-efficacy and self-concept are said to be two of the most widely used

and theoretically significant of these self-beliefs (Marsh et al., 2019). They are also two of the most commonly measured outcomes in OAE research that also demonstrate positive effects (e.g., Barrett & Greenaway, 1995; Bowen & Neill, 2013; Crompton & Sellar, 1981; Hattie et al., 1997; Marsh et al., 1986a, 1986b; Neill, 2008). Growth in self-confidence is often cited as an outcome of OAE, although it appears to be more commonly reported in qualitative studies (e.g., Davidson, 2001; Duerden, Taniguchi, et al., 2012; Richmond et al., 2018; Stott et al., 2015; cf., Neill, 2008). Increases in hope and optimism as outcome measures are more common in coaching research (e.g., Green, Grant, et al., 2007; Green et al., 2006; Madden et al., 2011), while resilience is an outcome that has been found to be positively influenced through both OAE and coaching interventions (e.g., Beightol et al., 2009; Bowen, Neill, & Crisp, 2016; Clough & Strycharczyk, 2015; Ewert & Yoshino, 2011; Grant et al., 2009; Hayhurst et al., 2015; Kelly, 2019; Neill & Dias, 2001; Ungar et al., 2005; Whittington & Aspelmeier, 2018). The theoretical features that distinguish these constructs have been detailed in Chapter Three.

Quantitative outcomes. Across the quantitative analysis, the measurement scales for Agency and Pathways Thinking (together Hope), General Self-Esteem, Self-Confidence, Self-Efficacy, and Life Resilience demonstrated significant positive effects that were echoed in the qualitative interview data. The short-term effects for these scales are set out in Table 8.1. Details on these measurement scales can be found in Chapter Three under the sections headed “Hope,” “Resilience,” “Self-Concept,” and “Life Effectiveness.” There is concern that the more generalised the measures are, the less distinct may be what they are measuring (referred to as a *jangle fallacy*; see Marsh et al., 2019). The results of the MTMM analysis that found some violations of discriminant validity for these scales (other than Life Resilience), lend some support to this concern. More research is required in this regard. Nevertheless, the constructs appear to have been referenced by the interview participants in distinct ways.

Table 8.1
Short-Term Effect Sizes for Global Positive Self-Belief Outcomes

Program/Analysis	Global Positive Self-Belief Effects: ES (SE)					
	Agency	Pathways	Gen Self-Esteem	Self-Confidence	Self-Efficacy	Life Resilience
<i>Adventure Programs</i>						
RCT	.266 (.087)**	.247 (.093)**	.084 (.112)	.215 (.100)*	.226 (.088)**	.057 (.084)
Within-Subjects	.207 (.064)***	.114 (.070)	.199 (.111)	.184 (.091)*	.198 (.088)*	.230 (.090)*
<i>Arctos Adventure Program</i>						
RCT	.178 (.139)	.119 (.142)	-.081 (.163)	.049 (.160)	.134 (.139)	-.005 (.122)
Within-Subjects	.376 (.105)***	.256 (.106)*	.328 (.169)	.342 (.132)**	.211 (.134)	.396 (.118)***
<i>JC Adventure Program</i>						
RCT	.275 (.124)*	.490 (.128)***	.309 (.145)*	.541 (.114)***	.406 (.120)***	.326 (.106)**
Within-Subjects	.194 (.082)*	.111 (.116)	.297 (.190)	.234 (.130)	.213 (.120)	.125 (.127)
<i>OB Adventure Program</i>						
RCT	.344 (.164)*	.131 (.169)	.023 (.183)	.053 (.142)	.139 (.085)	-.148 (.133)
Within-Subjects	.051 (.104)	-.025 (.121)	-.026 (.136)	-.022 (.139)	.169 (.129)	.168 (.165)
<i>Coaching Only Program</i>						
RCT	-.120 (.162)	-.082 (.156)	-.065 (.176)	.069 (.133)	.104 (.242)	-.182 (.192)
Within-Subjects	.567 (.251)*	.627 (.188)***	.199 (.200)	.284 (.240)	.574 (.242)*	.706 (.203)***

Note. ES = standardised effect size; SE = standard error; Pathways = Pathways Thinking; Gen = General; RCT = randomised controlled trial analysis results; Within-Subjects = waitlist control group data within-subjects pre-post analysis results; JC = James Craig; OB = Outward Bound. * p < .05; ** p < .01; *** p < .001.

Adventure Program participants reported significant positive short-term effects when compared to the control group, in Agency, Pathways Thinking, Self-Confidence, and Self-Efficacy. For the individual Adventure Programs, the James Craig Adventure Program participants reported significant positive short-term effects when compared to the control group, in Agency, Pathways Thinking, General Self-Esteem, Self-Confidence, Self-Efficacy, and Life Resilience. The Outward Bound Adventure Program participants also reported significant positive short-term effects for Agency when compared to the control group.

The within-subjects analysis supported the RCT results for the Adventure Program participants in Agency, Self-Confidence, and Self-Efficacy, with Life Resilience also demonstrating significant pre-post effects. Across the individual Adventure Programs, participants reported significant pre-post effects for Agency (Arctos and James Craig), Pathways Thinking (Arctos), Self-Confidence (Arctos), and Life Resilience (Arctos). Participants in the Coaching Only Program reported significant pre-post effects in Agency, Pathways Thinking, Self-Efficacy, and Life Resilience.

While none of these scales evidenced a significant long-term effect in the RCT analysis, none of them evidenced a significant decline over the follow-up period (from immediate post-test to three months post-test). Moreover, the outcomes measured by these scales were some of the most widely referenced outcomes in the qualitative interviews.

Qualitative outcomes. Eight of the 13 interview participants (from both the Adventure and Coaching Only Programs) explicitly used the word “confidence” to describe the positive change they experienced in their self-beliefs through participation in a THP program. Interview participants spoke about successfully meeting challenges, both physical and social, and how doing so provided a sense of accomplishment and changed the way that they saw themselves. By doing new things that were out of her comfort zone, Amy saw that she has “different skills” and “can improve in those skills and learn something different.” Another participant noted how prior to the THP program, he felt that he was “not worth anything,” that he was “just a failure,” and maybe he should “just drop out [of school]” (Alex). However, participation in the program helped him to build his self-esteem through “goal setting and achieving,” and to recognise that he “has stuff to offer.” Even for a participant who wasn’t “someone who has low self-esteem,” recognising that her “self-worth and capability” was greater than before meant that she could “do more,” that “you’re not useless; you can impact other people,” and “knowing that you have that influence and that impact hopefully motivates [you]” to use your capabilities (Daisy).

Participants described having “the confidence to do everything and not second guess” themselves after the program (Ben). Although the positive self-belief scales did not demonstrate significant long-term effects in the RCT analysis, the interview data does suggest participants maintained these positive self-beliefs beyond the program, and that such feelings had flow-on effects into other aspects of their lives. Interview participants described how their newfound confidence made them able “to take more initiative for other things” (Alex) and to do things they were “scared of previously” (Emily), particularly within their school and community. A number of these participants described how after the THP program they went on

to leadership positions within their school and were able to envision a hopeful future for themselves that they previously had been unable to see. This transfer of the positive self-beliefs gained by participants during the THP program to their everyday lives is consistent with findings from other studies (e.g., Duerden, Taniguchi, et al., 2012; Paxton & McAvoy, 2000).

Beyond increased confidence in their abilities to succeed, interview participants spoke about how the THP program experiences showed them that they could handle more than they expected, demonstrating increased self-efficacy. More than just a “can-do” attitude, they developed a “can-give-it-a-go” attitude (Charlie). They learned skills that would help them face a difficult situation and “overcome” (Amy), by “thinking of ways to solve it and make sure those get executed” (Ben). They also realised that “by doing, [they’re] able to accomplish something” (Amy). Interview participants also referenced their resilience when they spoke of learning to “handle bad things that happen and not just get bogged down” (Ben). It is suggested that opportunities to make mistakes and recover, together with supported reflection, may have facilitated future resilience. Moreover, learning to deal with the unexpected helped participants not to “judge [themselves] so harshly anymore” (Alex). While Bandura suggests that people with self-efficacy are more resilient (see also Ewert & Yoshino, 2011), it may be that resilience promotes self-efficacious beliefs or perhaps there is a reciprocal relationship between the two.

Bandura (1997a, 1997b) describes four sources of self-efficacy beliefs: (a) experiences of success or mastery in challenging tasks; (b) experiences of social modelling of success; (c) social persuasion to believe in oneself, including direct feedback; and (d) building physical strength, which signals personal capability and reduces stress, anxiety, and depression. Interview data demonstrates how the THP program experiences gave participants opportunities to meet challenges with ample support, and how these experiences appear to have contributed to improved self-efficacy for dealing with future challenges. It is suggested that the use of goal-setting and goal-striving as an explicit element of the program model provided participants with many opportunities to achieve proximal, concrete challenges relevant to the development of self-efficacy. Moreover, the coaching element of the program offered support for reflecting on those achievements. Many of the

interview participants also spoke to their experiences of social modelling and social persuasion from peers, teachers, coaches, and other program facilitators. Finally, some of the adventure experiences in the program may have assisted participants to build physical strength, providing them with additional evidence of their personal capabilities.

All of these results are consistent with existing studies in the fields of OAE and coaching that have found both OAE and coaching to be effective for improving various global positive self-beliefs.

Social competence. Beyond the global positive self-beliefs referred to above, THP program participants also evidenced strong, positive, and consistent improvement across the quantitative and qualitative analyses in specific self-perceptions of social competence. These findings are consistent with previous research that has found OAE programs to influence improvements across a range of social developmental outcomes, including communication, cooperation, teamwork, leadership, and relationships (e.g., Bowen & Neill, 2013; Cooley et al., 2014; Hattie et al., 1997; Mirkin & Middleton, 2014; Neill, 2008; Richmond et al., 2018; Whittington, 2011). OAE literature suggests that engaging in small group activities that require support, teamwork, and communication provides opportunities for positive social interactions that can help to develop various social skills (Jostad et al., 2015). Furthermore, it has been suggested that OAE experiences with intact school groups may provide additional opportunities for enhancing some of these social competencies (Richmond et al., 2018). Moreover, developing social competence can affect a “positive orientation toward the social world that sets in motion adaptive beliefs and behaviours that facilitate adjustment in a variety of contexts” (Mirkin & Middleton, 2014, p. 234). Outcomes included in this thesis that are related to social competence include self-perceptions of social effectiveness, cooperative teamwork, leadership ability, and self-concepts of various relationships.

Quantitative outcomes. Across the quantitative analysis, the measurement scale for Social Effectiveness evidenced significant positive short-term effects that were echoed in the qualitative interview data. In addition, while the measurement scales for Cooperative Teamwork and Opposite-Sex Relationships Self-Concept did

not demonstrate any significant positive short-term effects, both scales were found to have significant positive long-term effects. The short-term and long-term effects for these scales are set out in Table 8.2. Details on these measurement scales can be found in Chapter Three under the sections headed “Self-Concept” and “Life Effectiveness.”

Table 8.2
Short-Term and Long-Term Effect Sizes for Social Competence Outcomes

Program/Analysis	Social Competence Effects: ES (SE)		
	Social Effectiveness	Cooperative Teamwork	Opp-Sex Relationships SC
<i>Adventure Programs</i>			
RCT short-term	.184 (.090)*	.192 (.104)	.059 (.093)
RCT long-term	.049 (.089)	.187 (.089)*	.228 (.081)**
Within-Subjects	.113 (.085)	.192 (.098)*	.218 (.065)***
<i>Arctos Adventure Program</i>			
RCT short-term	-.027 (.147)	.193 (.119)	.176 (.115)
RCT long-term	-.205 (.108)	.109 (.110)	.219 (.125)
Within-Subjects	.267 (.098)**	.182 (.147)	.368 (.095)***
<i>JC Adventure Program</i>			
RCT short-term	.336 (.160)*	.276 (.160)	.097 (.115)
RCT long-term	.296 (.120)*	.352 (.162)*	.331 (.103)***
Within-Subjects	.360 (.141)*	.178 (.155)	.117 (.127)
<i>OB Adventure Program</i>			
RCT short-term	.244 (.128)	.107 (.144)	-.095 (.149)
RCT long-term	.055 (.132)	.100 (.125)	.133 (.113)
Within-Subjects	-.287 (.202)	.216 (.179)	.170 (.105)
<i>Coaching Only Program</i>			
RCT short-term	.234 (.080)*	-.038 (.082)	.097 (.159)
RCT long-term	.063 (.099)	.001 (.079)	.318 (.085)***
Within-Subjects	.436 (.177)*	.476 (.207)*	.192 (.114)

Note. ES = standardized effect size; SE = standard error; Opp = Opposite; SC = Self-Concept; RCT = randomised controlled trial analysis results; Within-Subjects = waitlist control group data within-subjects pre-post analysis results; JC = James Craig; OB = Outward Bound. * p < .05; ** p < .01; *** p < .001.

Social Effectiveness. Adventure Program participants reported significant positive short-term effects in Social Effectiveness when compared to the control group, however, a significant effect was not maintained over the longer-term. The individual program participants in the James Craig Adventure Program reported a significant positive effect in Social Effectiveness both in the short- and long-term, when compared to the control group. Although the Arctos Adventure Program participants did not report a significant positive short-term effect in Social Effectiveness, they did demonstrate a significant positive effect over the longer

term. The Coaching Only Program participants also reported a significant positive short-term effect.

While the participants in the within-subjects analysis did not evidence a significant pre-post effect in Social Effectiveness for the Adventure Programs (taken together), the participants in the Arctos and James Craig Adventure Programs reported significant pre-post effects, as did the Coaching Only Program participants.

Cooperative Teamwork. None of the program participants in the RCT analysis reported a significant positive short-term effect in Cooperative Teamwork. However, participants in the Adventure Programs (taken together), as well as the Coaching Only Program, reported significant positive pre-post effects for this outcome in the within-subjects analysis. Interestingly, participants in the Adventure Programs, as well as the individual James Craig Adventure Program, demonstrated significant positive long-term effects for this outcome. These delayed outcomes for Cooperative Teamwork may reflect the challenging nature of the group work experienced by some participants (as evidenced in the qualitative interviews), and it may have been only with additional experiences (e.g., Community Project or other opportunities for teamwork at school) that participants appreciated their competence in this area. It may also be that improvements in perceptions of general social effectiveness preceded the more specific perceptions of effectiveness in teamwork.

Opposite-Sex Relationships Self-Concept. It is also interesting to note that while no group of program participants in the RCT analysis reported a significant positive short-term effect in Opposite-Sex Relationships Self-Concept, the within-subjects analysis participants in the Adventure Programs, as well as the individual Arctos program, reported significant positive pre-post effects for this scale. While the lack of a control group raises the question of maturation, the program participants in the RCT analysis from the Adventure Programs, as well as the individual James Craig Adventure Program and Coaching Only Program, reported significant positive long-term effects in Opposite Sex-Relationships Self-Concept (as well as significant follow-up effects between immediate post-test and three months

later). Given that the program groups were single-sex and social competence was not a direct aim of the THP program, this scale was judged to be of low relevance. It is again suggested that improvements in perceptions of general social effectiveness may have preceded the more specific perceptions regarding social relationships with the opposite sex. Participants, however, reported no significant effects across the analyses for Same-Sex Relationships Self-Concept. Consequently, more research is needed to assess these associations.

Qualitative outcomes. Eight of the 13 interview participants spoke explicitly about improvement in their communication skills. In particular, participants spoke about how they were now able to meet and engage with new people, something many of them had previously found challenging. However, being part of an unfamiliar social group also can be challenging, especially for adolescents where interactions with people outside of their friendship groups may be new (Jostad et al., 2015). Participants recognised challenging elements of the program that required them to learn how to communicate within their group, including at times having to find new ways to make those connections (e.g., without words). Some of the participants described times during the program where they, or another team member, struggled with communication, and how that made them feel “angry” or “frustrated” (Daisy), and how it could impact the group dynamic. For example, Amy noticed how when she was explaining something, some group members “might not understand” and how she needed to “make it simpler for them.” As described in Chapter Seven, communication challenges were often influenced by a participant’s meaning-making capacity. However, with support these communication challenges could provide learning opportunities. These challenges also seemed to diminish as participants’ communication skills improved. One participant spoke about how even though communication wasn’t an explicit program “focus,” she realised how “through The Helmsman Project we needed to communicate and actually share ideas to know [group members] more and what their strengths are” (Holly). Another participant experienced that even though there was group conflict “through good communication we’re able to negotiate and

talk about it and try to sort things out” (Amy). More than just an improvement in communication skills, one participant noted:

I learnt that I value communication and people more than I thought. I’ve always referred to myself as someone who’s alone; who can handle things alone. I can’t, and that is what really, I think, got me to open my eyes to that. (Eric)

Participants also saw beyond communication to other important aspects of interpersonal relating. They learned that through communication “you understand each other” and can then be “more empathetic” (Daisy). The challenges in the THP program required support and teamwork, and these experiences helped participants to learn to rely on others, to “listen and understand” (Ben), to put themselves in the “shoes of others” (Dan), and to advocate for themselves, all of which are beneficial to their social competence. Daisy commented on how “speaking up about issues you’re having and communicating to people, and asking for help, that’s very important.”

For a number of the participants, the Community Project provided an opportunity to reinforce and build upon their improved social skills, particularly as these projects often involved people in their community who were unknown to them. For Grace, “presenting to the audience people” during her Community Project “helped build [her] confidence in talking.” Holly noted how after her Community Project at her local homeless shelter where she had an opportunity to engage with an “elderly” at the shelter, she “just felt the need to open up a bit more and just let people like... like meeting new people would feel more better for me and be better, like a better person.”

Elements of the THP program that required participants to interact and rely on others, as well as opportunities to practise a number of social skills, enabled participants to see the value in interpersonal relationships, to develop skills relevant to the establishment and maintenance of those relationships, and to build confidence in their social competence. This finding is consistent with previous research that has suggested that small-group interactions in OAE can facilitate the acquisition of important interpersonal skills (Hattie et al., 1997; Paisley et al., 2008).

Perspective-taking capacity. As detailed in Chapter Seven, the THP program participants evidenced growth in their meaning-making capacities through program participation, creating bigger perspectives. Evidence of this increase in perspective-taking is found not only in the qualitative results, but also within the quantitative analysis. While the research in this area is limited, these results are consistent with some related findings on the influence of OAE programs on participants' self-authorship (McGowan, 2016) and broader life perspectives (Sibthorp et al., 2011; Sibthorp et al., 2008). Perspective-taking capacity is related to intellectual flexibility, open thinking, and the meaning-making capacity with which Study 3 was concerned.

Quantitative outcomes. Perspective-taking was measured in the quantitative analysis through the Open Thinking scale. The results for this scale are set out in Table 8.3. Details on this measurement scales can be found in Chapter Three under the section headed "Life Effectiveness."

Table 8.3
Short-Term Effect Sizes for Open Thinking Outcomes

Program/Analysis	Open Thinking Effects: ES (SE)
<i>Adventure Programs</i>	
RCT	.213 (.108)*
Within-Subjects	.136 (.081)
<i>Arctos Adventure Program</i>	
RCT	-.010 (.144)
Within-Subjects	.271 (.117)*
<i>JC Adventure Program</i>	
RCT	.430 (.149)**
Within-Subjects	.162 (.097)
<i>OB Adventure Program</i>	
RCT	.219 (.149)
Within-Subjects	-.025 (.132)
<i>Coaching Only Program</i>	
RCT	-.056 (.136)
Within-Subjects	.314 (.340)

Note. ES = standardized effect size; SE = standard error; RCT = randomised controlled trial analysis results; Within-Subjects = waitlist control group data within-subjects pre-post analysis results; JC = James Craig; OB = Outward Bound. * $p < .05$; ** $p < .01$; *** $p < .001$.

Program participants in the Adventure Programs (taken together) reported significant positive short-term effects in Open Thinking, when compared to the control group, although this significant effect was not maintained in the longer term. The individual James Craig Adventure Program participants also reported significant positive short-term effects in Open Thinking. For the within-subjects analysis, the Arctos Adventure Program participants reported significant positive pre-post effects for this scale.

Qualitative outcomes. This openness to new ideas and adaptability in thinking were also outcomes that participants mentioned in the interview data. Through the THP program, Alex recognised that to be an effective member of a team, you need to be “open-minded,” “to work with other people even though you might not like them,” and “to be able to adapt to whatever role your team needs.” Holly also spoke about engaging in group work that required her to be “flexible with other people’s ideas.” In addition to confidence and resilience, Ben said that one of the learnings from the program that was most important to him, was “just having an open mind.” When asked what it meant to have an open mind, Ben noted that it involved being “open to other people’s ideas, being willing to listen and not just make sure everything is what you want.” What was most important about having an open mind, Ben said, was that “you’re able to listen and understand.” By having an open mind, participants could let more in, creating possibilities for new and different ways of understanding.

As indicated in Chapter Seven, there were also participants who grew in their meaning-making capacities, allowing them to “shift [their] gaze” and “be more open” (Eric), so as to be able to see and act on more. Some of the coaches seem to have been able to use reflective conversations to help participants notice their current ways of making meaning, to challenge them to do something that stimulated a new way of seeing their world, and to support them into new ways of understanding. Moreover, the group experience could also encourage constructive-developmental growth. Having to interact with others through challenging activities in a small team setting required some participants to act in a whole new way, and for others, it compelled them to examine themselves more closely.

Goal Self-Regulation. The Goal Self-Regulation scale was intended to measure participants' perceived capacities to use self-regulation to promote proactive behaviour, problem solving, and persistence in the pursuit of goals. The THP program had an explicit focus on developing cognitive self-regulation skills specific to goal pursuit, with participants learning to set appropriate goals and monitor and evaluate their goal progress through the experiential learning cycle. Moreover, research has demonstrated a positive connection between extracurricular programs and increased self-regulation, particularly among disadvantaged youth (Bandy & Moore, 2011). Research also has found coaching to significantly improve goal-directed self-regulation (Grant et al., 2009), although this research related to adults in a corporate environment. Indeed, the James Craig Adventure Program participants evidenced a significant positive effect in Goal Self-Regulation in the short-term RCT analysis. Nevertheless, it was surprising that there were not more significant positive results for this scale. While some of the interview participants referred to setting goals and the experience of goal achievement, they only made limited reference to self-regulatory skills used in the goal striving process.

It has been suggested that the ability to self-regulate is impacted by a number of different factors internal and external to a person, including that person's genetics, self-regulatory skillset, motivation (both intrinsic and extrinsic), social support, and environmental context (Murray et al., 2015). Accordingly, it can be a complex skill to empirically assess. There is also the relationship between self-regulation and constructive-developmental capacity. As has been discussed previously, self-regulation is a skill that can be challenging until one begins to move toward development of a self-authoring system. As most of the participants interviewed for this research did not demonstrate this level of constructive-developmental capacity, the expectation for development in this area may be somewhat unrealistic. Assuming there is capacity for self-regulatory development, the outcomes for this scale may have been influenced by participants' particular program experiences related to goal striving, including in relation to the Community Project, as also has been discussed. Additionally, it is possible that self-regulatory skills take practice, and additional experiences and time are required for

these skills to develop. Finally, these results may relate in part to the measure. This scale was derived from the Adolescent Self-Regulatory Inventory items when the a priori hypothesised scales did not demonstrate sound psychometric properties. While the analysis in Chapter Five found this scale to be psychometrically acceptable, it is the first time the scale has been used to measure this construct. Future research should explore further the development of self-regulation in adolescents and its measurement.

Summary. Notwithstanding the diversity in some of the research results, the quantitative and qualitative findings, taken together, suggest that the THP program results in some beneficial program outcomes for some participants. While the quantitative analysis did not find these effects to maintain longer term, the qualitative data indicates some participants did experience long-lasting program outcomes. This qualitative data adds depth to our understanding of the program participant experience. Special attention was given to program effects on a number of global positive self-beliefs, aspects of social competence, and perspective-taking capacities, which were outcomes consistently found across the quantitative and qualitative analyses. In the next section, the relationship between primary program elements, the different ways in which participants understood those program elements, and program outcomes is considered through the qualitative findings.

Primary Program Elements

The design and delivery of the THP program was based on existing evidence from research and practice in OAE and coaching psychology. The primary program elements include challenging individual and group activities experienced through an experiential learning framework, and program facilitation and support through developmental coaching. The interview data provides some insight into the relationship among these primary program elements, participants' meaning-making capacities, and program outcomes. Figure 8.1 summarises these relationships, highlighting the effective balance of challenge and support as an essential element of program success, and the important role that individual meaning-making capacity plays in that balancing process. These relationships are discussed further below.

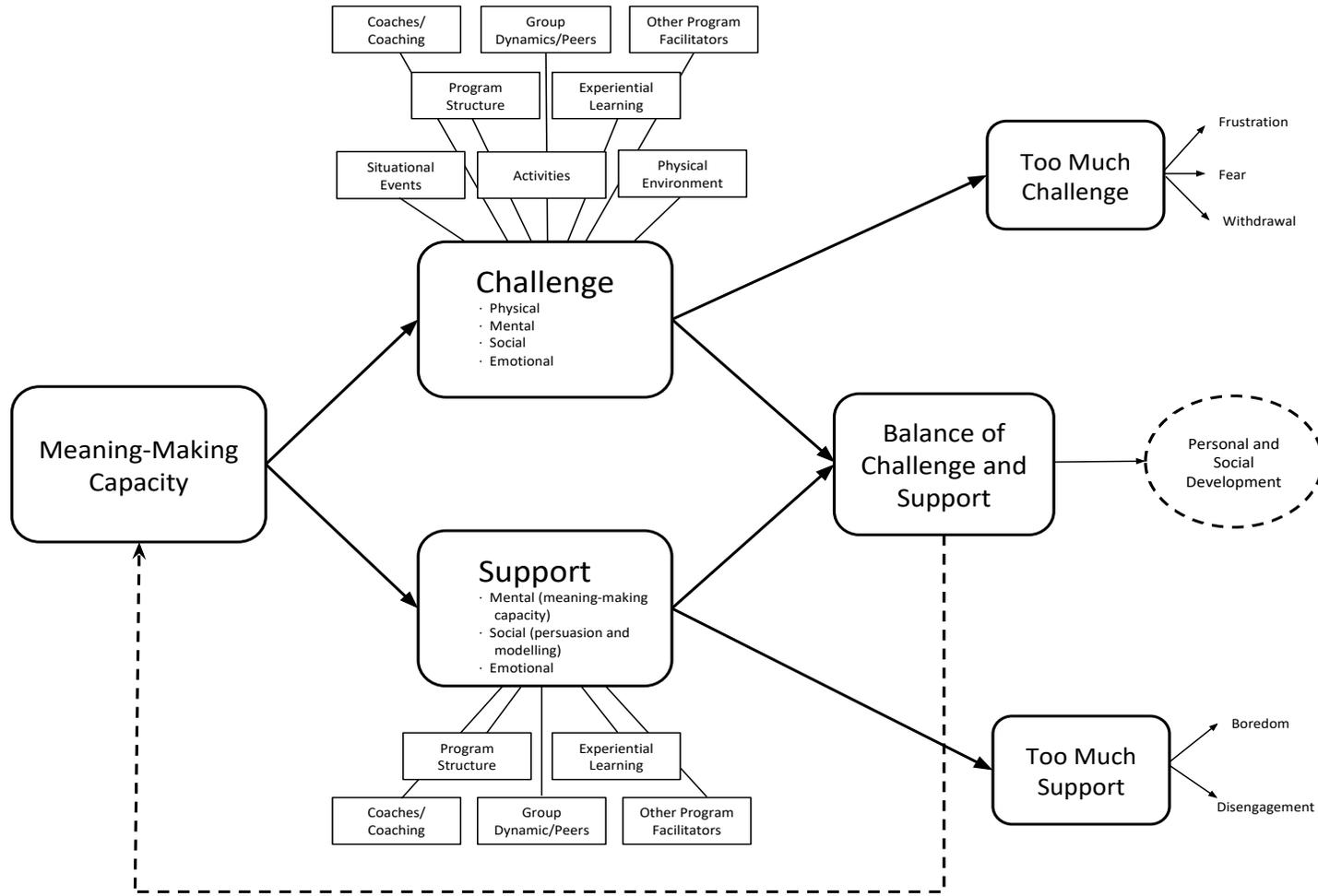


Figure 8.1. Diagram reflecting the relationship between primary program elements, participant meaning-making capacity, and program outcomes.

Challenge and support. Consistent with OAE literature and research findings (Martin & Leberman, 2005; e.g., McKenzie, 2003; Sheard & Golby, 2006; Walsh & Golins, 1976), the qualitative research indicates that challenging experiences in the context of a supported environment were an important program element in delivering positive program outcomes for the THP program participants. However, the interview data suggests that what was “challenging” and “supportive” in a THP program was subjective, being determined at least in part, by the constructive-developmental capacity of the participant. Where the level of challenge and support was effectively balanced from the perspective of an individual participant, both the achievement of the challenge and the experience of the support seemed to lead to positive growth for the participant. However, where the level of challenge exceeded the support from the perspective of the individual participant, and the challenge was not met, the participant could be left with a negative outcome. Although such a situation was not expressed in the interview data, it is hypothesised that where the level of support exceeds the level of challenge from the perspective of the individual participant, positive outcomes may be less likely. However, there is an open question as to whether any negative outcomes might arise.

Experiential learning framework. Kolb’s (1984, 2014) experiential learning cycle, which provides a learning framework for the THP program, promotes learning through an iterative process of experience, reflection, and abstraction. Reflection is a process of interpreting and assigning meaning to experience, and abstraction involves analysing those reflections to theorise concepts and generalisations (Kolb, 1984, 2014). This information is then used to inform future experiences. Being in a new and challenging environment (physically, mentally, emotionally, and socially), takes participants outside their comfort zone, and it is this challenging environment that can lead to transformational learning, resulting in new perspectives on experience. It provides an opportunity for participants to transform not only the experience, but the meaning-making framework itself; to notice *how* one is making meaning of experience and consider alternative ways of understanding experience. The reincorporation of learning into

future experiences that is part of the experiential learning model, allows participants to test new ways of making meaning, becoming potentially transformative of meaning-making capacity. However, this type of accommodative change can be quite difficult.

While the challenging experiences inherent in OAE make experiential learning a common framework in OAE programs, the interview data indicates that some participants (especially with Instrumental or Socializing ways of knowing) struggled with the reflection and abstraction essential to this process and to constructive-developmental growth. Consequently, the process must be well supported. Moreover, it has been suggested that time and space away from an experience can facilitate deeper reflection and enhance transfer of learning (Leberman & Martin, 2004). The interview data supports this idea and indicates that further experiences may also be important. While there were opportunities for facilitated reflection after the adventure experiences, these sessions were available only until the program completed (approximately 5 weeks later). This may not have been a sufficient time away from the experience to effectively process it. There also may not have been enough time during the program to engage in sufficient facilitated reflection on those experiences. For example, these subsequent sessions may have been more focused on other challenges, such as the Community Project pitch due at the end of the program. Although having intact school groups may encourage post-program reflection, whether that occurs is a function of the group dynamic, including the developmental levels of the group members, and opportunities to engage with that group. Therefore, structured and supported post-program reflection may also be an important program element for deriving positive long-term program outcomes. This is consistent with some more recent OAE literature suggesting post-program activities and support may help to maintain program outcomes over the longer term (Allison et al., 2012; Brown, 2010; Leberman & Martin, 2004; Schary, Lewis, & Cardinal, 2015; Schary et al., 2016).

Developmental coaching. Professional coaches can scaffold the experiential learning process. By understanding constructive-developmental theory and the meaning-making capacities of participants, coaches can proactively adapt

program delivery to better align program challenges and supports to meet different individual capacities for meaning making. As evidenced in the interview data, for some participants this scaffolding might require tightening the reigns of the program structure, while for other participants a loosening of those reigns may be what is required. Understanding participants' constructive-developmental capacities will also allow coaches to better support the reflective processes in the experiential learning model. Moreover, it is suggested that there is a role for coaches in helping participants explicitly notice, name, and expand their ways of seeing and talking about themselves and their experiences. Doing so may also better support post-program reflection.

Summary. In this section, the relationship among some of the key program elements, participant meaning-making capacity, and program outcomes was explored through the qualitative data. While challenging experiences in the context of a supported environment were important in delivering positive program outcomes for the THP program participants, the interview data reflected how the elements of challenge and support in a THP program were subjective, being determined at least in part, by the constructive-developmental capacity of the participant. It is suggested that these relationships may inform effective OAE program design for beneficial program outcomes.

Summary

The findings from Study 2 and Study 3 were juxtaposed in order to derive a deeper and more holistic picture of the research results. While the rigorous quantitative analysis for this thesis evidenced a number of significant positive outcomes for THP program participants, the varied results across the analyses indicated there might be unaccounted for variables influencing these results. The qualitative findings provided additional context for the quantitative results, suggesting that situational events, program facilitation and delivery, group dynamics, and post-program experiences, as well as participants' constructive-developmental capacities, may be relevant to participants' program experiences and program outcomes.

Despite the varied quantitative results, there were a number of positive outcomes consistently reflected in the quantitative and qualitative studies that fit with existing research in OAE and coaching psychology. In particular, the THP program participants reported increases in a number of global positive self-beliefs and perceptions of their social competence. Participants also evidenced a broadening of their perspectives and openness to new ideas, as well as development in the ways in which they understand their experiences.

Finally, while challenging experiences and a supportive environment were important elements in achieving the program outcomes, how participants made meaning of these program elements was relevant to those outcomes.

In the final sections of this chapter, all three studies are considered in order to outline the strengths, limitations, and directions for future research, as well as implications of the findings for policy and practice. This chapter will end with some concluding comments on the research for this thesis.

Strengths, Limitations, and Directions for Future Research

Strengths and Directions for Future Research

This thesis made a number of methodological and theoretical contributions to the existing research base in the fields of OAE and coaching psychology. Some of these contributions have been mentioned in the chapters on the studies to which they relate. An overview of these contributions are presented first, followed by a more detailed discussion.

Methodological contributions. A primary methodological strength of this thesis relates to how it addressed key criticisms of previous approaches to research design and measurement in the OAE and coaching literature (Cason & Gillis, 1994; Ewert & McAvoy, 2000; Hattie et al., 1997; Neill, 2008; Scrutton & Beames, 2015; Sibthorp, 2000). More specifically,

- This research used many well-established, robust outcome measures, and further established their psychometric properties.
- The research design and statistical methods were rigorous, including an RCT design and longitudinal analysis, as well as advanced statistical techniques.

- A mixed-methods approach was applied in order to provide the most complete picture of the THP program effects and help to further the evidence base for the benefits that OAE and coaching have to offer.

Each of these methodological contributions is expanded upon further below.

Robust outcome measures. This research used many well-established, robust outcome measures with strong connections to existing theory and practice in the areas of OAE, coaching, and personal and social development. Many of these measures had demonstrated established psychometric properties and suitability for use with an adolescent population. The psychometric properties of these measures were re-assessed to ensure they demonstrated sound psychometric properties for the data used in this thesis. Moreover, the psychometric analysis used rigorous statistical procedures, including omega estimates for internal consistency reliability, exploratory structural equation modelling in addition to confirmatory factor analysis, invariance testing, a multi-trait multi-method matrix for construct validity, and reporting of confidence intervals in accordance with best practice. The results of the psychometric analysis provided additional psychometric research on the measurement scales and extended existing validation research to the use of these measures with disadvantaged adolescent populations, as well as invariance testing in connection with a number of measures for which such testing was not previously found.

Rigorous research design and statistical methods. The research design and statistical methods used in connection with Study 2 were rigorous and met criticisms of existing research in the fields of OAE and coaching psychology. In particular, an RCT design was used, which provides a stronger test of potential program effects relative to simpler pre-post models of analysis. Data for the RCT analysis included 362 participants who were all in year nine in school at the time of program participation. Participants also included multiple cohorts from multiple schools, providing statistical power and greater generalizability of the results. The design also included an extended baseline for the control group to test stability of the measurement scales, as well as a longitudinal assessment to examine outcomes over time (including maintenance of short-term positive effects and sleeper effects).

A within-subjects pre-post analysis of the control group data also was undertaken to test the replicability of the RCT short-term analysis. This type of analysis reduces the influence of individual differences in the data since each person effectively serves as his or her own control. Moreover, rigorous and advanced statistical analysis was applied using best practice modelling techniques, including factor scores, multiple regression analysis, cluster-robust standard errors, and controls for certain potential biases and pre-existing differences. The measurement scales were grouped by program relevance to aid in the interpretation of the results, and standardised effect sizes were reported for easier interpretation and comparison across research studies.

Mixed-methods approach. The research design for this thesis includes both quantitative and qualitative analysis. The quantitative research methods have the capacity to analyse a larger, randomised sample, strengthening the inferences that can be made from the data and the generalizability of those inferences. While the rigorous quantitative design and analysis can account for many variables reducing random error in the analysis, there are often additional variables for which the quantitative analysis cannot account. The qualitative analysis, which seeks to understand the elements of the quantitative analysis, can consider these additional variables. It can also provide information that helps to answer *how* and *why* outcomes are or are not found in the quantitative analysis; information that is not often available through quantitative methods. Given the complex nature of OAE and coaching, a mixed-methods research approach provides a more nuanced and complete picture of participation in a THP program and its outcomes. This research also appears to be the first time an integrated developmental coaching and OAE intervention has been assessed with a constructive-developmental lens. This research design therefore satisfies the need for both quantitative and qualitative data in OAE and coaching research.

Theoretical contributions. The novel, school-based THP program, integrating developmental coaching with OAE, is a primary strength of this thesis. By using an evidence-based program structure and detailing the various elements of that structure, the intervention provided the opportunity to assess many aspects of

extracurricular programming and its potential benefits for disadvantaged youth.

More specifically,

- The research results across a broad range of outcomes provide support for, and add to, existing research on the effects of OAE on the personal and social development of adolescents.
- The research findings lend support to the proposition that the content of the adventure experience is less important than the processes involved.
- While there is no clear finding on the incremental benefits of the adventure experience for the THP program, this appears to be the first study to compare outcomes between adventure-oriented and non-adventure extracurricular school-based coaching programs within a single research study.
- The results of the aptitude-treatment interactions add to existing research on extracurricular programs with greater effects for those students most in need.
- The research findings add to existing research on the use of OAE with intact groups.
- The qualitative findings reaffirm the complex nature of OAE and the value of a mixed-methods approach to research in this field. In particular, these findings support, and add to, existing OAE research in a number of areas, including longitudinal effects, relevance of individual developmental characteristics, balance of challenge and support, and the role of developmentally trained coaches in OAE.

These theoretical contributions are discussed in more detail below.

Broad range of outcomes. The breadth of the outcome measures used in this thesis allowed for the impact of the THP program to be considered against a wide variety of outcomes, many of which had been the subject of prior research in the fields of OAE and coaching psychology. The findings of a number of significant, small-to-moderate positive effects across a range of outcomes fits within the research and literature indicating OAE has a positive influence on the personal and social development of adolescents (Bowen & Neill, 2013; Cason & Gillis, 1994; Ewert,

1987; Ewert et al., 2007; Hattie et al., 1997; Marsh et al., 1986a; Neill, 2008), including disadvantaged adolescents (Norton & Watt, 2014). Importantly, this research provides empirical support for OAE and developmental coaching with disadvantaged youth. Moreover, those measurement scales with the greatest relevance to the THP program's aims and design were found to have the strongest effects, consistent with findings by Marsh et al. (1986a, 1986b) in which the effects of Outward Bound on multiple dimensions of self-concept were larger for those scales that were most relevant to the goals of the program.

More particularly, the positive outcomes observed in self-esteem and self-efficacy, both quantitatively and qualitatively, add to extensive prior research that has found OAE to be beneficial in developing these important self-perceptions (Barrett & Greenaway, 1995; Bowen & Neill, 2013; Burton, 1981; Crompton & Sellar, 1981; Davidson, 2001; Ewert, 1987; Hattie et al., 1997; Hunter et al., 2010; Marsh et al., 1986a, 1986b; Paxton & McAvoy, 2000; Propst & Koesler, 1998; Rose et al., 2018; Scarf et al., 2018). It is suggested that future research continues to measure aspects of self-efficacy and self-concept in OAE and that future coaching research also consider these important outcomes. While self-confidence also demonstrated some of the strongest results across this thesis, there is less quantitative evidence in OAE and coaching psychology that measures this as an outcome variable (see, however, Neill, 2008 and other research that uses the LEQ or ROPELOC instruments). Consequently, future research in OAE and coaching might focus more closely on self-confidence as an outcome measure. In addition to these positive global self-beliefs, resilience and social competence are also important continued points of focus, as well as the interrelationships among these outcomes.

Furthermore, while the quantitative data did not find many program outcomes that were maintained in the longitudinal analysis, the qualitative findings tell a different story, indicating positive effects across a number of outcomes, including self-confidence, self-efficacy, self-esteem, hope, and resilience. These findings add further support to research that has found the effects of OAE to be long-lasting (e.g., Hattie et al., 1997; Marsh et al., 1986a; Norton & Watt, 2014) and that has indicated that some effects may take time (e.g., Davidson, 2001; Duerden, Witt, et al., 2012). While research should continue to assess the longitudinal

impacts of OAE, it is suggested that a longer timeframe for this assessment may demonstrate better results. Additionally, qualitative longitudinal research can provide a dual function of assessment and post-program reflection.

OAE content versus processes. The multiple modes through which the adventure experience was offered provided an opportunity to compare the effects of those adventure experiences. Prior literature and research have suggested that the content of the adventure experience may be less important than the processes involved (McKenzie, 2000; Ungar et al., 2005). The research results support the proposition that the program processes, rather than the type of adventure, appear most relevant to program outcomes. The qualitative findings highlighted additional variables that potentially influenced these distinct results, including situational events, program delivery, group dynamics, post-program experience, and individual developmental differences. Nevertheless, given that the two programs experiencing the most significant program outcomes were both water-based programs, there may be an open question as to whether being on the water (particularly where that is especially novel for the participants) has a greater influence on program outcomes. In addition to the novelty and excitement that being on the water provides, it may be that these programs necessitate greater group interaction that supports team-building and a sense of belonging. Further research is required in this regard.

Adventure-based versus non-adventure-based programming. An alternative coaching program without the adventure component was also offered. Recent calls have been made for further research evaluation of OAE programs in Australia not only to assess program outcomes, but also to contrast these outcomes with evidence-based research on other types of intervention (Bowen, Neill, Williams, et al., 2016). While prior meta-analytic research has been used to compare different extracurricular programming (e.g., Hattie et al., 1997; Neill, 2008), this appears to be the first study to compare outcomes between adventure-oriented and non-adventure extracurricular school-based coaching programs within a single research study. While the Adventure Programs demonstrated significantly stronger effects when compared to the Coaching Only Program on a number of outcomes in the short-term, most of these significant differences were not

maintained in the longer term. There were also outcomes for which the Coaching Only Program reported significantly stronger effects than the Adventure Programs. Given the variability in these results and the small sample size for the Coaching Only Program, further research is needed to better understand these outcomes and the incremental value of the adventure experience. Having an adventure only program (without the coaching element) would be valuable in order to fully appreciate all of the incremental effects.

Outcomes with greater effects for those most in need. An additional contribution of this thesis relates to the extensive aptitude-treatment interaction analysis, providing additional data on the effectiveness of the THP programs for those participants with the lowest baseline levels in the outcome variables of interest and, therefore, the greatest needs. These findings add to related research that has found the benefits of extracurricular activities that are school-based to be greater for disadvantaged than advantaged students (Marsh, 1992a; Marsh & Kleitman, 2002). Future research should aim to develop a deeper understanding of those program aspects that provide the greatest impact for more disadvantaged adolescents, as such effects provide an opportunity to reduce the disparity in educational outcomes for disadvantaged youth.

OAE with intact groups. This thesis contributes to a gap in the literature on OAE programming with intact groups from the same school. Although the quantitative outcomes did not demonstrate effect sizes significantly greater than more common OAE programs, some of the qualitative findings suggest there are benefits associated with connection and support from within one's community and school, and the negative effects that can arise where this is lacking. While further research is required, it is suggested that the qualitative findings support the literature on the important role the school and community can play in bridging the learning transfer from OAE (Richmond et al., 2018). However, this requires active school involvement, as well as collaboration and coordination between the school and OAE providers. Research should continue to investigate the benefits of school-based OAE programs, including whether outcomes generalise to school engagement and achievement. As part of this thesis, it was hoped to gather information on

participant attendance, as well as parent and teacher perspectives. However, there were financial, privacy, and other logistical circumstances that made gathering such information difficult in this instance. It is suggested that future studies consider collecting such data in order to provide additional insight into program outcomes over time and across domains. It will also be important to assess how schools and OAE providers can work together to better transfer learning following OAE experiences, allowing students to retain and further develop positive program outcomes.

Qualitative findings. The qualitative interviews provide valuable data that adds to, and informs, the quantitative data, reaffirming the complex nature of OAE. It may be that these subsequent self-perceptions of participant outcomes (made on average 1.24 years after program completion) are more representative of program efficacy than the longitudinal quantitative data, which was collected three months after program completion. The qualitative data also extends the limited prior research into the relevance of individual developmental characteristics as an important form of individual difference in OAE (Collins et al., 2012), providing further understanding of the influence of meaning-making capacity on OAE program experiences and outcomes. This is a key individual characteristic leading to differences in participants' experiences of OAE that has not been given adequate attention in the existing literature and research on OAE.

The qualitative findings also demonstrate how program facilitation through developmentally trained coaches can provide an effective scaffold both to manage program challenges and to stimulate constructive-developmental growth, supporting existing literature and research in OAE that emphasise the important role program facilitators plays in OAE outcomes (Ewert & McAvoy, 2000; Gass & Gillis, 1995; Hattie et al., 1997; Kemp, 2006; Luckner & Nadler, 1997; Martin & Leberman, 2005). These findings also add to existing research by providing a more nuanced perspective on the balance of challenge and support in OAE programming, the impact an imbalance can have on program outcomes, and the role of facilitators in managing that balance. This thesis also extends existing theory by connecting developmental coaching to constructive-developmental growth through OAE. It is

suggested that future research examine in more detail the characteristics of coaches that influence program outcomes, including constructive-developmental growth. Most notably, this thesis contributes more broadly to the understanding of the processes and methods through which OAE outcomes are achieved. It is hoped that there may be implications in these findings for the ways in which OAE programs are designed and implemented.

Limitations and Directions for Future Research

The findings in this thesis are subject to several limitations and suggest a number of avenues for future research. Some of these limitations have been mentioned in the chapters on the studies to which they relate. In general, these limitations relate to:

- the psychometric analysis of the Survey;
- the breadth of the outcomes included in the quantitative analysis;
- the sample of students for the research;
- implementation of the RCT study in a field setting; and
- aspects of the qualitative design.

These limitations are expanded upon below, together with some suggestions for future research.

Survey instrument. The Survey incorporated 41 scales from 11 different measurement instruments and doing so required modification of some of the instruments from their original form. While a significant level of psychometric evaluation was done on these scales, it may be difficult to generalise those results. Moreover, it was not possible to assess the structural integrity of the measurement instrument by modelling all of the scales together, owing to the model complexity in the context of sample size limitations. Furthermore, while the analysis in Study 1 of this thesis found overall support for the psychometric properties of the measurement scales, some scales did not evidence adequate internal consistency reliability and a few of the scales required structural adjustment in order to proceed. Consequently, future research should further evaluate the psychometric properties of these scales.

Quantitative outcomes. While the breadth of the outcomes included in the research are a strength of this thesis, they also are a limitation. Being so comprehensive resulted in a measurement instrument with over 200 items. The length of the instrument, coupled with the repeated measures design, may have caused participants to complete the Survey with less than their full attention and consideration (Davidson et al., 2016). Moreover, some of the Survey items may have been too complex or ambiguous for the participants, particularly at their developmental level, and other items have an ‘all or nothing’ context that do not seek incremental change of the type anticipated. Furthermore, some of the scales consisted entirely of negatively-worded items, which can be problematic, particularly for adolescents (Marsh, 1986a; Melnick & Gable, 1990).

The inclusion of so many scales created challenges in the statistical analysis and presentation of results for this research. Confidence in the generalisability of multiple analyses is weaker than for a single analysis. Moreover, presenting the results of so many scales makes it difficult for readers to fully appreciate the program effects. While we could have reported on a subset of the scales, for reasons of transparency and knowledge development we chose to report on the results for all scales. Another alternative would have been to use exploratory factor analysis to collapse the measurement items into fewer scales. However, that would have created a new measurement instrument, and a primary aim of this thesis was to evaluate the THP program with scales commonly used in OAE research. Multivariate omnibus tests of the high relevance scales, which revealed statistically significant intervention effects, support tests of intervention effects for the individual scales. Notwithstanding our approach, the findings in this thesis that suggest better effects for program outcomes that are more closely aligned with program aims and design should be borne in mind in future research, with an explicit focus on limiting the number of outcome measures. This echoes a recent suggestion by Dawson, Yeomans, and Brown (2018).

Research sample. This thesis is limited by its population of students from a specific area of socioeconomic disadvantage in Sydney, Australia. Therefore, it may be difficult to generalise the results to other populations. The requirement that

students apply to the program and provide parental consent raise the issue of potential bias and further affect the potential generalizability of the results. Moreover, while 362 students (186 intervention and 176 control) participated in the quantitative research, the sample sizes for the individual program modes were small (with an average of 46.5 intervention participants for an individual program), and measurement attrition further impacted sample size. In particular, the Coaching Only Program had only 28 intervention participants to begin. Furthermore, the within-subjects analysis of the control group's extended results through participation in a THP program was impacted by attrition, which is not uncommon in a study seeking to collect data at five timewaves across an entire school year (or more). These smaller samples affected statistical power and internal validity. Future studies with an increased sample size are suggested in order to verify these results. Moreover, future research should continue to compare the benefit of a coaching only program with an adventure-based program, particularly given the additional costs and logistics of implementing the adventure component. However, it is important to note that there are also costs and logistics of implementing a coaching program, given the small supply of professional coaches trained in developmental psychology.

RCT design. Applying an RCT research design in a field setting with an intervention provided by an industry partner had limitations. While RCTs are the “gold standard” in research, ethical and logistical considerations in the school context commonly impact the implementation of RCT research. In this research, groups had to be single-gender affecting random allocation and matching control groups. Gender-Treatment interaction analysis raised questions about whether gender had an effect on program outcomes. While the Adventure Programs (taken together) evidenced only a single significant Gender-Treatment interaction effect across the short-term and long-term analyses, the individual THP programs (considered separately) indicated females may have gained more than males for some of the outcome variables and some of the THP programs, and males may have declined more than females in some outcome variables and for some of the THP programs. As the sample sizes for this analysis were small, it is suggested that future

research should continue to consider the interaction between gender and OAE program effects. Mixed-methods approaches such as the approach used in this research, can help to overcome some of these issues.

Further related to the logistics of implementation, intervention and control groups were located within the same school and in one case, Adventure Programs and Coaching Only Programs were also offered in the same school, raising the possibility of contamination effects. Participation also required completion of an application and consent forms bringing into question issues of consent and selection bias.

There were also logistical issues in connection with quantitative data collection. This process required coordination with the schools, and this had an impact on timing and in some cases affected attrition as a direct result of timing or where collection was unable to be coordinated. Where the timing of the longitudinal data collection coincided with the Community Project, the experience of that project may have influenced those results. A longer follow-up period that occurs after completion of the Community Project, together with separate measurement of that program element, is necessary to understand the effects of the Community Project and the full benefits of the THP program. A longer follow-up period is particularly important for complex interventions where outcomes may take time to develop.

Furthermore, working with an industry partner to deliver the intervention affected the research design. The initial RCT design contemplated a single Adventure Program mode through a small yacht sailing adventure. However, subsequent program considerations required the Adventure Programs to be offered through multiple modes. While the various modes of the Adventure Programs were a strength of this thesis, they were also a limitation. They reduced the effective sample size and added additional elements that could not be controlled, complicating the analyses and interpretations of the results. Beyond the multiple program modes, the desire for an effective sample size while still maintaining small groups meant that research was collected from 54 different participant groups over the course of three years. While the statistical analyses used cluster-robust standard

errors, we were unable to measure the quality and consistency of program implementation.

One source of variance in program implementation is the coaching. Although coaches were selected on the basis of their coaching qualifications and experience, and specifically trained in the theories underlying the program, most of the coaches had not previously worked with adolescents and few of the coaches undertook coaching across multiple cohorts. Furthermore, the program structure intentionally provided flexibility in program implementation. The qualitative analysis used in this research provides insight into aspects of program implementation. However, future studies might also consider measuring implementation fidelity, as well as evaluating coach performance.

Beyond program implementation, it is difficult in an RCT to account for the myriad of variables that can arise in the complex environments of schools and the outdoors. Nevertheless, some of the effects of these variables are captured through the qualitative analysis. These qualitative findings can also inform ways to address such variables in future quantitative studies. As has been suggested in the OAE literature (Ewert & McAvoy, 2000; Ewert & Sibthorp, 2009; Rowley, 1987; Scrutton & Beames, 2015), these results underscore the importance of trying to account for more of these variables in future studies, and the important role qualitative research may play in doing so. However, future studies should consider also incorporating additional outcome measures (e.g., school reported data and third party reports) to balance biases and other limitations inherent in self-report measures.

Qualitative design. While the qualitative research in this thesis provides valuable data on the participant experience and longer-term outcomes associated with OAE programs, it also has limitations. Recruitment of interview participants through the schools and additional consent requirements may have resulted in consent and selection bias. Interview participants may also have been orienting to the particular demands of the interview, although it is suggested that the demands of the type of interview used should distract from any such orientation. Furthermore, financial and other logistical issues prevented pre-post analysis of constructive-developmental stage, as well as inclusion of a comparison group, both

of which would have strengthened the results. While this thesis provides an initial consideration of the connection between OAE program elements, program outcomes, and the constructive-developmental capacities of program participants, future research should aim to consider the constructive-developmental capacity of participants both before and after program participation, as doing so may provide a better understanding of whether and how such development occurs. Including such an assessment within a quantitative study would also allow analysis of the moderating effect of constructive-development on OAE program outcomes.

Implications for Educational Policy and OAE Research, Design, and Practice

This thesis expands previous research on OAE, coaching psychology, and school-based extracurricular programming through its examination of a novel program integrating developmental coaching and OAE for disadvantaged students. Evidence of program effectiveness is important to educational policymakers and those who fund non-profit programs such as The Helmsman Project. Accordingly, the application of rigorous, advanced, and holistic mixed-methodological approaches in this thesis, leading to important findings, should be of interest to educators, programmers, and policymakers in the fields of OAE, coaching, and education, as well as those broadly interested in assisting adolescents to develop the personal and social skills necessary to flourish and reach their full potential.

First, the findings from this thesis reinforce existing research demonstrating the positive effects OAE has on the personal and social development of adolescents, including disadvantaged adolescents. Some of these effects appear to be long-lasting and generalisable to other aspects of participants' everyday lives. These findings reinforce the value in providing school-based OAE programs. The further finding that the effects for some outcome variables were significantly greater for participants who had a lower baseline value in that outcome variable, suggests that OAE programs like the THP program may be able to offer compensatory benefits to disadvantaged students that may ultimately support a lessening of inequity in educational outcomes. This reinforces the need for such programs to be as readily available to disadvantaged students as they are to students with greater advantages.

Second, the findings of individual differences in constructive-developmental capacity and the influence this has on participants' experiences of elements of OAE programs (e.g., program structure, challenge, group work, and coaching), has implications for the design and implementation of OAE. OAE programs need to be careful not to create program designs that demand capacities that exceed the capacities of their participants without appropriate support. Equally, such program designs should not consider participants only in terms of their existing capacities. This is a difficult balance to achieve when participants have a range of constructive-developmental capacities. While flexibility in programming is recommended as an important program element, it is suggested that further consideration needs to be given to better matching program participants with appropriate experiences. It is also suggested that making some of the aims and theories underlying the program more transparent can be helpful for participants at every constructive-developmental level. These findings also have implications for evaluations of youth programs targeting personal and social development. As evidenced by the aptitude-treatment interaction results, the differing developmental capacities and needs of participants make standard outcome evaluation problematic. Future program evaluation needs to account for these differences.

Moreover, the distinct results across programs and analyses suggest a need for OAE providers to be more fully attuned to all of the different elements that potentially have an impact on OAE outcomes, as well as an understanding that these elements need to be considered in the context of the meaning that is made of them by the participants. These findings also potentially have flow-on implications for educators of our youth more broadly. By better understanding these meaning-making distinctions, OAE providers, coaches, and educators can match processes and expectations more closely to developmental capacities and provide a better holding environment for learning and developmental growth.

Finally, the long-lasting effects experienced in connection with the Community Project (both positive and negative) highlight the important role that the post-program experience has in program outcomes. While further research is required, it is suggested that schools and communities have important roles to play in bridging the learning transfer from OAE so that students can retain and further

develop positive program outcomes. However, this requires active school involvement, as well as collaboration and coordination between schools and program providers. Key to this process are opportunities for further challenges, reflection, and support.

Conclusion

Development during adolescence has a deep impact on functioning, health, and wellbeing throughout the rest of one's life. Consequently, there is an opportunity during this period to positively influence development in a way that will have lasting impact. Programs that provide opportunities to enhance positive beliefs about oneself and other non-cognitive skills can be important elements in this development. Evidence has found both OAE and coaching to be avenues for promoting positive development in youth.

The THP programs are novel extracurricular programs integrating a series of structured developmental coaching sessions with outdoor adventure experiences, which aim to foster personal and social development through a range of outcomes. A further program objective is for participants to develop broader perspective-taking capacities, enabling them to perceive different and bigger possibilities than before.

This thesis offers three interrelated studies to provide a holistic investigation into the effectiveness of the THP programs and to expand the understanding of the complex processes and outcomes associated with OAE programs, including the relationship among program participants, the various program elements, and the numerous outcome measures. This thesis also aimed to bolster confidence in existing OAE and coaching research findings by applying a rigorous multidimensional research design and advanced statistical analyses.

One important conclusion from this research is that an integrated OAE and developmental coaching program offers a valuable extracurricular framework for promoting growth in a range of outcomes related to participants' personal and social development, including a number of positive self-beliefs and other important life skills and qualities. Notably, some of the beneficial outcomes from this program may be greater for participants with lower baseline levels in the outcome variable,

making this an important point of focus for disadvantaged students who generally have less access to programs aimed at developing these outcomes.

Moreover, this thesis provides strong support for the theoretical claim that the constructive-developmental capacity of OAE program participants is a relevant individual difference influencing OAE program experiences and outcomes. Furthermore, this research suggests that there is an important role for developmentally trained coaches in providing an effective scaffold that both manages OAE program challenges and stimulates constructive-developmental growth. However, it is suggested that simply appreciating constructive-developmental theory can create an awareness of individual difference in experience that will allow OAE program providers to match processes and expectations more closely to developmental capacities and provide a holding environment for its participants that better supports their learning and developmental growth. Consequently, constructive-developmental theory provides a valuable and novel frame of reference for OAE researchers and practitioners.

Finally, the findings in this thesis evidence the value in a mixed-methods approach to OAE research. The quantitative and qualitative studies in this thesis each demonstrated significant strengths as well as limitations. Both studies reaffirm the complex nature of OAE and the challenges inherent in conducting field research on OAE programs and their outcomes. Taken together, these studies provide a more complete and holistic understanding of the relationship among OAE program elements, participants, and outcomes. The methodologically rigorous results from this multidimensional investigation contribute to literature, research, and practice in the fields of OAE and coaching psychology and provide a platform for future study.

REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: SAGE.
- Alessandri, G., Vecchione, M., Fagnani, C., Bentler, P. M., Barbaranelli, C., Medda, E., . . . Caprara, G. V. (2010). Much more than model fitting? Evidence for the heritability of method effect associated with positively worded items of the Life Orientation Test Revised. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(4), 642-653. doi:10.1080/10705511.2010.510064
- Allen, K. (2016). *Theory, research, and practical guidelines for family life coaching*. Switzerland: Springer International.
- Allison, P. (2000). *Research from the ground up: Post-expedition adjustment*. Brathay Hall, UK: Brathay Hall Trust.
- Allison, P., Davis-Berman, J., & Berman, D. (2012). Changes in latitude, changes in attitude: analysis of the effects of reverse culture shock - a study of students returning from youth expeditions. *Leisure Studies*, *31*(4), 487-503. doi:10.1080/02614367.2011.619011
- Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: SAGE.
- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 545-557. doi:10.1037/0021-843X.112.4.545
- American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.
- Anderman, L. H., & Freeman, T. M. (2004). Students' sense of belonging in school. In P. R. Pintrich & M. L. Maehr (Eds.), *Advances in motivation and achievement series (Vol. 13): Motivating students, improving schools: The legacy of Carol Midgley* (pp. 27-63). Oxford, England: Elsevier.
- Appaneal, R. N. (2012). A confirmatory factor analysis of the Life Orientation Test-Revised with competitive athletes. *Research Quarterly for Exercise and Sport*, *83*(4), 592-596. doi:10.1080/02701367.2012.10599885

- Aron, A., Coups, E. J., & Aron, E. (2013). *Statistics for psychology* (6th ed.). Boston, MA: Pearson.
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397-438. doi:10.1080/10705510903008204
- Asparouhov, T., Muthén, B., & Morin, A. J. S. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances: Comments on Stromeier et al. *Journal of Management*, *41*(6), 1561-1577. doi:10.1177/0149206315591075
- Association for Psychological Science. (22 October, 2018). Submission Guidelines. Retrieved from https://www.psychologicalscience.org/publications/psychological_science/p-s-submissions
- Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin*, *120*(3), 338-375. doi:10.1037/0033-2909.120.3.338
- Bache, S. M., & Wickham, H. (2014). magrittr: A Forward-Pipe Operator for R [Computer software]. Retrieved from <https://CRAN.R-project.org/package=magrittr>
- Bachkirova, T. (2010). The cognitive-developmental approach to coaching. In E. Cox, T. Bachkirova, & D. Clutterbuck (Eds.), *The complete handbook of coaching* (pp. 132-145). London, England: SAGE.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1997a). Insights: Self-efficacy. *Harvard Mental Health Letter*, *13*(9), 4.
- Bandura, A. (1997b). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Bandura, A. (2013). The role of self-efficacy in goal-based motivation. In E. A. Locke & G. P. Latham (Eds.), *New developments in goal setting and task performance* (pp. 147-157). New York, NY: Routledge.
- Bandy, T., & Moore, K. A. (2011). *What works for promoting and enhancing positive social skills: Lessons from experimental evaluations of programs and*

- interventions [Fact sheet]*. Retrieved from https://www.childtrends.org/wp-content/uploads/2011/03/child_trends_2011_03_02_RB_WWSocialSkills.pdf
- Barnette, J. J. (2000). Effects of stem and likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, *60*(3), 361-370.
- Barrett, J., & Greenaway, R. (1995). *Why adventure? The role and value of outdoor adventure in young people's personal and social development: A review of research*. Retrieved from <https://www.englishoutdoorcouncil.org/publications/why-adventure>
- Bartram, D., Sinclair, J., & Baldwin, D. (2013). Further validation of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) in the UK veterinary profession: Rasch analysis. *An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation - Official Journal of the International Society of Quality of Life Research*, *22*(2), 379-391. doi:10.1007/s11336-012-0144-4
- Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research*, *40*, 373-400.
- Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*(6), 351-355. doi:10.1111/j.1467-8721.2007.00534.x
- Baxter Magolda, M. B., Creamer, E. G., & Meszaros, P. S. (2010). *Development and assessment of self-authorship: Exploring the concept across cultures* (1st ed.). Sterling, VA: Stylus.
- Beard, C., & Wilson, J. P. (2013). *Experiential learning: A handbook for education, training and coaching* (3rd ed.). London: Kogan Page.
- Beightol, J., Jeverson, J., Gray, S., Carter, S., & Gass, M. (2009). The effect of an experiential adventure-based "anti-bullying initiative" on levels of resilience: A mixed methods study. *Journal of Experiential Education*, *31*(3), 420-424.

- Bender, R., & Lange, S. (2001). Adjusting for multiple testing—when and how? *Journal of Clinical Epidemiology*, *54*(4), 343-349. doi:10.1016/S0895-4356(00)00314-0
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289-300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238-246. doi:10.1037/0033-2909.107.2.238
- Berger, J. (2002). *Exploring the connection between teacher education practice and adult development theory*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3055838)
- Berger, J. G. (2003). A summary of the constructive-developmental theory of Robert Kegan. Retrieved from <https://wiki.canterbury.ac.nz/download/attachments/6358104/berger+on+kegan+narrative.doc>
- Berger, J. G. (2004). Dancing on the threshold of meaning: Recognizing and understanding the growing edge. *Journal of Transformative Education*, *2*(4), 336-351. doi:10.1177/1541344604267697
- Berger, J. G. (2010). Using the subject-object interview to promote and assess self-authorship. In M. B. Baxter Magolda, E. G. Creamer, & P. S. Meszaros (Eds.), *Development and assessment of self-authorship: Exploring the concept across cultures* (1st ed.). Sterling, VA: Stylus.
- Berger, J. G., & Atkins, P. W. (2009). Mapping complexity of mind: Using the subject-object interview in coaching. *Coaching: An International Journal of Theory, Research and Practice*, *2*(1), 23-36. doi:10.1080/17521880902780989
- Bergman, M. M. (2010). On concepts and paradigms in mixed methods research. *Journal of Mixed Methods Research*, *4*(3), 171-175. doi:10.1177/1558689810376950
- Bergsteiner, H., Avery, G. C., & Neumann, R. (2010). Kolb's experiential learning model: Critique from a modelling perspective. *Studies in Continuing Education*, *32*(1), 29-46. doi:10.1080/01580370903534355

- Blackman, A., Moscardo, G., & Gray, D. (2016). Challenges for the theory and practice of business coaching: A systematic review of empirical evidence. *Human Resource Development Review, 15*(4), 459-486. doi:10.1177/1534484316673177
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263. doi:10.1111/j.1467-8624.2007.00995.x
- Bloom, B. (1976). *Human characteristics and school learning*. New York, NY: McGraw Hill.
- Bodkin-Andrews, G. H., Ha, M. T., Craven, R. G., & Yeung, A. S. (2010). Factorial invariance testing and latent mean differences for the Self-Description Questionnaire II (Short Version) with indigenous and non-indigenous Australian secondary school students. *International Journal of Testing, 10*(1), 47-79. doi:10.1080/15305050903352065
- Bowen, D. J., & Neill, J. T. (2013). A meta-analysis of adventure therapy outcomes and moderators. *The Open Psychology Journal, 6*, 28-53. doi:10.2174/1874350120130802001
- Bowen, D. J., & Neill, J. T. (2015). Effects of the PCYC Catalyst outdoor adventure intervention program on youths' life skills, mental health, and delinquent behaviour. *International Journal of Adolescence and Youth, 21*(1), 1-22. doi:10.1080/02673843.2015.1027716
- Bowen, D. J., Neill, J. T., & Crisp, S. J. R. (2016). Wilderness adventure therapy effects on the mental health of youth participants. *Evaluation and Program Planning, 58*, 49-59. doi:10.1016/j.evalprogplan.2016.05.005
- Bowen, D. J., Neill, J. T., Williams, I. R., Mak, A. S., Allen, N. B., & Olsson, C. A. (2016). A profile of outdoor adventure interventions for young people in Australia. *Journal of Outdoor Recreation, Education and Leadership, 8*(1), 26-40. doi:10.18666/JOREL-2016-V8-I1-7281
- Boyatzis, R., Boyatzis, R. E., & Akrivou, K. (2006). The ideal self as the driver of intentional change. *Journal of Management Development, 25*(7), 624-642. doi:10.1108/02621710610678454

- Breckler, S. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin*, *107*(2), 260-273.
- Brown, M. (2010). Transfer: Outdoor adventure education's Achilles heel? Changing participation as a viable option. *Australian Journal of Outdoor Education*, *14*(1), 13-22.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.
- Browne, M., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research*, *21*(2), 230-258.
- Buck, M., & Inman, S. (1998). Personal and social development at the crossroads. In S. Inman, M. Buck, & H. Burke (Eds.), *Assessing personal and social development: Measuring the unmeasurable?* (pp. 1-16). London, England: Falmer Press.
- Burke, D., & Linley, P. A. (2007). Enhancing goal self-concordance through coaching. *International Coaching Psychology Review*, *2*(1), 62-69.
- Burton, L. M. (1981). *A critical analysis and review of the research on Outward Bound and related programs*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 8122147)
- Byrne, B., Shavelson, R., & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456-466.
- Byrne, B. M. (1984). The general/academic self-concept nomological network: A review of construct validation research. *Review of Educational Research*, *54*(3), 427-456. doi:10.3102/00346543054003427
- Byrne, B. M. (1996). *Measuring self-concept across the life span: Issues and instrumentation* (1st ed.). Washington, DC: American Psychological Association.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Routledge.
- Cabin, R. J., & Mitchell, R. J. (2000). To Bonferroni or not to Bonferroni: When and how are the questions. *Bulletin of the Ecological Society of America*, *81*(3), 246-248.

- Campbell, D., & Fiske, D. (1959). Convergent and discriminatn validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81-105.
- Campbell, D. T., & O'Connell, E. J. (1967). Methods factors in multitrait-multimethod matrices: Multiplicative rather than additive? *Multivariate Behavioral Research*, *2*(4), 409-426. doi:10.1207/s15327906mbr0204_1
- Capurso, M., & Borsci, S. (2013). Effects of a tall ship sail training experience on adolescents' self-concept. *International Journal of Educational Research*, *58*(C), 15-24. doi:10.1016/j.ijer.2013.01.004
- Carpenter, C., & Harper, N. (2016). Health and wellbeing benefits of activities in the outdoors. In B. Humberstone, H. Prince, & K. A. Henderson (Eds.), *Routledge International Handbook of outdoor studies*. New York, NY: Routledge.
- Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. Cambridge, England: Cambridge University Press.
- Carver, C. S., & Scheier, M. F. (2011). Self-regulation of action and affect. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (2nd ed., pp. 3-21). New York, NY: Guilford Press.
- Carver, C. S., & Scheier, M. F. (2014). Dispositional optimism. *Trends in Cognitive Sciences*, *18*(6), 293-299. doi:10.1016/j.tics.2014.02.003
- Carver, C. S., Scheier, M. F., & Segerstrom, S. C. (2010). Optimism. *Clinical Psychology Review*, *30*(7), 879-889. doi:10.1016/j.cpr.2010.01.006
- Cason, D., & Gillis, H. L. L. (1994). A meta-analysis of outdoor adventure programming with adolescents. *Journal of Experiential Education*, *17*(1), 40-47. doi:10.1177/105382599401700109
- Castejón, A., & Zancajo, A. (2015). Educational differentiation policies and the performance of disadvantaged students across OECD countries. *European Educational Research Journal*, *14*(3-4), 222-239. doi:10.1177/1474904115592489
- Cavanagh, M. J. (2006). Coaching from a systemic perspective: A complex adaptive conversation. In D. R. Stober & A. M. Grant (Eds.), *Evidence based coaching handbook: Putting best practices to work for your clients* (pp. 313-354). Hoboken, NJ: Wiley.

- Cavanagh, M. J. (2016). The coaching engagement in the twenty-first century: New paradigms for complex times. In S. David, D. Clutterbuck, & D. Megginson (Eds.), *Beyond goals: Effective strategies for coaching and mentoring* (pp. 151-184). London, England: Routledge.
- Chan, D. W. (2009). Orientations to happiness and subjective well-being among Chinese prospective and in-service teachers in Hong Kong. *Educational Psychology, 29*(2), 139-151. doi:10.1080/01443410802570907
- Chang, E. C. (1998). Hope, problem-solving ability, and coping in a college student population: Some implications for theory and practice. *Journal of Clinical Psychology, 54*(7), 953-962.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464-504. doi:10.1080/10705510701301834
- Cheung, A. C. K., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher, 45*(5), 283-292. doi:10.3102/0013189X16656615
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.
- Ciarrochi, J., Heaven, P. C. L., & Davies, F. (2007). The impact of hope, self-esteem, and attributional style on adolescents' school grades and emotional well-being: A longitudinal study. *Journal of Research in Personality, 41*(6), 1161-1178. doi:10.1016/j.jrp.2007.02.001
- Ciarrochi, J., Kashdan, T. B., & Harris, R. (2013). The foundations of flourishing. In T. B. Kashdan & J. Ciarrochi (Eds.), *Mindfulness, acceptance, and positive psychology: The seven foundations of well-being* (pp. 1-29). Oakland, CA: Context Press.
- Clarke, A., Friede, T., Putz, R., Ashdown, J., Martin, S., Blake, A., . . . Stewart-Brown, S. (2011). Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Validated for teenage school students in England and Scotland. A mixed methods assessment. *BMC Public Health, 11*, 487-487. doi:10.1186/1471-2458-11-487

- Claro, S., Paunesku, D., & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(31), 8664.
doi:10.1073/pnas.1608207113
- Clough, P., & Strycharczyk, D. (2015). *Developing mental toughness: Coaching strategies to improve performance, resilience and wellbeing*. London: England: Kogan Page.
- Coats, E., Janoff-Bulman, R., & Alpert, N. (1996). Approach versus avoidance goals: Differences in self-evaluation and well-being. *Personality and Social Psychology Bulletin*, *22*(10), 1057-1067.
doi:https://doi.org/10.1177/01461672962210009
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.
- Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2013). *Psychological testing and assessment : An introduction to tests and measurement* (8th ed.). New York, NY: McGraw Hill.
- Coley, R. J., & Baker, B. (2013). *Poverty and education: Finding the way forward*. Retrieved from
https://www.ets.org/s/research/pdf/poverty_and_education_report.pdf
- Collins, R., Paisley, K., Sibthorp, J., & Gookin, J. (2012). "Black and white thinkers" and "colorful problems": Understanding student thinking in outdoor education. *Journal of Outdoor Recreation, Education and Leadership*, *4*(3), 185-198.
- Cook-Greuter, S. R. (2004). Making the case for a developmental perspective. *Industrial and Commercial Training*, *36*(7), 275-281.
doi:10.1108/00197850410563902
- Cooley, S. J., Holland, M. J. G., Cumming, J., Novakovic, E. G., & Burns, V. E. (2014). Introducing the use of a semi-structured video diary room to investigate students' learning experiences during an outdoor adventure education groupwork skills course. *Higher Education*, *67*(1), 105. doi:10.1007/s10734-013-9645-5

- Coopersmith, S. (1967). *The antecedents of self-esteem*. San Francisco, CA: Freeman.
- Covay, E., & Carbonaro, W. (2010). After the bell: Participation in extracurricular activities, classroom behavior, and academic achievement. *Sociology of Education*, 83(1), 20-45. doi:10.1177/0038040709356565
- Cowen, E. L. (1994). The enhancement of psychological wellness: Challenges and opportunities. *American Journal of Community Psychology*, 22(2), 149-179. doi:10.1007/BF02506861
- Crane, D., Hattie, J., & Houghton, S. (1997). Goal setting and the adventure experience. *Australian Journal of Psychology*, 49(1), 6-13. doi:10.1080/00049539708259844
- Credé, M., Tynan, M. C., & Harms, P. D. (2016). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, 113(3), 492-511. doi:10.1037/pspp0000102
- Creed, P. A., Patton, W., & Bartrum, D. (2002). Multidimensional properties of LOT-R: Effects of optimism and pessimism on career and well-being related variables in adolescents. *Journal of Career Assessment*, 10(1), 42-61. doi:10.1177/1069072702010001003
- Crescioni, A. W., & Baumeister, R. F. (2013). The four needs for meaning, the value gap, and how (and whether) society can fill the void. In J. Hicks & C. Routledge (Eds.), *The experience of meaning in life: Classical perspectives, emerging themes, and controversies* (pp. 3-15). New York, NY: Springer.
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: SAGE.
- Crompton, J. L., & Sellar, C. (1981). Do outdoor education experiences contribute to positive development in the affective domain? *Journal of Environmental Education*, 12(4), 21-29. doi:10.1080/00958964.1981.9942638
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. doi:10.1007/BF02310555
- Csikszentmihalyi, M. (1992). The flow experience and its significance for human psychology. In M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 15-35). Cambridge, England: Cambridge University Press.

- Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research, 18*(2), 147-167. doi:10.1207/s15327906mbr1802_2
- Culhane, J. D. (2004). *The effects of a fifth-grade adventure-based cooperative physical education program on life effectiveness and locus of control*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3139620)
- Cumming, G., Fidler, F., Kalinowski, P., & Lai, J. (2012). The statistical recommendations of the American Psychological Association Publication Manual: Effect sizes, confidence intervals, and meta-analysis. *Australian Journal of Psychology, 64*(3), 138-146. doi:10.1111/j.1742-9536.2011.00037.x
- Cutler, D. M., & Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics, 29*(1), 1-28. doi:10.1016/j.jhealeco.2009.10.003
- Davidson, C., Ewert, A., & Chang, Y. (2016). Multiple methods for identifying outcomes of a high challenge adventure activity. *Journal of Experiential Education, 39*(2), 164-178. doi:10.1177/1053825916634116
- Davidson, L. (2001). Qualitative research and making meaning from adventure: A case study of boys' experiences of outdoor education at school. *Journal of Adventure Education and Outdoor Learning, 1*(2), 11-20.
- Dawson, A., Yeomans, E., & Brown, E. R. (2018). Methodological challenges in education RCTs: reflections from England's Education Endowment Foundation. *Educational Research, 60*(3), 292-310. doi:10.1080/00131881.2018.1500079
- de Boer, M. R., Waterlander, W. E., Kuijper, L. D. J., Steenhuis, I. H. M., & Twisk, J. W. R. (2015). Testing for baseline differences in randomized controlled trials: An unhealthy research behavior that is hard to eradicate. *Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate, 12*(1), 162.
- Delle Fave, A. (2009). Optimal experience and meaning: Which relationship? *Psihologijske Teme, 18*(2), 285-302.

- Delle Fave, A., Brdar, I., Freire, T., Vella-Brodrick, D., & Wissing, M. P. (2011). The eudaimonic and hedonic components of happiness: Qualitative and quantitative findings. *Social Indicators Research, 100*(2), 185-207. doi:10.1007/s11205-010-9632-5
- Desmond, B., & Jowitt, A. (2012). Stepping into the unknown: Dialogical experiential learning. *Journal of Management Development, 31*(3), 221-230. doi:10.1108/02621711211208853
- Dew-Reeves, S. E., Michele Athay, M., & Kelley, S. D. (2012). Validation and use of the Children's Hope Scale-Revised PTPB Edition (CHS-PTPB): High initial youth hope and elevated baseline symptomatology predict poor treatment outcomes. *Administration and Policy in Mental Health and Mental Health Services Research, 39*(1), 60-70. doi:10.1007/s10488-012-0411-2
- Dias, P. C., del Castillo, J. A. G., & Moilanen, K. L. (2014). The adolescent self-regulatory inventory (ASRI) adaptation to portuguese context. *Paideia, 24*(58), 155-164. doi:10.1590/1982-43272458201403
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment, 49*(1), 71-75. doi:10.1207/s15327752jpa4901_13
- Dietrich, J., Parker, P., & Salmela-Aro, K. (2012). Phase-adequate engagement at the post-school transition. *Developmental Psychology, 48*(6), 1575-1593. doi:10.1037/a0030188
- Dillon, J. (2012). Barriers and benefits to learning in natural environments: Towards a reconceptualisation of the possibilities for change. *Cosmos, 08*(02), 153-166. doi:10.1142/s0219607712300056
- Dixson, D. D. (2017). Hope across achievement: Examining psychometric properties of the Children's Hope Scale across the range of achievement. *SAGE Open, 7*(3), 1-11. doi:10.1177/2158244017717304
- Dixson, D. D., Keltner, D., Worrell, F. C., & Mello, Z. (2017). The magic of hope: Hope mediates the relationship between socioeconomic status and academic achievement. *Journal of Educational Research, 1-9*. doi:10.1080/00220671.2017.1302915

- Drago-Severson, E., Helsing, D., Kegan, R., Portnow, K., Popp, N., & Broderick, M. (2001). Describing the NCSALL adult development research. *Focus on Basics*, 5(B), 3-6.
- Duckworth, A. L., Kirby, T. A., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2011). Deliberate practice spells success. *Social Psychological and Personality Science*, 2(2), 174-181. doi:10.1177/1948550610385872
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087-1101. doi:10.1037/0022-3514.92.6.1087
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, 91(2), 166-174. doi:10.1080/00223890802634290
- Duerden, M. D., Taniguchi, S., & Widmer, M. (2012). Antecedents of identity development in a structured recreation setting: A qualitative inquiry. *Journal of Adolescent Research*, 27(2), 183-202. doi:10.1177/0743558411417869
- Duerden, M. D., Witt, P. A., & Taniguchi, S. (2012). The impact of postprogram reflection on recreation program outcome. *Journal of Park & Recreation Administration*, 30(1), 36-50.
- Duncan, G. J., & Murnane, R. J. (Eds.). (2011). *Whither opportunity?: Rising inequality, schools, and children's life chances*. New York, NY: Russell Sage Foundation.
- Dunn, T. J., Baguley, T., & Brunsdon, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399-412. doi:10.1111/bjop.12046
- Durlak, J. A., & Dupre, E. P. (2008). Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *American Journal of Community Psychology*, 41(3-4), 327-350. doi:10.1007/s10464-008-9165-0
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.

- Eatough, V., & Smith, J. A. (2008). Interpretative phenomenological analysis. In C. Willig & W. Stainton-Rogers (Eds.), *The SAGE handbook of qualitative research in psychology* (pp. 179-194). London, England: SAGE.
- Eccles, J. S., Barber, B. L., Stone, M., & Hunt, J. (2003). Extracurricular activities and adolescent development. *Journal of Social Issues, 59*(4), 865-889.
doi:10.1046/j.0022-4537.2003.00095.x
- Edwards, L., Ong, A., & Lopez, S. (2007). Hope measurement in Mexican American youth. *Hispanic Journal of Behavioral Sciences, 29*(2), 225-241.
doi:10.1177/0739986307299692
- Edwards, L., Rand, K. L., Lopez, S. J., & Snyder, C. R. (2007). Understanding hope: A review of measurement and construct validity research. In A. D. Ong & M. H. M. van Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 83-95). New York, NY: Oxford University Press.
- Ekstrom, R. B., Goertz, M. E., Pollack, M. M., & Rock, D. A. (1986). Who drops out of high school and why? Findings from a national study. In G. Natriello (Ed.), *School dropouts: Patterns and policies* (pp. 52-69). New York, NY: Teachers College Press.
- Elliot, A., & McGregor, H. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501-519. doi:10.1037//0022-3514.80.3.501
- Ellis, L., Marsh, H. W., & Craven, R. (2009). Addressing the challenges faced by early adolescents: A mixed-method evaluation of the benefits of peer support. *American Journal of Community Psychology, 44*(1), 54-75.
doi:10.1007/s10464-009-9251-y
- Ellis, L. A., Marsh, H. W., & Richards, G. E. (2002). A brief version of the Self Description Questionnaire II. In R. G. Craven, H. W. Marsh, & K. B. Simpson (Eds.), *Self-Concept research: Driving international research agendas: Proceedings of the 2nd Biennial SELF Research Centre International Conference*. Sydney, Australia: SELF Research Centre, University of Western Sydney.
- Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

- Epskamp, S., & Stuber, S. (2017). semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output [Computer software]. Retrieved from <https://CRAN.R-project.org/package=semPlot>
- Ewert, A. (1987). Research in outdoor adventure: Overview and analysis. *The Bradford Papers Annual*, 2, 15-28.
- Ewert, A., Garvey, D., Prouty, D., Panicucci, J., & Collinson, R. (2007). Philosophy and theory of adventure education. In D. Prouty, J. Panicucci, & R. Collinson (Eds.), *Adventure education: Theory and applications* (pp. 19-32). Champaign, IL: Human Kinetics.
- Ewert, A., & McAvoy, L. (2000). The effects of wilderness settings on organized groups: A state-of-knowledge paper. In S. F. McCool, D. N. Cole, W. T. Borrie, & J. O'Loughlin (Eds.), *Wilderness science in a time of change conference-Volume 3: Wilderness as a place for scientific inquiry* (pp. 13-26). Missoula, MT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Ewert, A., & Sibthorp, J. (2009). Creating outcomes through experiential education: The challenge of confounding variables. *Journal of Experiential Education*, 31(3), 376-389. doi:10.1177/105382590803100305
- Ewert, A., & Yoshino, A. (2011). The influence of short-term adventure-based experiences on levels of resilience. *Journal of Adventure Education and Outdoor Learning*, 11(1), 35-50. doi:10.1080/14729679.2010.532986
- Fan, X., & Thompson, B. (2001). Confidence intervals about score reliability coefficients, please: An EPM guidelines editorial. *Educational and Psychological Measurement*, 61(4), 517-531. doi:10.1177/00131640121971365
- Feldman, D. B., & Kubota, M. (2015). Hope, self-efficacy, optimism, and academic achievement: Distinguishing constructs and levels of specificity in predicting college grade-point average. *Learning and Individual Differences*, 37, 210-216. doi:10.1016/j.lindif.2014.11.022
- Fenwick, T. J. (2001). *Experiential learning: A theoretical critique from five perspectives*. Retrieved from Columbus, OH:

- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532-538.
doi:10.1037/a0015808
- Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Los Angeles, CA: SAGE.
- Fox, J., & Weisberg, S. (2011). *An {R} companion to applied regression*. Retrieved from <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>
- Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. Los Angeles, CA: SAGE.
- Gabb, S., Tinberg, H., & Weisberger, R. (2011). Kegan's theory of development applied to community college students. In M. London (Ed.), *The Oxford Handbook of lifelong learning* (pp. 102-116). New York, NY: Oxford University Press.
- Garcia, E. (2015). *Inequalities at the starting gate: Cognitive and noncognitive skills gaps between 2010-2011 kindergarten classmates*. Retrieved from <https://www.epi.org/publication/inequalities-at-the-starting-gate-cognitive-and-noncognitive-gaps-in-the-2010-2011-kindergarten-class/>
- Gass, M., & Gillis, L. (1995). Focusing on the "solution" rather than the "problem": Empowering client change in adventure experiences. *Journal of Experiential Education*, 18(2), 7. doi:10.1177/105382599501800202
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, England: Cambridge University Press.
- Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, 5(2), 189-211. doi:10.1080/19345747.2011.618213
- Gilbertson, K., Bates, T., McLaughlin, T., & Ewert, A. (2006). *Outdoor education: Methods and strategies*. Champaign, IL: Human Kinetics.
- Gillham, J. E., Reivich, K. J., & Shatté, A. J. (2001). *Building optimism and preventing depressive symptoms in children*. Washington, DC: American Psychological Association.

- Gillis, H. L., & Speelman, E. (2008). Are challenge (ropes) courses an effective tool? A meta-analysis. *Journal of Experiential Education*, 31(2), 111-135.
- Gilman, R., Laughlin, J. E., & Huebner, E. S. (1999). Validation of the Self-Description Questionnaire-II with an American sample. *School Psychology International*, 20(3), 300-307. doi:10.1177/0143034399203005
- Glaesmer, H., Rief, W., Martin, A., Mewes, R., Brahler, E., Zenger, M., & Hinz, A. (2012). Psychometric properties and population-based norms of the Life Orientation Test Revised (LOT-R). *British Journal of Health Psychology*, 17(2), 432-445. doi:10.1111/j.2044-8287.2011.02046.x
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. London, England: SAGE.
- Glickman, M. E., Rao, S. R., & Schultz, M. R. (2014). False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *Journal of Clinical Epidemiology*, 67(8), 850-857. doi:10.1016/j.jclinepi.2014.03.012
- Goldstein, S., & Brooks, R. B. (2013). Why study resilience? In S. Goldstein & R. B. Brooks (Eds.), *Handbook of resilience in children* (pp. 3-14). Boston, MA: Springer.
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, 24(6), 645-662. doi:10.1016/j.appdev.2003.09.002
- Gorard, S. (2010). Research design, as independent of methods. In A. Tashakkori & C. Teddlie (Eds.), *SAGE Handbook of mixed methods in social & behavioral research* (2nd ed., pp. 237-252). Thousand Oaks, CA: SAGE. Retrieved from <http://methods.sagepub.com/book/sage-handbook-of-mixed-methods-social-behavioral-research-2e>. doi:10.4135/9781506335193
- Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement*, 66(6), 930-944. doi:10.1177/0013164406288165
- Grant, A. M. (2002). *Towards a psychology of coaching: The impact of coaching on metacognition, mental health and goal attainment*. (Doctoral dissertation).

Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 305462830)

Grant, A. M. (2003). The impact of life coaching on goal attainment, metacognition and mental health. *Social Behavior and Personality, 31*(3), 253-264.

doi:10.2224/sbp.2003.31.3.253

Grant, A. M. (2006). An integrative goal-focused approach to executive coaching. In D. R. Stober & A. M. Grant (Eds.), *Evidence based coaching handbook: Putting best practices to work for your clients* (pp. 153-192). Hoboken, NJ: John Wiley & Sons.

Grant, A. M. (2007). A model of goal striving and mental health for coaching populations. *International Coaching Psychology Review, 2*(3), 248-262.

Grant, A. M. (2012a). The efficacy of coaching. In J. Passmore, D. B. Peterson, & T. Freire (Eds.), *The Wiley-Blackwell handbook of the psychology of coaching and mentoring* (pp. 13-39). Chichester, England: John Wiley & Sons.

Grant, A. M. (2012b). An integrated model of goal-focused coaching: An evidence-based framework for teaching and practice. *International Coaching Psychology Review, 7*(2), 146-165.

Grant, A. M. (2014). Autonomy support, relationship satisfaction and goal focus in the coach-coachee relationship: Which best predicts coaching success? *Coaching: An International Journal of Theory, Research and Practice, 7*(1), 18-38. doi:10.1080/17521882.2013.850106

Grant, A. M. (2016a). The contribution of qualitative research to coaching psychology: Counting numbers is not enough, qualitative counts too. *Journal of Positive Psychology, 12*(3), 317-318. doi:10.1080/17439760.2016.1262616

Grant, A. M. (2016b). What can Sydney tell us about coaching? Research with implications for practice from down under. *Consulting Psychology Journal: Practice and Research, 68*(2), 105. doi:10.1037/cpb0000047

Grant, A. M. (2017). The third 'generation' of workplace coaching: Creating a culture of quality conversations. *Coaching: An International Journal of Theory, Research and Practice, 10*(1), 37-53.

doi:10.1080/17521882.2016.1266005

- Grant, A. M., & Cavanagh, M. J. (2007). Evidence-based coaching: Flourishing or languishing? *Australian Psychologist*, *42*(4), 239-254.
doi:10.1080/00050060701648175
- Grant, A. M., & Cavanagh, M. J. (2011). Coaching and positive psychology. In K. M. Sheldon, T. B. Kashdan, & M. F. Steger (Eds.), *Designing positive psychology: Taking stock and moving forward*. (pp. 293-309). New York, NY: Oxford University Press.
- Grant, A. M., Curtayne, L., & Burton, G. (2009). Executive coaching enhances goal attainment, resilience and workplace well-being: A randomised controlled study. *Journal of Positive Psychology*, *4*(5), 396-407.
doi:10.1080/17439760902992456
- Grant, A. M., Passmore, J., Cavanagh, M. J., & Parker, H. M. (2010). The state of play in coaching today: A comprehensive review of the field. *International Review of Industrial and Organizational Psychology*, *25*(1), 125-167.
doi:10.1002/9780470661628.ch4
- Grant, A. M., & Stober, D. R. (2006). Introduction. In D. R. Stober & A. M. Grant (Eds.), *Evidence based coaching handbook* (pp. 1-14). Hoboken, NJ: Wiley.
- Green, J., Martin, A. J., & Marsh, H. W. (2007). Motivation and engagement in English, mathematics and science high school subjects: Towards an understanding of multidimensional domain specificity. *Learning and Individual Differences*, *17*(3), 269-279. doi:10.1016/j.lindif.2006.12.003
- Green, S. L., Grant, A. M., & Rynsaardt, J. (2007). Evidence-based life coaching for senior high school students: Building hardiness and hope. *International Coaching Psychology Review*, *2*(1), 24-32.
- Green, S. L., Oades, L. G., & Grant, A. M. (2006). Cognitive-behavioral, solution-focused life coaching: Enhancing goal striving, well-being, and hope. *Journal of Positive Psychology*, *1*(3), 142-149. doi:10.1080/17439760600619849
- Green, S. L., Oades, L. G., & Robinson, P. L. (2012). Positive education programmes: Integrating coaching and positive psychology in schools. In C. van Nieuwerburgh (Ed.), *Coaching in education: Getting better results for students, educators, and parents* (pp. 115-132). London, England: Karnac Books.

- Grover, S., & Furnham, A. (2016). Coaching as a Developmental Intervention in Organisations: A Systematic Review of Its Effectiveness and the Mechanisms Underlying It. *PLoS One*, *11*(7), e0159137. doi:10.1371/journal.pone.0159137
- Gutman, L., & Schoon, I. (2013). *The impact of non-cognitive skills on outcomes for young people: Literature review*. Retrieved from http://educationendowmentfoundation.org.uk/uploads/pdf/Non-cognitive_skills_literature_review.pdf
- Hahn, K. (1957). *Outward bound*. Paper presented at the Address given to a conference at Harrogate. Outward Bound Trust.
- Hallquist, M., & Wiley, J. (2017). MplusAutomation: Automating Mplus Model Estimation and Interpretation [Computer software]. Retrieved from <https://CRAN.R-project.org/package=MplusAutomation>
- Hamoudi, A., Murray, D. W., Sorensen, L., & Fontaine, A. (2015). *Self-regulation and toxic stress: A review of ecological, biological, and developmental studies of self-regulation and stress (OPRE Report #2015-30)*. Retrieved from <https://www.acf.hhs.gov/opre/resource/self-regulation-and-toxic-stress-a-review-of-ecological-biological-and-developmental-studies-of-self-regulation-and-stress>
- Hans, T. A. (2000). A meta-analysis of the effects of adventure programming on locus of control. *Journal of Contemporary Psychotherapy*, *30*(1), 33-60. doi:10.1023/A:1003649031834
- Harper, N. (2010). Future paradigm or false idol: A cautionary tale of evidence-based practice for adventure education and therapy. *The Journal of Experiential Education*, *33*(1), 38-55. doi:10.5193/JEE.33.1.38
- Hatch, K., & McCarthy, C. (2005). Exploration of challenge courses' long-term effects on members of college student organizations. *The Journal of Experiential Education*, *27*(3), 245-264.
- Hattie, J., Marsh, H. W., Neill, J. T., & Richards, G. E. (1997). Adventure education and Outward Bound: Out-of-class experiences that make a lasting difference. *Review of Educational Research*, *67*(1), 43-87. doi:10.2307/1170619
- Hayhurst, J., Hunter, J. A., Kafka, S., & Boyes, M. (2015). Enhancing resilience in youth through a 10-day developmental voyage. *Journal of Adventure*

- Education and Outdoor Learning*, 15(1), 40-52.
doi:10.1080/14729679.2013.843143
- Heaven, P., & Ciarrochi, J. (2008). Parental styles, gender and the development of hope and self-esteem. *European Journal of Personality*, 22(8), 707-724.
doi:10.1002/per.699
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900-1902.
- Helmes, E., Holden, R. R., & Ziegler, M. (2015). Response bias, malingering, and impression management. In G. J. Boyle, D. H. Saklofske, & G. Matthews (Eds.), *Measures of personality and social psychological constructs* (pp. 16-43). London, England: Elsevier.
- Helsing, D., Broderick, M., & Hammerman, J. (2001). A developmental view of ESOL students' identity transitions in an urban community college. In The Adult Development Research Group (Ed.), *Research Monograph: NCSALL Reports #19: Toward a new pluralism in ABE/ESOL classrooms: Teaching to multiple "cultures of mind"*. Cambridge, MA: National Center for the Study of Adult Learning and Literacy.
- Helsing, D., Drago-Severson, E., Kegan, R., Portnow, K., Popp, N., & Broderick, M. (2001). Three different types of change. *Focus on Basics*, 5, 10-14.
- Henderson, L. (2009). *Necessary solitude: Solo, gender and personal growth*. (Master's thesis). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 1465616)
- Henderson, L. W., Knight, T., & Richardson, B. (2013). An exploration of the well-being benefits of hedonic and eudaimonic behaviour. *Journal of Positive Psychology*, 8(4), 322-336. doi:10.1080/17439760.2013.803596
- Herzberg, P. Y., Glaesmer, H., & Hoyer, J. (2006). Separating optimism and pessimism: A robust psychometric analysis of the Revised Life Orientation Test (LOT-R). *Psychological Assessment*, 18(4), 433-438. doi:10.1037/1040-3590.18.4.433
- Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice*. New York, NY: Guilford Press.

- Hinz, A., Sander, C., Glaesmer, H., Brähler, E., Zenger, M., Hilbert, A., & Kocalevent, R.-D. (2017). Optimism and pessimism in the general population: Psychometric properties of the Life Orientation Test (LOT-R). *International Journal of Clinical and Health Psychology, 17*(2), 161-170.
doi:10.1016/j.ijchp.2017.02.003
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140. doi:10.3102/00346543067001088
- Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2003). Reliability methods: A note on the frequency of use of various types. In B. Thompson (Ed.), *Score reliability: Contemporary thinking on reliability issues* (pp. 59-68). Thousand Oaks, CA: SAGE. Retrieved from <http://methods.sagepub.com/book/score-reliability>.
doi:10.4135/9781412985789
- Hone, L. C., Jarden, A., Schofield, G., & Duncan, S. (2014). Measuring flourishing: The impact of operational definitions on the prevalence of high levels of wellbeing. *International Journal of Wellbeing, 4*(1), 62-90.
doi:10.5502/ijw.v4i1.4
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*(3), 117-144.
doi:10.1080/03610739208253916
- Howard, S., & Johnson, B. (2000). What makes the difference? Children and teachers talk about resilient outcomes for children at risk. *Educational Studies, 26*(3), 321-337. doi:10.1080/03055690050137132
- Howell, A. J. (2009). Flourishing: Achievement-related correlates of students' well-being. *Journal of Positive Psychology, 4*(1), 1-13.
doi:10.1080/17439760802043459
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55.
doi:10.1080/10705519909540118
- Huble, A. M., & Zumbo, B. D. (2013). Psychometric characteristics of assessment procedures: An overview. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J. I.

- C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology* (pp. 3-19). Washington, DC: American Psychological Association.
- Hunter, J. A., Kafka, S., Hayhurst, J., Clark, H., Dickerson, D., Harold, G., . . . Stringer, M. (2010). Increased self-efficacy following a ten-day developmental voyage. *Journal of Child & Adolescent Mental Health, 22*(1), 63-65. doi:10.2989/17280583.2010.496943
- Hunter, S. C., Houghton, S., & Wood, L. (2015). Positive mental well-being in Australian adolescents: Evaluating the Warwick-Edinburgh Mental Well-Being Scale. *Australian Educational and Developmental Psychologist, 32*(2), 93-104. doi:10.1017/edp.2015.12
- Huppert, F. A. (2004). A population approach to positive psychology: The potential for population interventions to promote well-being and prevent disorder. In P. A. Linley & S. Joseph (Eds.), *Positive psychology in practice* (pp. 693-709). Hoboken, NJ: Wiley.
- Huppert, F. A. (2009). Psychological well-being: Evidence regarding its causes and consequences. *Applied Psychology: Health and Well-Being, 1*(2), 137-164. doi:10.1111/j.1758-0854.2009.01008.x
- Huppert, F. A., & So, T. T. (2013). Flourishing across Europe: Application of a new conceptual framework for defining well-being. *Social Indicators Research, 110*(3), 837-861.
- Huppert, F. A., & So, T. T. C. (2009). What percentage of people in Europe are flourishing and what characterises them? *Paper prepared for OECD/ISQOLS meeting: Measuring subjective well-being: An opportunity for NSOs?* Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.550.8290&rep=rep1&type=pdf>
- Huta, V. (2013). Pursuing eudaimonia versus hedonia: Distinctions, similarities, and relationships. In A. S. Waterman (Ed.), *The best within us: Positive psychology perspectives on eudaimonia* (pp. 139-158). Washington, DC: American Psychological Association.

- Huta, V., & Ryan, R. M. (2010). Pursuing pleasure or virtue: The differential and overlapping well-being benefits of hedonic and eudaimonic motives. *Journal of Happiness Studies, 11*(6), 735-762. doi:10.1007/s10902-009-9171-4
- Huta, V., & Waterman, A. S. (2014). Eudaimonia and its distinction from hedonia: Developing a classification and terminology for understanding conceptual and operational definitions. *Journal of Happiness Studies, 15*(6), 1425-1456. doi:10.1007/s10902-013-9485-0
- Iacobucci, D., & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. *Journal of Consumer Psychology, 13*(4), 478-487. doi:10.1207/S15327663JCP1304_14
- Imholt, R. D. (2009). *The impact of adolescent expatriate participation in experiential education on locus of control*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses global database. (UMI No. 305077033)
- Inman, S., & Buck, M. (1998). Learning outcomes for personal and social development. In S. Inman, M. Buck, & H. Burke (Eds.), *Assessing personal and social development: Measuring the unmeasurable?* (pp. 49-64). London, England: Falmer Press.
- Jackson, D. L., Gillaspay, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods, 14*(1), 6-23. doi:10.1037/a0014694
- Johnson, J. W. (2012). *The effect of high school outdoor-based adventure leadership programs in independent schools on personal effectiveness and locus of control*. (Master's thesis). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 1510668)
- Jones, R. J., Woods, S. A., & Guillaume, Y. R. F. (2016). The effectiveness of workplace coaching: A meta-analysis of learning and performance outcomes from coaching. *Journal of Occupational and Organizational Psychology, 89*(2), 249-277. doi:10.1111/joop.12119
- Jostad, J., Paisley, K., Sibthorp, J., & Gookin, J. (2013). The multi-dimensionality of group cohesion: a social network analysis of NOLS courses.(Report). *Journal*

- of Outdoor Recreation, Education and Leadership*, 5(2), 131. doi:10.7768/1948-5123.1211
- Jostad, J., Sibthorp, J., Pohja, M., & Gookin, J. (2015). The adolescent social group in outdoor adventure education: Social connections that matter. *Research in Outdoor Education*, 13(1), 16-37. doi:10.1353/roe.2015.0002
- Kalisch, K. R., Bobilya, A. J., & Daniel, B. (2011). The Outward Bound solo: A study of participants and perceptions. *Journal of Experiential Education*, 34(1), 1-18. doi:10.5193/JEE34.1.1
- Kashdan, T. B., Biswas-Diener, R., & King, L. A. (2008). Reconsidering happiness: The costs of distinguishing between hedonics and eudaimonia. *Journal of Positive Psychology*, 3(4), 219-233. doi:10.1080/17439760802303044
- Kayes, D. C. (2002). Experiential Learning and Its Critics: Preserving the Role of Experience in Management Learning and Education. *Academy of Management Learning & Education*, 1(2), 137-149. doi:10.5465/AMLE.2002.8509336
- Kegan, R. (1982). *The evolving self*. Cambridge, MA: Harvard University Press.
- Kegan, R. (1994). *In over our heads: The mental demands of modern life*. Cambridge, MA: Harvard University Press.
- Kegan, R. (2000). What “form” transforms? A constructive-developmental approach to transformative learning. In J. Mezirow (Ed.), *Learning as transformation: Critical perspectives on a theory in progress* (1st ed.). San Francisco, CA: Jossey-Bass.
- Kelley, K. (2018). MBESS: The MBESS R Package [Computer software]. Retrieved from <https://CRAN.R-project.org/package=MBESS>
- Kelley, K., & Pornprasertmanit, S. (2016). Confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. *Psychological Methods*, 21(1), 69-92. doi:10.1037/a0040086
- Kelly, J. (2019). Influence of outdoor and adventure activities on subjective measures of resilience in university students. *Journal of Experiential Education*, 16. doi:10.1177/1053825919831724

- Kemp, T. J. (2006). An adventure-based framework for coaching. In D. R. Stober & A. M. Grant (Eds.), *Evidence based coaching handbook: Putting best practices to work for your clients*. Hoboken, NJ: Wiley.
- Kemp, T. J. (2008). Searching for the elusive model of coaching: Could the “Holy Grail” be right in front of us. *International Coaching Psychology Review*, 3(3), 219-226.
- Keyes, C. L. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior*, 43(2), 207-222.
doi:10.2307/3090197
- Keyes, C. L., & Annas, J. (2009). Feeling good and functioning well: Distinctive concepts in ancient philosophy and contemporary science. *Journal of Positive Psychology*, 4(3), 197-201. doi:10.1080/17439760902844228
- Keyes, C. L., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: The empirical encounter of two traditions. *Journal of Personality and Social Psychology*, 82(6), 1007-1022. doi:10.1037/0022-3514.82.6.1007
- Keyes, C. L. M. (2003). Complete mental health: An agenda for the 21st century. In C. L. M. Keyes & J. Haidt (Eds.), *Flourishing: Positive psychology and the life well-lived* (pp. 293-312). Washington, DC: American Psychological Association.
- Kirk, J. (1986). *Reliability and validity in qualitative research*. Newbury Park, CA: SAGE.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford Press.
- Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. Retrieved from <http://methods.sagepub.com/book/psychological-testing>
doi:10.4135/9781483385693
- Klint, K. A. (1990). New directions for inquiry into self-concept and adventure experiences. In J. Miles & S. Priest (Eds.), *Adventure education* (pp. 163-171). State College, PA: Venture.
- Kolb, D. [David] A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

- Kolb, D. [David] A. (2014). *Experiential learning: Experience as the source of learning and development* (2nd ed.). Upper Saddle River, NJ: Pearson Education.
- Kolb, D. [Daryl] G. (1991). Meaningful methods: Evaluation without the crunch. *Journal of Experiential Education*, 14(1), 40-44.
doi:10.1177/105382599101400108
- Lahey, L., Souvaine, E., Kegan, R., Goodman, R., & Felix, S. (2011). *A guide to the subject-object interview: Its administration and interpretation*. Cambridge, MA: Minds at Work.
- Lane, J. (2008). *The effectiveness of an adventure travel summer camp program on the life effectiveness of adolescents*. (Master's thesis). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 1459504)
- Larkin, M., Watts, S., & Clifton, E. (2006). Giving voice and making sense in interpretative phenomenological analysis. *Qualitative Research in Psychology*, 3(2), 102-120. doi:10.1191/1478088706qp062oa
- Laske, O. (2007). Contributions of evidence-based developmental coaching to coaching psychology and practice. *International Coaching Psychology Review*, 2(2), 202-212.
- Latham, G. P., & Brown, T. C. (2006). The effect of learning vs. outcome goals on self-efficacy, satisfaction and performance in an MBA program. *Applied Psychology*, 55(4), 606-623. doi:10.1111/j.1464-0597.2006.00246.x
- Latham, G. P., & Locke, E. A. (1979). Goal setting: A motivational technique that works. *Organizational Dynamics*, 8(2), 68-80. doi:10.1016/0090-2616(79)90032-9
- Lawrence, D., Johnson, S., Hafekost, J., Bortherhoven De Hann, K., Sawyer, M., Ainley, J., & Zubrick, S. R. (2015). *The mental health of children and adolescents: Report on the second Australian child and adolescent survey of mental health and wellbeing*. Retrieved from <https://www.health.gov.au/internet/main/publishing.nsf/Content/9DA8CA21306FE6EDCA257E2700016945/%24File/child2.pdf>
- Leberman, S. I., & Martin, A. J. (2004). Enhancing transfer of learning through post-course reflection. *Journal of Adventure Education and Outdoor Learning*, 4(2), 173-184.

- Leung, K. C., Marsh, H. W., Yeung, A. S., & Abduljabbar, A. S. (2015). Validity of social, moral and emotional facets of Self-Description Questionnaire II. *Journal of Experimental Education, 83*(1), 1-23.
doi:10.1080/00220973.2013.876229
- Lewis, P., Forsythe, G., Sweeney, P., Bartone, P., Bullis, C., & Snook, S. (2005). Identity development during the college years: Findings from the West Point longitudinal study. *Journal of College Student Development, 46*(4), 357-373.
doi:10.1353/csd.2005.0037
- Linley, P. A. (2006). Coaching research: Who? what? where? when? why. *International Journal of Evidence Based Coaching and Mentoring, 4*(2), 1-7.
- Linley, P. A., Maltby, J., Wood, A. M., Osborne, G., & Hurling, R. (2009). Measuring happiness: The higher order factor structure of subjective and psychological well-being measures. *Personality and Individual Differences, 47*(8), 878-884.
doi:10.1016/j.paid.2009.07.010
- Little, T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guildford Press.
- Lloyd, K., & Devine, P. (2012). Psychometric properties of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) in Northern Ireland. *Journal of Mental Health, 2012, Vol.21*(3), p.257-263, 21(3), 257-263.
doi:10.3109/09638237.2012.670883
- Locke, E. A. (1996). Motivation through conscious goal setting. *Applied and Preventive Psychology, 5*(2), 117-124.
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Englewood Cliffs, NJ: Prentice-Hall.
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. *American Psychologist, 57*(9), 705-717.
doi:10.1037/0003-066X.57.9.705
- Locke, E. A., & Latham, G. P. (2009). Has goal setting gone wild, or have its attackers abandoned good scholarship. *Academy of Management Perspectives, 23*(1), 17-23. doi:10.5465/AMP.2009.37008000
- Lopez, S. J., Ciarlelli, R., Coffman, L., Stone, M., & Wyatt, L. (2000). Diagnosing for strengths: On measuring hope building blocks. In C. R. Snyder (Ed.),

- Handbook of hope: Theory, measures, and applications*. San Diego, CA: Academic Press.
- Louv, R. (2005). *Last child in the woods: Saving our children from nature deficit disorder*. Chapel Hill, NC: Algonquin Books.
- Lowman, R. L. (2005). Executive coaching: The road to Dodoville needs paving with more than good assumptions. *Consulting Psychology Journal: Practice and Research*, 57(1), 90-96. doi:10.1037/1065-9293.57.1.90
- Lubke, G. H., & Muthen, B. O. (2004). Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural Equation Modeling*, 11(4), 514-534.
- Luckner, J. L., & Nadler, R. S. (1997). *Processing the experience: Strategies to enhance and generalize learning* (2nd ed.). Dubuque, IA: Kendall/Hunt.
- Luo, Y. (2011). *Outward Bound outcome model validation and multilevel modeling*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3491492)
- MacCann, C., & Roberts, R. D. (2010). Do time management, grit, and self-control relate to academic achievement independently of conscientiousness? In R. Hicks (Ed.), *Personality and individual differences: Current directions* (pp. 79-90). Queensland, Australia: Australian Academic Press.
- MacCann, C., Ziegler, M., & Roberts, R. D. (2012). Faking in personality assessment: Reflections and recommendations. In M. Ziegler, C. MacCann, & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 309-329). New York, NY: Oxford University Press.
- Madden, W., Green, S. L., & Grant, A. M. (2011). A pilot study evaluating strengths-based coaching for primary school students: Enhancing engagement and hope. *International Coaching Psychology Review*, 6(1), 71-83.
- Magaletta, P. R., & Oliver, J. M. (1999). The hope construct, will, and ways: Their relations with self-efficacy, optimism, and general well-being. *Journal of Clinical Psychology*, 55(5), 539-551. doi:10.1002/(SICI)1097-4679(199905)55:5<539::AID-JCLP2>3.0.CO;2-G
- Maheswaran, H., Weich, S., Powell, J., & Stewart-Brown, S. (2012). Evaluating the responsiveness of the Warwick Edinburgh Mental Well-Being Scale

- (WEMWBS): Group and individual level analysis. *Health and Quality of Life Outcomes*, 10(1). doi:10.1186/1477-7525-10-156
- Maller, C., Townsend, M., St Leger, L., Pryor, A., Prosser, L., & Moore, M. (2008). *Healthy parks, healthy people: The health benefits of contact with nature in a park context* (2nd ed.). Melbourne, Australia: Deakin University.
- Marsh, H. W. (1986a). Negative item bias in ratings scales for preadolescent children: A cognitive-developmental phenomenon. *Developmental Psychology*, 22(1), 37-49. doi:10.1037/0012-1649.22.1.37
- Marsh, H. W. (1986b). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23(1), 129-149. doi:10.3102/00028312023001129
- Marsh, H. W. (1988). *Self-Description Questionnaire: A theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept—A test manual and research monograph*. San Antonio, TX: The Psychological Corporation.
- Marsh, H. W. (1990a). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology*, 82(4), 646-656. doi:10.1037/0022-0663.82.4.646
- Marsh, H. W. (1990b). Confirmatory factor analysis of multitrait-multimethod data: The construct validation of multidimensional self-concept responses. *Journal of personality*, 58(4), 661-692.
- Marsh, H. W. (1990c). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, 2(2), 77-172. doi:10.1007/BF01322177
- Marsh, H. W. (1992a). Extracurricular activities: Beneficial extension of the traditional curriculum or subversion of academic goals? *Journal of Educational Psychology*, 84(4), 553-562. doi:10.1037/0022-0663.84.4.553
- Marsh, H. W. (1992b). *Self-Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: A test manual and research monograph*. Sydney, Australia: University of Western Sydney.

- Marsh, H. W. (1992c). *Self-Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: A test manual and research monograph*. Sydney, Australia: University of Western Sydney.
- Marsh, H. W. (1993). Academic self-concept: Theory measurement and research. In J. Suls (Ed.), *Psychological perspectives on the self: The self in social perspective* (Vol. 4, pp. 59-98). Hillsdale, NJ: Erlbaum.
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, 70(4), 810-819. doi:10.1037/0022-3514.70.4.810
- Marsh, H. W. (2005). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Paper presented at the Annual Meeting of the Education Section of The British Psychological Society: 25th Vernon-Wall Lecture, Durham University, Durham, UK.
- Marsh, H. W., & Balla, J. R. (1994). Goodness of fit in confirmatory factor analysis: The effects of sample size and model parsimony. *International Journal of Methodology*, 28(2), 185-217. doi:10.1007/BF01102761
- Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103(3), 391-410. doi:10.1037/0033-2909.103.3.391
- Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, 80(3), 366-380. doi:10.1037/0022-0663.80.3.366
- Marsh, H. W., & Craven, R. (1997). Academic self-concept: Beyond the dustbowl. In G. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 131-198). San Diego, CA: Academic Press.
- Marsh, H. W., Craven, R., & Martin, A. J. (2006). What is the nature of self-esteem? Unidimensional and multidimensional perspectives. In M. Kernis (Ed.), *Self-esteem issues and answers: A sourcebook of current perspectives* (pp. 16-24). New York, NY: Psychology Press.

- Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G., & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment, 17*(1), 81-102. doi:10.1037/1040-3590.17.1.81
- Marsh, H. W., & Gouvenet, P. J. (1989). Multidimensional self-concepts and perceptions of control: Construct validation of responses by children. *Journal of Educational Psychology, 81*(1), 57-69. doi:10.1037/0022-0663.81.1.57
- Marsh, H. W., & Hattie, J. (1996). Theoretical perspectives on the structure of self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 38-90). New York, NY: Wiley.
- Marsh, H. W., & Hau, K.-T. (1996). Assessing goodness of fit: Is parsimony always desirable? *Journal of Experimental Education, 64*(4), 364-390.
- Marsh, H. W., & Hau, K.-T. (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology, 32*(1), 151-170.
- Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary psychometrics: A Festschrift for Roderick P. McDonald* (pp. 225-340). Mahwah, NJ: Erlbaum.
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320-341. doi:10.1207/s15328007sem1103_2
- Marsh, H. W., & Kleitman, S. (2002). Extracurricular school activities: The good, the bad, and the nonlinear. *Harvard Educational Review, 72*(4), 464-514. doi:10.17763/haer.72.4.051388703v7v7736
- Marsh, H. W., Liem, G. A. D., Martin, A. J., Morin, A. J. S., & Nagengast, B. (2011). Methodological measurement fruitfulness of exploratory structural equation modeling (ESEM): New approaches to key substantive issues in motivation and engagement. *Journal of Psychoeducational Assessment, 29*(4), 322-346. doi:10.1177/0734282911406657

- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22*(3), 471-491. doi:10.1037/a0019227
- Marsh, H. W., & Martin, A. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*(1), 59-77. doi:10.1348/000709910X503501
- Marsh, H. W., Martin, A. J., & Jackson, S. A. (2010). Introducing a short version of the physical self description questionnaire: New strategies, short-form evaluative criteria, and applications of factor analyses. *Journal of Sport and Exercise Psychology, 32*(4), 438-482. doi:10.1123/jsep.32.4.438
- Marsh, H. W., Martin, A. J., Yeung, A. S., & Craven, R. (2017). Competence of self-perceptions. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation: Theory and application* (2nd ed., pp. 85-115). New York: NY: Guilford Press.
- Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014a). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85-110. doi:10.1146/annurev-clinpsy-032813-153700
- Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014b). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis: Supplemental material. *Annual Review of Clinical Psychology, 10*, 85-110. Retrieved from http://www.annualreviews.org/doi/suppl/10.1146/annurev-clinpsy-032813-153700/suppl_file/cp10_marshall_supmat.pdf
- Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling: A Multidisciplinary Journal, 16*(3), 439-476. doi:10.1080/10705510903008220
- Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity,

- maturity, and la dolce vita effects. *Developmental Psychology*, 49(6), 1194-1218. doi:10.1037/a0026913
- Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology*, 111(2), 331-353. doi:10.1037/edu0000281
- Marsh, H. W., Richards, G. E., & Barnes, J. (1986a). Multidimensional self-concepts: A long-term follow-up of the effect of participation in an Outward Bound program. *Personality and Social Psychology Bulletin*, 12(4), 475-492. doi:10.1177/0146167286124011
- Marsh, H. W., Richards, G. E., & Barnes, J. (1986b). Multidimensional self-concepts: The effect of participation in an Outward Bound Program. *Journal of Personality and Social Psychology*, 50(1), 195-204. doi:10.1037/0022-3514.50.1.195
- Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology*, 89(1), 41-54. doi:10.1037/0022-0663.89.1.41
- Marsh, H. W., & Yeung, A. S. (1998). Top-down, bottom-up, and horizontal models: The direction of causality in multidimensional, hierarchical self-concept models. *Journal of Personality and Social Psychology*, 75(2), 509. doi:10.1037/0022-3514.75.2.509
- Martin, A. J. (2001). The Student Motivation Scale: A tool for measuring and enhancing motivation. *Australian Journal of Guidance and Counselling*, 11, 1-20. doi:10.1017/S1037291100004301
- Martin, A. J. (2002). Motivation and academic resilience: Developing a model for student enhancement. *Australian Journal of Education*, 46(1), 34-49. doi:10.1177/000494410204600104
- Martin, A. J. (2003). The Student Motivation Scale: Further testing of an instrument that measures school students and motivation. *Australian Journal of Education*, 47(1), 88-106. doi:10.1177/000494410304700107

- Martin, A. J. (2005). Exploring the effects of a youth enrichment program on academic motivation and engagement. *An International Journal*, 8(2), 179-206. doi:10.1007/s11218-004-6487-0
- Martin, A. J. (2007). Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology*, 77(2), 413-440. doi:10.1348/000709906X118036
- Martin, A. J. (2009). Motivation and engagement across the academic life span: A developmental construct validity study of elementary school, high school, and university/college students. *Educational and Psychological Measurement*, 69(5), 794-824. doi:10.1177/0013164409332214
- Martin, A. J. (2013). Academic buoyancy and academic resilience: Exploring everyday and classic resilience in the face of academic adversity. *School Psychology International*, 34(5), 488-500. doi:10.1177/0143034312472759
- Martin, A. J., Colmar, S. H., Davey, L. A., & Marsh, H. W. (2010). Longitudinal modelling of academic buoyancy and motivation: Do the 5Cs hold up over time? *The British journal of educational psychology*, 80(3), 473-496. doi:10.1348/000709910X486376
- Martin, A. J., & Hau, K.-T. (2010). Achievement motivation among Chinese and Australian school students: Assessing differences of kind and differences of degree. *International Journal of Testing*, 10(3), 274-294. doi:10.1080/15305058.2010.482220
- Martin, A. J., & Leberman, S. I. (2005). Personal learning or prescribed educational outcomes: A case study of the Outward Bound experience. *Journal of Experiential Education*, 28(1), 44-59. doi:10.1177/105382590502800106
- Martin, A. J., & Legg, S. (2002). Investigating the inward sounds of Outward Bound. *Australian Journal of Outdoor Education*, 6(2), 27-36.
- Martin, A. J., Malmberg, L.-E., & Liem, G. A. D. (2010). Multilevel motivation and engagement: Assessing construct validity across students and schools. *Educational and Psychological Measurement*, 70(6), 973-989. doi:10.1177/0013164410378089

- Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools, 43*(3), 267-281. doi:10.1002/pits.20149
- Martin, A. J., & Marsh, H. W. (2008). Academic buoyancy: Towards an understanding of students' everyday academic resilience. *Journal of School Psychology, 46*(1), 53-83. doi:10.1016/j.jsp.2007.01.002
- Martin, A. J., & Marsh, H. W. (2009). Academic resilience and academic buoyancy: multidimensional and hierarchical conceptual framing of causes, correlates and cognate constructs. *Oxford Review of Education, 35*(3), 353-370. doi:10.1080/03054980902934639
- Marx, R. W., & Winne, P. H. (1978). Construct interpretations of three self-concept inventories. *American Educational Research Journal, 15*(1), 99-109. doi:10.3102/00028312015001099
- Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist, 56*(3), 227-238.
- Masten, A. S. (2014). Global perspectives on resilience in children and youth. *Child Development, 85*(1), 6-20. doi:10.1111/cdev.12205
- Maxted, J. (2005). Coming home: Adolescents and the nature-based solo. In C. Knapp & T. Smith (Eds.), *Exploring the power of solo, silence, and solitude* (pp. 121-136). Boulder, CO: Association for Experiential Education.
- McAvoy, L. H., Mitten, D. S., Stringer, L. A., Steckart, J. P., & Sproles, K. (1996). Group development and group dynamics in outdoor education. In L. H. McAvoy, L. A. Stringer, M. D. Bialeschki, & A. B. Young (Eds.), *Coalition for Education in the Outdoors: Third Research Symposium Proceedings (January 12-14, 1996, Bradford Woods, Indiana)* (pp. 51-62). Cortland, NY: Coalition for Education in the Outdoors.
- McCann, T. (2005). Adolescent experience 'in over our heads'. In P. Heywood, T. McCann, B. Neville, & P. Willis (Eds.), *Towards re-enchantment: Education, imagination and the getting of wisdom* (pp. 61-70). Queensland, Australia: Post Pressed.
- McCloskey, D. N. (2010). *Cult of statistical significance: How the standard error costs us jobs, justice, and lives*. Ann Arbor, MI: University of Michigan Press.

- McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51*(6), 882-888.
doi:10.1037/0022-006X.51.6.882
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- McGowan, A. L. (2016). Impact of one-semester outdoor education programs on adolescent perceptions of self-authorship. *Journal of Experiential Education, 39*(4), 386-411. doi:10.1177/10538259166668902
- McInerney, D. M., Roche, L. A., McInerney, V., & Marsh, H. W. (1997). Cultural perspectives on school motivation: The relevance and application of goal theory. *American Educational Research Journal, 34*(1), 207-236.
doi:10.3102/00028312034001207
- McKay, M. T., & Andretta, J. R. (2017). Evidence for the psychometric validity, internal consistency and measurement invariance of Warwick Edinburgh Mental Well-being Scale scores in Scottish and Irish adolescents. *Psychiatry Research, 255*, 382-386. doi:10.1016/j.psychres.2017.06.071
- McKenzie, M. D. (2000). How are adventure education program outcomes achieved?: A review of the literature. *Australian Journal of Outdoor Education, 5*(1), 19-27.
- McKenzie, M. D. (2003). Beyond "the Outward Bound Process": Rethinking student learning. *Journal of Experiential Education, 26*(1), 8-23.
doi:10.1177/105382590302600104
- McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*(1), 114-140. doi:10.1037/met0000078
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568-592. doi:10.1037/0021-9010.93.3.568
- Melnick, S. A., & Gable, R. K. (1990). The use of negative item stems: A cautionary note. *Educational Research Quarterly, 14*(3), 31-36.
- Merrell, B. (2009). *The effect of LIFT on life effectiveness and locus of control*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3400787)

- Midkiff, B., Langer, M., Demetriou, C., & Panter, A. T. (2017). Measuring grit among first-generation college students: A psychometric analysis. In L. A. Van der Ark, M. Wiberg, S. A. Culpepper, J. A. Douglas, & W. C. Wang (Eds.), *Quantitative psychology: The 81st annual meeting of the Psychometric Society, Asheville, North Carolina, 2016* (pp. 407-420). New York, NY: Springer.
- Miettinen, R. (2000). The concept of experiential learning and John Dewey's theory of reflective thought and action. *International Journal of Lifelong Education*, *19*(1), 54-72. doi:10.1080/026013700293458
- Mirkin, B. J., & Middleton, M. J. (2014). The social climate and peer interaction on outdoor courses. *Journal of Experiential Education*, *37*(3), 232-247. doi:10.1177/1053825913498370
- Mishler, E. G. (1991). *Research interviewing: Context and narrative*. Cambridge, MA: Harvard University Press.
- Moilanen, K. L. (2007). The Adolescent Self-Regulatory Inventory: The development and validation of a questionnaire of short-term and long-term self-regulation. *Journal of Youth and Adolescence*, *36*(6), 835-848. doi:10.1007/s10964-006-9107-9
- Moilanen, K. L. (2011). *The Adolescent Self-Regulatory Inventory scoring instructions*. Provided by the author.
- Moilanen, K. L. (2014). Short- and long-term self-regulation and sexual risk-taking behaviors in unmarried heterosexual young adults. *Journal of Sex Research*, *52*(7), 1-12. doi:10.1080/00224499.2014.959881
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). Hoboken, NJ: Wiley.
- Monzani, D., Steca, P., & Greco, A. (2014). Brief report: Assessing dispositional optimism in adolescence – Factor structure and concurrent validity of the Life Orientation Test – Revised. *Journal of Adolescence*, *37*(2), 97-101. doi:10.1016/j.adolescence.2013.11.006
- Moon, C., & Snyder, C. R. (2000). The Children's Hope Scale. In J. Maltby, C. A. Lewis, & A. Hill (Eds.), *Commissioned reviews of 250 psychological tests* (Vol. 1, pp. 160-166). Lewiston, NY: Edwin Mellen Press.

- Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling, 23*(1), 116-139. doi:10.1080/10705511.2014.961800
- Murray, D. W., Rosanbalm, K., & Christopoulos, C. (2016). *Self-regulation and toxic stress report 3: A comprehensive review of self-regulation interventions from birth through young adulthood (OPRE Report #2016-34)*. Retrieved from https://www.acf.hhs.gov/sites/default/files/opre/acf_report_3_approved_fromword_b508.pdf
- Murray, D. W., Rosanbalm, K., Christopoulos, C., & Hamoudi, A. (2015). *Self-regulation and toxic stress: Foundations for understanding self-regulation from an applied developmental perspective (OPRE Report #2015-21)*. Retrieved from https://www.acf.hhs.gov/sites/default/files/opre/report_1_foundations_paper_final_012715_submitted_508.pdf
- Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.
- Mygind, L., Kjeldsted, E., Hartmeyer, R., Mygind, E., Bølling, M., & Bentsen, P. (2019). Mental, physical and social health benefits of immersive nature-experience for children and adolescents: A systematic review and quality assessment of the evidence. *Health and Place, 58*, 102136. doi:10.1016/j.healthplace.2019.05.014
- Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology, 15*(6), 1044-1045.
- Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi* (pp. 239-263). Dordrecht, Netherlands: Springer.
- Neill, J. T. (2007, April 26, 2007). Factors which influence the effects of outdoor education programs. *Guide to outdoor education research & evaluation*. Retrieved from <http://www.wilderdom.com/research/researchfactors.html>

- Neill, J. T. (2008). *Enhancing life effectiveness: The impacts of outdoor education programs*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 10309583)
- Neill, J. T., & Dias, K. L. (2001). Adventure education and resilience: The double-edged sword. *Journal of Adventure Education and Outdoor Learning*, 1(2), 35-42. doi:10.1080/14729670185200061
- Neill, J. T., Marsh, H. W., & Richards, G. E. (2003). The Life Effectiveness Questionnaire: Development and psychometrics. Retrieved from <http://www.wilderdom.com/abstracts/NeillMarchRichards2003LEQDevelopmentPsychometrics.htm>
- Neill, J. T., Richards, G., & Badenoch, A. (1997). A research evaluation of short-term camps for youth at risk. Canberra, Australia: National Outdoor Education & Leadership Services.
- Neill, J. T., & Richards, G. E. (1998). Does outdoor education really work? A summary of recent meta-analyses. *Australian Journal of Outdoor Education*, 3(1), 2-9.
- Noble, W. S. (2009). How does multiple testing correction work? *Nature Biotechnology*, 27(12), 1135-1137. doi:10.1038/nbt1209-1135
- Norton, C. L., & Watt, T. T. (2014). Exploring the impact of a wilderness-based positive youth development program for urban youth. *Journal of Experiential Education*, 37(4), 335-350. doi:10.1177/1053825913503113
- O'Keefe, D. J. (2003). Colloquy: Should familywise alpha be adjusted? *Human Communication Research*, 29(3), 431-447. doi:10.1111/j.1468-2958.2003.tb00846.x
- O'Brien, K., & Lomas, T. (2017). Developing a Growth Mindset through outdoor personal development: Can an intervention underpinned by psychology increase the impact of an outdoor learning course for young people? *Journal of Adventure Education and Outdoor Learning*, 17(2), 133-147. doi:10.1080/14729679.2016.1232199
- O'Mara, A. J., Green, J., & Marsh, H. W. (2006). Administering self-concept interventions in schools: No training necessary? A meta-analysis. *International Education Journal*, 7(4), 524-533.

- O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41*(3), 181-206.
doi:10.1207/s15326985ep4103_4
- Oney, E., & Oksuzoglu-Guven, G. (2015). A critical review of the literature and an alternative perspective for general and specific self-confidence. *Psychological Reports, 116*(1), 149-163. doi: 10.2466/07.PRO.116k14w0
- Oort, F. (2005). Using structural equation modeling to detect response shifts and true change. *Quality of Life Research, 14*(3), 587-598. doi:10.1007/s11136-004-0830-y
- Ord, J., & Leather, M. (2011). The substance beneath the labels of experiential learning: The importance of John Dewey for outdoor educators. *Australian Journal of Outdoor Education, 15*(2), 13-23.
- Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *The Academy of Management Perspectives, 23*(1), 6-16.
doi:10.5465/AMP.2009.37007999
- Organisation for Economic Cooperation and Development. (2016). *Education at a glance 2016: OECD indicators*. Retrieved from <http://www.oecd.org/edu/education-at-a-glance-19991487.htm>
- Osborne, J. W. (2014). *Best practices in exploratory factor analysis*. (n.p.): Createspace.
- Paisley, K., Furman, N., Sibthorp, J., & Gookin, J. (2008). Student learning in outdoor education: A case study from the National Outdoor Leadership School. *Journal of Experiential Education, 30*(3), 201-222.
doi:10.1177/105382590703000302
- Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. In R. Riding & S. Rayner (Eds.), *Perception* (pp. 239-266). London: England: Ablex.
- Palmer, S., & Whybrow, A. (2008). The art of facilitation: Putting the psychology into coaching. *Psychologist, 21*(2), 136-137.

- Park, N., Peterson, C., & Ruch, W. (2009). Orientations to happiness and life satisfaction in twenty-seven nations. *Journal of Positive Psychology, 4*(4), 273-279. doi:10.1080/17439760902933690
- Parker, P. D., Jerrim, J., Schoon, I., & Marsh, H. W. (2016). A multination study of socioeconomic inequality in expectations for progression to higher education: The role of between-school tracking and ability stratification. *American Educational Research Journal, 53*(1), 6-32. doi:10.3102/0002831215621786
- Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., & Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: A multicontext study. *Developmental Psychology, 48*(6), 1629-1642. doi:10.1037/a0029167
- Passmore, J., & Brown, A. (2009). Coaching non-adult students for enhanced examination performance: A longitudinal study. *Coaching: An International Journal of Theory, Research and Practice, 2*(1), 54-64. doi:10.1080/17521880902783124
- Patton, M. Q. (2002). Two decades of developments in qualitative inquiry. *Qualitative Social Work, 1*(3), 261-283. doi:10.1177/1473325002001003636
- Patton, M. Q. (2015). *Qualitative research and evaluation methods: Integrating theory and practice* (4th ed.). Thousand Oaks, CA: SAGE.
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*(6), 784-793. doi:10.1177/0956797615571017
- Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychological Assessment, 5*(2), 164-172. doi:10.1037/1040-3590.5.2.164
- Pavot, W., & Diener, E. (2008). The Satisfaction With Life Scale and the emerging construct of life satisfaction. *Journal of Positive Psychology, 3*(2), 137-152. doi:10.1080/17439760701756946
- Pavot, W., & Diener, E. (2009). Review of the Satisfaction with Life Scale. In E. Diener (Ed.), *Social indicators research series: Vol. 39. Assessing well-being: The collected works of Ed Diener* (pp. 101-117). New York, NY: Springer.

- Paxton, T., & McAvoy, L. (2000). Social psychological benefits of a wilderness adventure program. In S. F. McCool, D. N. Cole, W. T. Borrie, & J. O'Loughlin (Eds.), *Wilderness Science in a Time of Change conference-Volume 3: Wilderness as a Place for Scientific Inquiry* (pp. 202-206). Missoula, MT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.
- Perkins-Gough, D. (2013). The significance of grit: A conversation with Angela Lee Duckworth [Interview]. *Educational Leadership*, 71(1), 14-20.
- Perneger, T. V. (1998). What's wrong with Bonferroni adjustments? *British Medical Journal*, 316(7139), 1236-1238.
- Peterson, C. (2000). The future of optimism. *American Psychologist*, 55(1), 44-55. doi:10.1037/0003-066X.55.1.44
- Peterson, C., Park, N., & Seligman, M. E. (2005). Orientations to happiness and life satisfaction: The full life versus the empty life. *Journal of Happiness Studies*, 6(1), 25-41. doi:10.1007/s10902-004-1278-z
- Piaget, J. (1976). Piaget's theory. In B. Inhelder, H. H. Chipman, & C. Zwingmann (Eds.), *Piaget and his school: A reader in developmental psychology* (pp. 11-23). Berlin, Germany: Springer.
- Piaget, J. (1983). Piaget's theory. In W. Kessen & P. H. Mussen (Eds.), *Handbook of child psychology, Vol. 1: History, theory and methods* (4th ed.). Chichester, NY: Wiley. doi:10.1002/1520-6696(198607)22:3<254::AID-JHBS2300220317>3.0.CO;2-Q
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95(4), 667-686. doi:10.1037/0022-0663.95.4.667
- Pizzolato, J. (2003). Developing self-authorship: Exploring the experiences of high-risk college students. *Journal of College Student Development*, 44(6), 797-812.
- Popp, N., & Portnow, K. (2001). Our developmental perspective on adulthood. In The Adult Development Research Group (Ed.), *Research Monograph: NCSALL Reports #19: Toward a new pluralism in ABE/ESOL classrooms:*

Teaching to multiple "cultures of mind". Cambridge, MA: National Center for the Study of Adult Learning and Literacy.

- Powers, D. L. (2004). *The effects of an outdoor education program on life effectiveness skills of participants*. (Master's thesis). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 1423059)
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31*, 437-448.
- Priest, S. (1993). A new model for risk taking. *Journal of Experiential Education, 16*(1), 50-53. doi:10.1177/105382599301600111
- Pritchard, M., & van Nieuwerburgh, C. (2016). The perceptual changes in life experience of at-risk adolescent girls following an integrated coaching and positive psychology intervention group programme: An interpretative phenomenological analysis. *International Coaching Psychology Review, 11*(1), 57-74.
- Propst, D. B., & Koesler, R. A. (1998). Bandura goes outdoors: Role of self-efficacy in the outdoor leadership development process. *Leisure Sciences, 20*(4), 319-344. doi:10.1080/01490409809512289
- Putz, R., O'Hara, K., Taggart, F., & Stewart-Brown, S. (2012). *Using WEMWBS to measure the impact of your work on mental wellbeing: A practice-based user guide*. Retrieved from http://www.corc.uk.net/media/1244/wemwbs_practitioneruserguide.pdf
- Qualtrics. (2017). Qualtrics [Computer software]. Retrieved from <http://www.qualtrics.com>
- R Core Team. (2017). R: A language and environment for statistical computing. Retrieved from <https://www.R-project.org/>
- Rathunde, K. (1996). Family context and talented adolescents' optimal experience in school-related activities. *Journal of Research on Adolescence, 6*(4), 605-628.
- Rauch, W. A., Schweizer, K., & Moosbrugger, H. (2007). Method effects due to social desirability as a parsimonious explanation of the deviation from

- unidimensionality in LOT-R scores. *Personality and Individual Differences*, 42(8), 1597-1607. doi:10.1016/j.paid.2006.10.035
- Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32(4), 329-353. doi:10.1207/s15327906mbr3204_2
- Raykov, T. (2001). Bias of coefficient alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25(1), 69-76.
- Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy*, 35(2), 299-331. doi:10.1016/S0005-7894(04)80041-8
- Raykov, T., & Shrout, P. E. (2002). Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Structural Equation Modeling*, 9(2), 195-212. doi:10.1207/S15328007SEM0902_3
- Reid, K., Flowers, P., & Larkin, M. (2005). Exploring lived experience: An introduction to interpretative phenomenological analysis. *Psychologist*, 18(1), 20-23.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667-696. doi:10.1080/00273171.2012.715555
- Reivich, K., Gillham, J. E., Chaplin, T. M., & Seligman, M. E. P. (2013). From helplessness to optimism: The role of resilience in treating and preventing depression in youth. In S. Goldstein & R. B. Brooks (Eds.), *Handbook of resilience in children* (pp. 201-214). Boston, MA: Springer.
- Revelle, W. (2018). psych: Procedures for Personality and Psychological Research [Computer software]. Evanston, IL: Northwestern University. Retrieved from <https://CRAN.R-project.org/package=psych> Version = 1.8.12.
- Richards, G. E., Ellis, L. A., & Neill, J. T. (2002). The ROPELOC: Review of Personal Effectiveness and Locus of Control: A comprehensive instrument for reviewing life effectiveness. In R. G. Craven, H. W. Marsh, & K. B. Simpson (Eds.), *Self-Concept research: Driving international research agendas: Proceedings of the 2nd Biennial SELF Research Centre International*

Conference. Sydney, Australia: SELF Research Centre, University of Western Sydney.

Richmond, D., Sibthorp, J., Gookin, J., Annarella, S., & Ferri, S. (2018).

Complementing classroom learning through outdoor adventure education: Out-of-school-time experiences that make a difference. *Journal of Adventure Education and Outdoor Learning*, 18(1), 36-52.

doi:10.1080/14729679.2017.1324313

Richter, L. M. (2006). Studying adolescence. *Science*, 312(5782), 1902-1905.

doi:10.1126/science.1127489

Robson-Kelly, L., & van Nieuwerburgh, C. (2016). What does coaching have to offer to young people at risk of developing mental health problems? A grounded theory study. *International Coaching Psychology Review*, 11(1), 75-92.

Rose, L., Williams, I. R., Olsson, C. A., & Allen, N. B. (2018). Promoting adolescent health and wellbeing through outdoor youth programs: Results from a multisite Australian study. *Journal of Outdoor Recreation, Education, and Leadership*, 10(1), 33-51. doi:10.18666/JOREL-2018-V10-I1-808

Rose, S., & Sieben, N. (2018). Hope measurement. In M. W. Gallagher & S. J. Lopez (Eds.), *The Oxford handbook of hope* (1st ed., pp. 1-23). New York, NY: Oxford University Press.

Rothman, J. K. (1990). No adjustments are needed for multiple comparisons.

Epidemiology, 1(1), 43-46. doi:10.1097/00001648-199001000-00010

Rotter, J. B. (1954). *Social learning and clinical psychology*. New York, NY: Prentice-Hall.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs*, 80(1), 1. doi:10.1037/h0092976

Rowe, K. J. E., & Rowe, K. S. E. (1999). Investigating the relationship between students' attentive-inattentive behaviors in the classroom and their literacy progress. *International Journal of Educational Research*, 31, 1-16.

Rowley, J. (1987). Adventure education and qualitative research. *Journal of Experiential Education*, 10(2), 8-12. doi:10.1177/105382598701000202

RStudio Team. (2018). RStudio: Integrated Development for R [Computer software]. Boston, MA: RStudio, Inc. Retrieved from <http://www.rstudio.com/>

- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581-592.
doi:10.2307/2335739
- Rutter, M. (2006). Implications of resilience concepts for scientific understanding. *Annals of the New York Academy of Sciences*, 1094, 1-12.
doi:10.1196/annals.1376.001
- Ryan, T. A. (1970). *Intentional behavior: An approach to human motivation*. New York, NY: Ronald.
- Sackett, P. R. (2012). Faking in personality assessment: Where do we stand? In M. Ziegler, C. MacCann, & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 330-344). New York, NY: Oxford University Press.
- Sanford, N. (1962). *The American college: A psychological and social interpretation of the higher learning*. New York, NY: Wiley.
- Sanford, N. (1966). *Self and society: Social change and individual development*. New York, NY: Atherton.
- Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(2), 167-180. doi:10.1080/10705511.2014.882658
- Satorra, A., & Bentler, P. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507-514.
doi:10.1007/BF02296192
- Satorra, A., & Bentler, P. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75(2), 243-248. doi:10.1007/s11336-009-9135-y
- Scarf, D., Kafka, S., Hayhurst, J., Jang, K., Boyes, M., Thomson, R., & Hunter, J. A. (2018). Satisfying psychological needs on the high seas: Explaining increases in self-esteem following an adventure education programme. *Journal of Adventure Education and Outdoor Learning*, 18(2), 165-175.
doi:10.1080/14729679.2017.1385496
- Schafermeyer, H. (1978). Adventure programming: Wilderness and urban. *Journal of Physical Education and Recreation*, 49(1), 30-32.

- Schary, D. P., Lewis, A. B., & Cardinal, B. J. (2015). Learning goals and the challenge course experience: An exploratory study. *Recreational Sports Journal, 39*, 59-68. doi:10.1123/rsj.2014-0036
- Schary, D. P., Wozniak, T., Jenny, S. E., & Morrow, G. S. (2016). Short- and long-term retention of challenge course outcomes: A classroom-based longitudinal study. *Recreational Sports Journal, 40*(2), 152-164. doi:10.1123/rsj.2016-0013
- Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health psychology, 4*(3), 219-247. doi:10.1037/0278-6133.4.3.219
- Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing Optimism From Neuroticism (and Trait Anxiety, Self-Mastery, and Self-Esteem): A Reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67*(6), 1063-1078. doi:10.1037/0022-3514.67.6.1063
- Schriesheim, C. A., & Hill, K. D. (1981). Controlling acquiescence response bias by item reversals: The effect on questionnaire validity. *Educational and Psychological Measurement, 41*(4), 1101-1114. doi:10.1177/001316448104100420
- Schueller, S. M., & Seligman, M. E. (2010). Pursuit of pleasure, engagement, and meaning: Relationships to subjective and objective measures of well-being. *Journal of Positive Psychology, 5*(4), 253-263. doi:10.1080/17439761003794130
- Schulz, K., & Grimes, D. (2005). Epidemiology 4: Multiplicity in randomised trials I: Endpoints and treatments. *The Lancet, 365*(9470), 1591-1595.
- Schunk, D. H. (2014). *Motivation in education: Theory, research and applications* (4th ed.). Harlow, England: Pearson Education.
- Scrutton, R., & Beames, S. (2015). Measuring the unmeasurable: Upholding rigor in quantitative studies of personal and social development in outdoor adventure education. *Journal of Experiential Education, 38*(1), 8-25. doi:10.1177/1053825913514730
- Seaman, J. (2008). Experience, reflect, critique: The end of the "learning cycles" era. *Journal of Experiential Education, 31*(1), 3-18. doi:10.1177/105382590803100103
- Seligman, M. E. (2002). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment*. New York, NY: Free Press.

- Seligman, M. E. (2011). *Flourish: A visionary new understanding of happiness and well-being*. New York, NY: Free Press.
- Seligman, M. E., & Csikszentmihalyi, M. (2000). Positive psychology. An introduction. *Am Psychol*, *55*(1), 5-14.
- semTools Contributors. (2016). semTools: Useful tools for structural equation modeling [Computer software]. Retrieved from <https://CRAN.R-project.org/package=semTools>
- Senn, S. (2013). Seven myths of randomisation in clinical trials. *Statistics in Medicine*, *32*(9), 1439-1450. doi:10.1002/sim.5713
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*(3), 407-441. doi:10.3102/00346543046003407
- Sheard, M., & Golby, J. (2006). The efficacy of an outdoor adventure education curriculum on selected aspects of positive psychological development. *Journal of Experiential Education*, *29*(2), 187-209. doi:10.1177/105382590602900208
- Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, *76*(3), 482-497. doi:10.1037/0022-3514.76.3.482
- Sheldon, K. M., Elliot, A. J., Ryan, R. M., Chirkov, V., Kim, Y., Wu, C., . . . Sun, Z. (2004). Self-concordance and subjective well-being in four cultures. *Journal of cross-cultural psychology*, *35*(2), 209-223. doi:<https://doi.org/10.1177/0022022103262245>
- Sheldon, K. M., & Houser-Marko, L. (2001). Self-concordance, goal attainment, and the pursuit of happiness: Can there be an upward spiral? *Journal of Personality and Social Psychology*, *80*(1), 152. doi:10.1037//0022-3514.80.1.152
- Sheldon, K. M., & Kasser, T. (1998). Pursuing personal goals: Skills enable progress, but not all progress is beneficial. *Personality and Social Psychology Bulletin*, *24*(12), 1319-1331. doi:10.1177/01461672982412006
- Shvarts, A., & Bakker, A. (2019). The early history of the scaffolding metaphor: Bernstein, Luria, Vygotsky, and before. *Mind, Culture, and Activity*, *26*(1), 4-23. doi:10.1080/10749039.2019.1574306

- Sibthorp, J. (2000). Measuring weather... and adventure education: Exploring the instruments of adventure education research. *Journal of Experiential Education*, 23(2), 99-107. doi:10.1177/105382590002300208
- Sibthorp, J. (2003). An empirical look at Walsh and Golins' adventure education process model: Relationships between antecedent factors, perceptions of characteristics of an adventure education experience, and changes in self-efficacy. *Journal of Leisure Research*, 35(1), 80-106. doi:10.18666/JLR-2003-V35-11-611
- Sibthorp, J. (2010). A letter from the editor: Positioning outdoor and adventure programs within positive youth development. *Journal of Experiential Education*, 33(2), vi-ix.
- Sibthorp, J., & Arthur-Banning, S. (2004). Developing life effectiveness through adventure education: The roles of participant expectations, perceptions of empowerment, and learning relevance. *Journal of Experiential Education*, 27(1), 32-50. doi:10.1177/105382590402700104
- Sibthorp, J., Collins, R., Rathunde, K., Paisley, K., Schumann, S., Pohja, M., . . . Baynes, S. (2015). Fostering experiential self-regulation through outdoor adventure education. *Journal of Experiential Education*, 38(1), 26-40. doi:10.1177/1053825913516735
- Sibthorp, J., Furman, N., Paisley, K., Gookin, J., & Schumann, S. (2011). Mechanisms of learning transfer in adventure education: Qualitative results from the NOLS transfer survey. *Journal of Experiential Education*, 34(2), 109-126. doi:10.5193/JEE34.2.109
- Sibthorp, J., & Jostad, J. (2014). The social system in outdoor adventure education programs. *Journal of Experiential Education*, 37(1), 60-74. doi:10.1177/1053825913518897
- Sibthorp, J., Paisley, K., Furman, N., & Gookin, J. (2008). Long-term impacts attributed to participation in wilderness education: Preliminary findings from NOLS (abstract). In K. Liddicoat, S. Todd, & A. B. Young (Eds.), *Abstracts from the Coalition for Education in the Outdoors Ninth Biennial Research Symposium*. Bradford Woods, IN: Coalition for Education in the Outdoors.

- Sibthorp, J., Paisley, K., & Gookin, J. (2007). Exploring participant development through adventure-based programming: A model from the National Outdoor Leadership School. *Leisure Sciences, 29*(1), 1-18.
doi:10.1080/01490400600851346
- Sibthorp, J., Witter, E., Wells, M., Ellis, G., & Voelkl, J. (2004). Hierarchical linear modeling in park, recreation, and tourism research. *Journal of Leisure Research, 36*(1), 89-100. doi:10.1080/00222216.2004.11950012
- Smith, J. A. (1996). Beyond the divide between cognition and discourse: Using interpretative phenomenological analysis in health psychology. *Psychology & Health, 11*(2), 261-271. doi:10.1080/08870449608400256
- Smith, J. A. (2004). Reflecting on the development of interpretative phenomenological analysis and its contribution to qualitative research in psychology. *Qualitative Research in Psychology, 1*(1), 39-54.
doi:10.1191/1478088704qp004oa
- Snyder, C. R. (1995). Conceptualizing, measuring, and nurturing hope. *Journal of Counseling & Development, 73*(3), 355-360. doi:10.1002/j.1556-6676.1995.tb01764.x
- Snyder, C. R. (2002). Hope theory: Rainbows in the mind. *Psychological Inquiry, 13*(4), 249-275.
- Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., . . . Harney, P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of Personality and Social Psychology, 60*(4), 570-585. doi:10.1037/0022-3514.60.4.570
- Snyder, C. R., Hoza, B., Pelham, W. E., Rapoff, M., Ware, L., Danovsky, M., . . . Stahl, K. J. (1997). The development and validation of the Children's Hope Scale. *Journal of pediatric psychology, 22*(3), 399-421. doi:10.1093/jpepsy/22.3.399
- Snyder, C. R., Rand, K. L., & Sigmon, D. R. (2002). Hope theory: A member of the positive psychology family. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 257-276). New York, NY: Oxford University Press.
- Snyder, C. R., & Shorey, H. (2002). Hope in the classroom: The role of positive psychology in academic achievement and psychology curriculum. *Psychology Teacher Network, 12*(1), 1-4.

- Snyder, C. R., Shorey, H., & Rand, K. L. (2006). Using hope theory to teach and mentor academically at-risk students. In W. Buskist & S. F. Davis (Eds.), *Handbook of the teaching of psychology* (pp. 170-174). Hoboken, NJ: Blackwell Publishing.
- Spence, G. B., Stout-Rostron, S., Van Reenen, M., & Glashoff, B. (2019). Exploring the delayed effects of leadership coaching: A pilot study. *Coaching: An International Journal of Theory, Research and Practice*, *12*(2), 125-146. Retrieved from <https://www.tandfonline.com/doi/abs/10.1080/17521882.2019.1574308>
doi:10.1080/17521882.2019.1574308
- Spinelli, E. (2005). *The interpreted world: An introduction to phenomenological psychology* (2nd ed.). London, England: SAGE.
- Sproule, J., Martindale, R., Wang, J., Allison, P., Nash, C., & Gray, S. (2013). Investigating the experience of outdoor and adventurous project work in an educational setting using a self-determination framework. *European Physical Education Review*, *19*(3), 315-328. doi:10.1177/1356336X13495629
- Standards Australia. (2011). *Coaching in organizations* (HB 332-2011). Sydney, Australia: SAI Global.
- Stankov, L., Morony, S., & Lee, Y. P. (2013). Confidence: The best non-cognitive predictor of academic achievement? *Educational Psychology*, *34*(1), 1-20. doi:10.1080/01443410.2013.814194
- Steenkamp, J. B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*(1), 78-107. doi:10.1086/209528
- Steger, M. F., Kashdan, T. B., & Oishi, S. (2008). Being good by doing good: Daily eudaimonic activity and well-being. *Journal of Research in Personality*, *42*(1), 22-42. doi:10.1016/j.jrp.2007.03.004
- Steiger, J. H., & Lind, J. C. (1980). *Statistically-based tests for the number of common factors*. Paper presented at the Annual Meeting of the Psychometric Society, Iowa City, IA.
- Steinmetz, H. (2013). Analyzing observed composite differences across multiple groups: Is partial measurement invariance enough? *Methodology: European*

Journal of Research Methods for the Behavioral and Social Sciences, 9, 1-12.

doi:10.1027/1614-2241/a000049

- Stenger, T. (2001). *Sequence of adventure-based resident outdoor education programs and middle school students' perceptions of life effectiveness*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3023956)
- Stewart-Brown, S., Tennant, R., Tennant, A., Platt, S., Parkinson, J., & Weich, S. (2009). Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): A Rasch analysis using data from the Scottish Health Education Population Survey. *Health and Quality of Life Outcomes*, 7(1), 15. doi:10.1186/1477-7525-7-15
- Stober, D. R., & Grant, A. M. (Eds.). (2006). *Evidence based coaching handbook*. Hoboken, NJ: Wiley.
- Stone, C. A. (1998). The metaphor of scaffolding: Its utility for the field of learning disabilities. *Journal of learning disabilities*, 31(4), 344-364. doi:10.1177/002221949803100404
- Stott, T., Allison, P., Felter, J., & Beames, S. (2015). Personal development on youth expeditions: A literature review and thematic analysis. *Leisure Studies*, 34(2), 197-229. doi:10.1080/02614367.2013.841744
- Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80(1), 99-103. doi:10.1207/S15327752JPA8001_18
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of personality*, 72(2), 271. doi:10.1111/j.0022-3506.2004.00263.x
- Taniguchi, S., & Freeman, P. A. (2004). Outdoor education and meaningful learning: Finding the attributes to meaningful learning experiences in an outdoor education program. *Journal of Experiential Education*, 26(3), 210-211. doi:10.1177/105382590402600318
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55.

- Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., . . . Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes*, 5(1), 63. doi:10.1186/1477-7525-5-63
- Tesiny, E. P., Lefkowitz, M. M., & Gordon, N. H. (1980). Childhood depression, locus of control, and school achievement. *Journal of Educational Psychology*, 72(4), 506-510.
- The Adult Development Research Group. (2001). *Research Monograph NCSALL Reports #19: Toward a New Pluralism in ABE/ESOL Classrooms: Teaching to Multiple "Cultures of Mind"*. Cambridge, MA: National Center for the Study of Adult Learning and Literacy.
- Theeboom, T., Beersma, B., & van Vianen, A. E. (2014). Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *Journal of Positive Psychology*, 9(1), 1-18. doi:10.1080/17439760.2013.837499
- Thompson, B. (1994). Guidelines for authors. *Educational and Psychological Measurement*, 54, 837-847.
- Thompson, B. (2003a). Guidelines for authors reporting score reliability estimates. In B. Thompson (Ed.), *Score reliability: Contemporary thinking on reliability issues* (pp. 91-101). Thousand Oaks, CA: SAGE. Retrieved from <http://methods.sagepub.com/book/score-reliability>. doi:10.4135/9781412985789
- Thompson, B. (2003b). Understanding reliability and coefficient alphas, really. In B. Thompson (Ed.), *Score reliability: Contemporary thinking on reliability issues*. Thousand Oaks, CA: SAGE. Retrieved from <http://methods.sagepub.com/book/score-reliability>. doi:10.4135/9781412985789
- Tucker, L., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1-10. doi:10.1007/BF02291170
- Twenge, J. M., Zhang, L., & Im, C. (2004). It's beyond my control: A cross-temporal meta-analysis of increasing externality in locus of control, 1960-2002.

- Personality and Social Psychology Review*, 8(3), 308-319.
doi:10.1207/s15327957pspr0803_5
- Ungar, M., Dumond, C., & McDonald, W. (2005). Risk, resilience and outdoor programmes for at-risk children. *Journal of Social Work*, 5(3), 319-338.
doi:10.1177/1468017305058938
- Uziel, L. (2010). Rethinking social desirability scales: From impression management to interpersonally oriented self-control. *Perspectives on Psychological Science*, 5(3), 243-262. doi:10.1177/1745691610369465
- Vacek, K. R., Coyle, L. D., & Vera, E. M. (2010). Stress, self-esteem, hope, optimism, and well-being in urban, ethnic minority adolescents. *Journal of Multicultural Counseling and Development*, 38(2), 99-111.
- Valentine, J. C., Cooper, H., Bettencourt, B. A., & Dubois, D. L. (2002). Out-of-school activities and academic achievement: The mediating role of self-beliefs. *Educational Psychologist*, 37(4), 245-256.
doi:10.1207/S15326985EP3704_4
- Valentine, J. C., Dubois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39(2), 111-133. doi:10.1207/s15326985ep3902_3
- Valle, M. F., Huebner, E. S., & Suldo, S. M. (2004). Further evaluation of the Children's Hope Scale. *Journal of Psychoeducational Assessment*, 22(4), 320-337. doi:10.1177/073428290402200403
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R [Computer software]. *Journal of Statistical Software*, 45(3), 1-67.
- Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271-296. doi:10.1007/s10648-010-9127-6
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4-70. doi:10.1177/109442810031002

- Vassar, M. (2008). A note on the score reliability for the Satisfaction with Life Scale: An RG study. *Social Indicators Research*, 86(1), 47-57. doi:10.1007/s11205-007-9113-7
- Vassar, M., & Bradley, G. (2010). A reliability generalization study of coefficient alpha for the Life Orientation Test. *Journal of Personality Assessment*, 92(4), 362-370. doi:10.1080/00223891.2010.482016
- Vautier, S., Raufaste, E., & Cariou, M. (2003). Dimensionality of the Revised Life Orientation Test and the status of filler items. *International Journal of Psychology*, 38(6), 390-400. doi:10.1080/00207590344000222
- Veazie, P. J. (2006). When to combine hypotheses and adjust for multiple tests. *Health Services Research*, 41(3p1), 804-818. doi:10.1111/j.1475-6773.2006.00512.x
- Vella-Brodrick, D. A., Park, N., & Peterson, C. (2009). Three ways to be happy: Pleasure, engagement, and meaning—Findings from Australian and US samples. *Social Indicators Research*, 90(2), 165-179. doi:10.1007/s11205-008-9251-6
- Villegas-Reimers, E. (1996). Self development of Venezuelan adolescents: A test of Kegan's theory and subject-object interview in another culture. *Journal of cross-cultural psychology*, 24(1), 25.
- Vince, R. (1998). Behind and beyond Kolb's learning cycle. *Journal of Management Education*, 22(3), 304-319. doi:10.1177/105256299802200304
- Vincent, S. M. (1995). Emotional safety in adventure therapy programs: Can it be defined? *Journal of Experiential Education*, 18(2), 76-81. doi:10.1177/105382599501800204
- Vitterso, J., & Soholt, Y. (2011). Life satisfaction goes with pleasure and personal growth goes with interest: Further arguments for separating hedonic and eudaimonic well-being. *Journal of Positive Psychology*, 6(4), 326-335. doi:10.1080/17439760.2011.584548
- Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
- Walsh, V., & Golins, G. (1976). The exploration of the Outward Bound Process. Retrieved from <https://wilderdom.com>

- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2), 129-133. doi:10.1080/00031305.2016.1154108
- Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Statistical inference in the 21st Century: A world beyond $p < 0.05$ [Special issue]. *The American Statistician*, 73(S1).
- Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, 64(4), 678-691. doi:10.1037/0022-3514.64.4.678
- Waterman, A. S., Schwartz, S. J., & Conti, R. (2008). The implications of two conceptions of happiness (hedonic enjoyment and eudaimonia) for the understanding of intrinsic motivation. *Journal of Happiness Studies*, 9(1), 41-79. doi:10.1007/s10902-006-9020-7
- Weems, G. H., Onwuegbuzie, A. J., & Lustig, D. (2003). Profiles of respondents who respond inconsistently to positively- and negatively-worded items on rating scales. *Evaluation & Research in Education*, 17(1), 45-60. doi:10.1080/14664200308668290
- Weis, R., & Speridakos, E. (2011). A meta-analysis of hope enhancement strategies in clinical and community settings. *Theory, Research and Practice*, 1(1), 1-16. doi:10.1186/2211-1522-1-5
- Whittington, A. (2011). Life after the river: Long-term impacts of girls' participation in an adventure program. *Journal of Outdoor Recreation, Education, and Leadership*, 3(1), 40-52.
- Whittington, A., & Aspelmeier, J. E. (2018). Resilience, peer relationships, and confidence: Do girls' programs promote positive change? *Journal of Outdoor Recreation, Education, and Leadership*, 10(2), 124-138. doi:10.18666/JOREL-2018-V10-I2-7876
- Whybrow, A. (2008). Coaching psychology: Coming of age? *International Coaching Psychology Review*, 3(3), 14.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag. Retrieved from <https://ggplot2.tidyverse.org>

- Wickham, H. (2017). stringr: Simple, Consistent Wrappers for Common String Operations [Computer software]. Retrieved from <https://CRAN.R-project.org/package=stringr>
- Wickham, H., Francois, R., Henry, L., & Müller, K. (2017). dplyr: A Grammar of Data Manipulation [Computer software]. Retrieved from <https://CRAN.R-project.org/package=dplyr>
- Wickham, H., & Henry, L. (2017). tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions [Computer software]. Retrieved from <https://CRAN.R-project.org/package=tidyr>
- Winters, D., & Latham, G. P. (1996). The effect of learning versus outcome goals on a simple versus a complex task. *Group & Organization Management, 21*(2), 236-250. doi:10.1177/1059601196212007
- Worrell, F. C., & Hale, R. L. (2001). The relationship of hope in the future and perceived school climate to school completion. *School Psychology Quarterly, 16*(4), 370-388. doi:10.1521/scpq.16.4.370.19896
- Wrosch, C., Scheier, M. F., Carver, C. S., & Schulz, R. (2003). The Importance of Goal Disengagement in Adaptive Self-Regulation: When Giving Up is Beneficial. *Self and Identity, 2*(1), 1-20. doi:10.1080/15298860309021
- Wyszyńska, P., Ponikiewska, K., Karaś, D., Najderska, M., & Rogoza, R. (2017). Psychometric properties of the Polish version of the Short Grit Scale. *Polish Psychological Bulletin, 48*(2), 8. doi:10.1515/ppb-2017-0026
- Yoon, M., & Kim, E. S. (2014). A comparison of multiple-indicators, multiple-causes- and item response theory-based analyses of subgroup differences. *Behavior Research Methods, 46*, 1199-1206. doi:10.3758/s13428-013-0430-2
- Zaff, J. F., Moore, K. A., Papillo, A. R., & Williams, S. (2003). Implications of extracurricular activity participation during adolescence on positive outcomes. *Journal of Adolescent Research, 18*(6), 599-630. doi:10.1177/0743558403254779
- Zimmerman, D. W. (1972). Test reliability and the Kuder-Richardson formulas: Derivation from probability theory. *Educational and Psychological Measurement, 32*(4), 939-954. doi:10.1177/001316447203200408

Zinbarg, R., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α , Revelle's β , and McDonald's ω_h : Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123-133.

doi:10.1007/s11336-003-0974-7

Zinbarg, R. E., Revelle, W., & Yovel, I. (2007). Estimating ω_h for structures containing two group factors: Perils and prospects. *Applied Psychological Measurement*, *31*(2), 135-157. doi:10.1177/0146621606291558

APPENDICES

APPENDIX A

SURVEY ITEMS: CROSS-REFERENCES AND WORDING

Table A.1

Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
Children's Hope Scale			
Agency			
Ag ₁	1	1	I think I am doing pretty well
Ag ₂	3	27	I am doing just as well as other kids my age
Ag ₃	5	46	I think the things I have done in the past will help me in the future
Pathways Thinking			
Pth ₁	2	15	I can think of many ways to get the things in life that are most important to me
Pth ₂	4	38	When I have a problem, I can come up with lots of ways to solve it
Pth ₃	6	60	Even when others want to quit, I know that I can find ways to solve the problem
Life Orientation Test, Revised			
Optimism			
Opt ₁	1	200	In uncertain times, I usually expect the best
Opt ₂	4	147	I'm always optimistic about my future
Opt ₃	10	52	Overall I expect more good things [than bad]
Pessimism			
Ps ₁ ^a	3	163	If something can go wrong for me, it will
Ps ₂ ^a	7	95	I hardly ever expect things to go my way
Ps ₃ ^a	9	67	I rarely count on good things happening to me
Filler items			
F ₁	2	178	It's easy for me to relax
F ₂	5	126	I enjoy my friends a lot
F ₃	6	109	It's important for me to keep busy
F ₄	8	84	I don't get upset too easily

^a Item is negatively worded and was reverse-scored prior to analysis.

(continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
Adolescent Self-Regulatory Inventory			
<i>Long-Term Self-Regulation</i>			
LT1	3	29	If something isn't going according to my plans, I change my actions to try and reach my goal
LT2	4	41	I can find ways to make myself study, even when my friends want to go out
LT3 ^{ab}	7	64	It's hard for me to get started on big projects that require planning in advance
LT4 ^a	12	33/76 (item repeated)	I lose control whenever I don't get my way
LT5 ^a	15	24	If I really want something, I have to have it right away
LT6	20	89	When I have a serious disagreement with someone, I can talk calmly about it without losing control
LT7	23	18	I can stay focused on my work even when it's dull
LT8	25	120	I can stop myself from doing things like throwing objects when I'm mad
LT9	26	133	I work carefully when I know something will be tricky
LT10	27/28*	143	I'm usually aware of my feelings before I let them out
LT11	28/29*	154	In class, I can concentrate on my work even if my friends are talking
LT12	29/20*	116	When I'm excited about reaching a goal (e.g., getting my driver's license, going to college), it's easy to start working on it
LT13	30/31*	129	I can find a way to stick with my plans and goals, even when it's tough
LT14	31/32*	137	When I have a big project, I can keep working on it
LT15 ^{ab}	34/35*	62	I have trouble getting excited about something that's really special when I am tired
LT16	36/37*	13	I can resist doing something when I know I shouldn't do it

Note. LT = Long-Term Self-Regulation.

*Inconsistency in the numbering of the original measurement instrument (see Moilanen, 2007).

^a Item is negatively worded and was reverse-scored prior to analysis.

^b Item was not scored in original measurement instrument (Moilanen, 2011).

(continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
<i>Short-Term Self-Regulation</i>			
ST1 ^{a b}	1	21	It's hard for me to notice when I have had enough (sweets, food, etc.)
ST2	2	7	When I'm sad, I can usually start doing something that will make me feel better
ST3 ^{a b}	5	56	I lose track of the time when I'm doing something fun
ST4 ^a	6	48	When I'm bored, I fidget or can't sit still
ST5	8	71	I can usually act normal around everybody if I am upset with someone
ST6	9	87	I am good at keeping track of lots of things going on around me, even when I'm feeling stressed
ST7 ^b	10	103	When I'm having a tough day, I stop myself from whining about it to my family and friends
ST8	11	92	I can start a new task, even if I'm already tired
ST9 ^a	13	11	Little problems detract me from my long-term plans
ST10 ^a	14	4	I forget about whatever else I need to do when I'm doing something really fun
ST11 ^a	16	35	During a dull task, I have trouble forcing myself to start paying attention
ST12	17	43	After I'm interrupted or distracted, I can easily continue working where I left off
ST13 ^a	18	59	If there are other things going on around me, I find it hard to keep my attention focused on whatever I'm doing
ST14 ^a	19	80	I never know how much more work I have to do
ST15 ^a	21	99	It's hard to start making plans to deal with a big project or problem, especially when I'm feeling stressed
ST16	22	105	I can calm myself down when I'm excited or all wound up

Note. ST = Short-Term Self-Regulation.

^{*}Inconsistency in the numbering of the original measurement instrument (see Moilanen, 2007).

^a Item is negatively worded and was reverse-scored prior to analysis.

^b Item was not scored in original measurement instrument (Moilanen, 2011).

(continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
ST17 ^{a b}	24	113	I usually know when I'm going to start crying
ST18 ^{a b}	32/33*	151	I can usually tell when I'm getting tired or frustrated.
ST19 ^a	33/34*	82	I get carried away emotionally when I get excited about something
ST20 ^{a b}	35/36*	69	It's hard for me to keep focused on something I find unpleasant or upsetting
Adolescent Self-Regulatory Inventory, Revised (with original ASRI item reference indicated in parentheses)			
Focus			
Foc1 (LT11)	28/29*	154	In class, I can concentrate on my work even if my friends are talking
Foc2 (ST12)	17	43	After I'm interrupted or distracted, I can easily continue working where I left off
Foc3 (LT7)	23	18	I can stay focused on my work even when it's dull
Foc4 (LT2)	4	41	I can find ways to make myself study, even when my friends want to go out
Goal Self-Regulation			
GSR1 (ST8)	11	92	I can start a new task, even if I'm already tired
GSR2 (LT13)	30/31*	129	I can find a way to stick with my plans and goals, even when it's tough
GSR3 (LT14)	31/32*	137	When I have a big project, I can keep working on it
GSR4 (LT1)	3	29	If something isn't going according to my plans, I change my actions to try and reach my goal
Emotional Self-Regulation			
ESR1 (ST2)	2	7	When I'm sad, I can usually start doing something that will make me feel better
ESR2 (ST5)	8	71	I can usually act normal around everybody if I am upset with someone
ESR3 (LT6)	20	89	When I have a serious disagreement with someone, I can talk calmly about it without losing control
ESR4 (LT8)	25	120	I can stop myself from doing things like throwing objects when I'm mad

Note. ST = Short-Term Self-Regulation; Foc = Focus; LT = Long-Term; GSR = Goal Self-Regulation; ESR = Emotion Self-Regulation.

*Inconsistency in the numbering of the original measurement instrument (see Moilanen, 2007).

^a Item is negatively worded and was reverse-scored prior to analysis.

^b Item was not scored in original measurement instrument (Moilanen, 2011).

(continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
Academic Resilience Scale			
AR1	1	9	I believe I'm mentally tough when it comes to exams
AR2	2	17	I don't let study stress get on top of me
AR3	3	25	I'm good at bouncing back from a poor mark in my schoolwork
AR4	4	36	I think I'm good at dealing with schoolwork pressures
AR5	5	107	I don't let a bad mark affect my confidence
AR6	6	118	I'm good at dealing with setbacks at school (e.g. bad mark, negative feedback on my work)
Life Resilience Scale			
LR1	1	5	I believe I am mentally tough when it comes to overcoming life challenges
LR2	2	31	I don't usually let life stresses get on top of me
LR3	3	40	I'm good at bouncing back from disappointments in my life
LR4	4	50	I think I'm good at dealing with sources of pressure in my life
LR5	5	65	I don't let difficulties and disappointments in life affect my confidence
LR6	6	73	I'm good at dealing with setbacks (e.g. negative feedback on what I do, disappointing outcomes)
Short Grit Scale			
Consistency of Interest			
COI1 ^a	1	77	I often set a goal but later chose to pursue a different one
COI2 ^a	2	85	New ideas and projects sometimes distract me from previous ones
COI3 ^a	3	93	I have been obsessed with a certain idea or project for a short time but later lost interest
COI4 ^a	4	123	I have difficulty maintaining my focus on projects that take more than a few months to complete
Perseverance of Effort			
POE1	5	131	I finish whatever I begin
POE2	6	90	Setbacks don't discourage me
POE3	7	97	I am diligent
POE4	8	54	I am a hard worker

Note. AR = Academic Resilience; LR = Life Resilience; COI = Consistency of Interest; POE = Perseverance of Effort.

^aItem is negatively worded and was reverse-scored prior to analysis.

(continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
Motivation and Engagement Scale - Short (no access to original instrument)			
Booster Thoughts			
BT ₁	1	190	I believe I can do a good job in my schoolwork (Self-belief)
BT ₂	2	174	What I learn in my schoolwork is important and useful (Valuing)
BT ₃	3	158	In my schoolwork, I am focused on learning and improving more than competing and being the best (Learning focus)
Booster Behaviours			
BB ₁	4	140	I plan out how I will do my schoolwork and study (Planning)
BB ₂	5	121	I use my study/homework time well and try to study and do homework under conditions that bring out my best (Task management)
BB ₃	6	110	I persist at schoolwork even when it is challenging or difficult (Persistence)
Mufflers			
Mf ₁ ^a	7	101	I get quite anxious about schoolwork and tests (Anxiety)
Mf ₂ ^a	8	94	I mainly do my schoolwork to avoid failing or disapproval from parents or the teacher/s (Failure avoidance)
Mf ₃ ^a	9	78	I don't think I have much control over how well I do in my schoolwork (Uncertain control)
Guzzlers			
Gz ₁ ^a	10	57	In my schoolwork I sometimes reduce my chances of doing well (e.g. waste time, not study, disrupt others, procrastinate) (Self-sabotage)
Gz ₂ ^a	11	44	I often feel like giving up in my schoolwork (Disengagement)
Satisfaction with Life Scale			
SL ₁	1	104	In most ways my life is close to my ideal
SL ₂	2	145	The conditions of my life are excellent
SL ₃	3	149	I am satisfied with life
SL ₄	4	167	So far I have gotten the important things I want in life
SL ₅	5	184	If I could lead my life over, I would change almost nothing

Note. BT = Booster Thoughts; BB = Booster Behaviours; Mf = Mufflers; Gz = Guzzlers; SL = Satisfaction with Life.

^a Item is negatively worded and was reverse-scored prior to analysis.

(continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
Warwick-Edinburgh Mental Well-Being Scale			
WB1	1	108	I've been feeling optimistic about the future
WB2	2	68	I've been feeling useful
WB3	3	22	I've been feeling relaxed
WB4	4	170	I've been feeling interested in other people
WB5	5	176	I've had energy to spare
WB6	6	187	I've been dealing with problems well
WB7	7	195	I've been thinking clearly
WB8	8	124	I've been feeling good about myself
WB9	9	181	I've been feeling close to other people
WB10	10	198	I've been feeling confident
WB11	11	161	I've been able to make up my own mind on things
WB12	12	165	I've been feeling loved
WB13	13	156	I've been interested in new things
WB14	14	141	I've been feeling cheerful
Self-Description Questionnaire II-S			
<i>Non-Academic Self-Concept</i>			
<i>Physical Abilities SC</i>			
Pab1	5	135	I enjoy things like sports, gym and dance
Pab2	16	193	I am good at things like sports, gym and dance
Pab3 ^a	27	53	I am awkward at things like sports, gym and dance
Pab4	38	148	I am better than most of my friends at sports, gym and dance
<i>Physical Appearance SC</i>			
Pap1	2	114	I have a nice looking face
Pap2	13	172	I am good looking
Pap3	24	88	Other people think I am good looking
Pap4	35	210	I have a good looking body
<i>Opposite-Sex Relationships SC</i>			
OS1 ^a	10	164	I am not very popular with members of the opposite sex
OS2 (females only)	21	182	I make friends easily with boys
OS2 (males only)	22	155	I make friends easily with girls
OS3 (females only) ^a	43	72	I do not get along very well with boys
OS3 (males only) ^a	44	136	I do not get along very well with girls
OS4	32	168	I have lots of friends of the opposite sex

Note. SC = Self-Concept; Pab = Physical Abilities Self-Concept; Pap = Physical Appearance Self-Concept; OS = Opposite-Sex Relationships Self-Concept.

^aItem is negatively worded and was reverse-scored prior to analysis.

(continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
Same-Sex Relationships SC			
SS1 ^a	11	122	It is difficult to make friends with members of my own sex
SS2 (males only)	21	182	I make friends easily with boys
SS2 (females only)	22	155	I make friends easily with girls
SS3 (males only) ^a	43	72	I do not get along very well with boys
SS3 (females only) ^a	44	136	I do not get along very well with girls
SS4	49	201	I make friends easily with members of my own sex
SS5	33	212	Not many people of my own sex like me
Parent Relationships SC			
PR1	8	153	I get along well with my parents
PR2	19	185	My parents treat me fairly
PR3	30	188	My parents understand me
PR4 ^a	41	100	I do not like my parents very much
Honesty/Trustworthiness SC			
Ho1	4	127	I am honest
Ho2 ^a	15	186	I often tell lies
Ho3 ^a	26	61	I sometimes cheat
Ho4 ^a	46	177	I sometimes take things that belong to other people
Ho5 ^a	51	199	I sometimes tell lies to stay out of trouble
Ho6	37	214	I always tell the truth
Emotional Stability SC			
ES1 ^a	7	150	I worry more than I need to
ES2 ^a	18	202	I am a nervous person
ES3 ^a	29	175	I often feel confused and mixed up
ES4 ^a	40	115	I get upset easily
ES5 ^a	48	194	I worry about a lot of things
Academic SC			
Math SC			
Mh1	1	111	Mathematics is one of my best subjects
Mh2	12	125	I get good marks in mathematics
Mh3	23	98	I have always done well in mathematics
Mh4 ^a	34	211	I do badly in tests in mathematics

Note. SC = Self-Concept; SS = Same-Sex Relationships Self-Concept; PR = Parent Relationships Self-Concept; Ho = Honesty/Trustworthiness Self-Concept; ES = Emotional Stability Self-Concept; Mh = Math Self-Concept.

^aItem is negatively worded and was reverse-scored prior to analysis.

(continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
Verbal SC			
V ₁ ^a	6	144	I am hopeless in English classes
V ₂	17	139	Work in English classes is easy for me
V ₃	28	51	English is one of my best subjects
V ₄	39	146	I get good marks in English
V ₅	47	192	I learn things quickly in English classes
School SC			
Sch ₁ ^a	9	157	I get bad marks in most school subjects
Sch ₂	20	197	I learn things quickly in most school subjects
Sch ₃	42	83	I am good at most school subjects
Sch ₄	31	213	I do well in tests in most school subjects
Global SC			
General Self-Esteem/SC			
GS ₁	3	117	Overall, I have a lot to be proud of
GS ₂	14	180	Most things I do, I do well
GS ₃	25	70	Overall most things I do turn out well
GS ₄	36	204	I [can] do things as well as most people
GS ₅	45	171	If I really try I can do almost anything I want to
GS ₆ ^a	50	134	Overall, I am a failure
Review of Personal Effectiveness and Locus of Control			
Personal Abilities/Beliefs			
Self-Confidence			
SC ₁	10	42	I am confident that I have the ability to succeed in anything I want to do
SC ₂	24	132	When I apply myself to something I am confident I will succeed
SC ₃	39	130	I am confident in my ability to be successful
Self-Efficacy			
SF ₁	3	8	No matter what the situation is, I can handle it
SF ₂	18	96	No matter what happens I can handle it
SF ₃	32	183	I can handle things, no matter what happens
Stress Management			
SM ₁	12	58	I am calm in stressful situations
SM ₂	27	159	I can stay calm and overcome anxiety in almost all situations
SM ₃	42	173	I am calm when things go wrong

Note. SC = Self-Concept; V = Verbal Self-Concept; Sch = School Self-Concept; GS = Global Self-Esteem/Self-Concept; SC = Self-Confidence; SF = Self-Efficacy; SM = Stress Management.

^a Item is negatively worded and was reverse-scored prior to analysis.

(continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
<i>Open Thinking</i>			
OT ₁	7	26	I am open to different thinking if there is a
OT ₂	22	119	I am open to new thoughts and ideas
OT ₃	36	206	I can adapt my thinking and ideas
<i>Social Skills</i>			
<i>Social Effectiveness</i>			
SE ₁	11	47	I am effective in social situations
SE ₂	26	152	I am confident and effective in social situations
SE ₃	40	138	I communicate effectively in social situations
<i>Cooperative Teamwork</i>			
CT ₁	2	3	I like cooperating in a team
CT ₂	16	86	I cooperate well when working in a team
CT ₃	31	179	I am good at cooperating with team members
<i>Leadership Ability</i>			
LA ₁	4	12	I can be a good leader
LA ₂	19	102	I am capable of being a good leader
LA ₃	34	196	I am seen as a capable leader
<i>Organisational Skills</i>			
<i>Time Efficiency</i>			
TE ₁	14	75	I plan and use my time efficiently
TE ₂	28	162	I am efficient and do not waste time
TE ₃	43	191	I am efficient in the way I use my time
<i>Quality Seeking</i>			
QS ₁	8	32	In everything I do I try my best to get the
QS ₂	23	128	I try to get the best possible results when I do
QS ₃	38	209	I try to get the very best results in everything I
<i>Coping with Change</i>			
CH ₁	15	81	I cope well with changing situations
CH ₂	30	169	When things around me change, I cope well
CH ₃	44	203	I can cope well when things change

Note. OT = Open Thinking; SE = Social Effectiveness; CT = Cooperative Teamwork; LA = Leadership Ability; TE = Time Efficiency; QS = Quality Seeking; CH = Coping with Change. (continues)

Table A.1 (continued)
Survey Items with Cross-References to Original Instrument and Survey Instrument

Scale Item Reference	Original Item Number	Survey Item Number	Item Wording
Energy			
Active Involvement			
AI1	6	20	I prefer to be actively involved in things
AI2	20	106	I like being active and energetic
AI3	35	205	I like to get into things and make action
Overall Effectiveness			
OE1	13	63	My overall effectiveness in life is very high
OE2	29	166	Overall, in all things in life, I am effective
OE3	45	207	Overall, in my life I am a very effective person
Locus of Control			
Internal LoC			
IL1	5	16	My own efforts and actions are what will
IL2	21	112	What I do and how I do it will determine my
IL3	37	208	If I succeed in life it will be because of my
External LoC			
EL1 ^a	9	37	Luck, other people and events control most of my life
EL2 ^a	25	142	My future is mostly in the hands of other people
EL3 ^a	41	160	My life is mostly controlled by external things
Control Items			
CI1	1	2	When I have spare time I always use it to paint
CI2	17	91	I prefer things that taste sweet instead of bitter
CI3	33	189	I solve all my mathematics problems easily

Note. AI = Active Involvement; OE = Overall Effectiveness; IL = Internal Locus of Control; EL = External Locus of Control; CI = Control Items.

^a Item is negatively worded and was reverse-scored prior to analysis.

APPENDIX B
COMPARISON OVERVIEW OF ADVENTURE PROGRAMS

Table B.1
Adventure Programs: Adventure Component Descriptions

	<u>Adventure Programs</u>		
	Arctos	James Craig	Outward Bound
Total Adventure Duration	7 days (2 -day & 5-day)	7 days (2 -day & 5-day)	7 days (2 -day & 5-day)
No. Participants	8	40	16
No. Coaches	2	10	4
No. Teachers	1	5	2
No. Crew/Staff	2	16	1-2
Inter-school Participation	nil	5	2
Adventure Description	Participants learn to sail a 55 ft. yacht under the instruction of two crew. Over the duration of the adventure they take on increasing responsibility for the sailing to earn a competent crew certificate. Participants are also responsible for menu planning, meal preparation, serving and clean up.	Participants participate in sailing an 1874 restored square rigger (Tall Ship) under instruction from a large crew. Participants work in small groups to complete various sailing tasks. Other (non-sailing) activities are included to promote interaction between the different school groups. Meals are cooked and served by ship's crew.	Participants participate on a hiking-based wilderness adventure that includes activities such as rock climbing and abseiling. Outward Bound supplies all food with participants responsible for meal preparation. Participants also carry all their own equipment and construct a sleeping shelter each night.
Location	Middle Harbour Yacht Club	Pymont	Canberra

(continues)

Table B.1 (continued)

Adventure Programs: Adventure Component Descriptions

	Adventure Programs		
	Arctos	James Craig	Outward Bound
Points of difference	Small group/single school	Large group/multiple schools	Medium size group/two schools
	Increasing responsibility as a team for sailing	Students learn individual tasks of sailing a large ship	Variety of challenging physical and practical survival tasks including shelter construction, building a fire and preparing meals
	Safe sailing in varied weather	Safe sailing is weather restricted	
	Menu planning and meal preparation	Limited responsibility for meals	
		More group activities	

APPENDIX C
COMPARISON OF ADVENTURE PROGRAMS AND COACHING ONLY
PROGRAM

Table C.1

Adventure Programs and Coaching Only Program: Week-by-Week Comparison

Week	Adventure Programs	Coaching Only Program
1 Launch Event and Team Coaching	<p>Theme: setting and discovering expectations for The Helmsman Project.</p> <p>Establish trust and rapport with participants and between participants.</p> <p>Establish expectations about coaching, confidentiality, and talk about ideas of hope, self-regulation, and resilience.</p> <p>Introduce Kolb's experiential learning cycle, HAVE/DO/BE/WOW goals, and student led community project.</p>	<p>Theme: setting and discovering expectations for The Helmsman Project.</p> <p>Establish trust and rapport with participants and between participants.</p> <p>Establish expectations about coaching, confidentiality, and talk about ideas of hope, self-regulation, and resilience.</p> <p>Introduce Kolb's experiential learning cycle, HAVE/DO/BE/WOW goals, and student led community project, as well as the personal project that is core to this program.</p>
2 Individual Coaching	<p>Theme: building the foundations of an effective coaching relationship and discovering the unique individual that is the participant.</p> <p>Establish trust and rapport in a one-to-one relationship with the individual participants. Begin case conceptualisation.</p> <p>Explore strengths and values and use these to discover what is important to the participant as a person and how they see THP supporting their goals and values.</p> <p>Consider BE goals for the program.</p>	<p>Theme: building the foundations of an effective coaching relationship and discovering the unique individual that is the participant.</p> <p>Establish trust and rapport in a one-to-one relationship with the individual participants. Begin case conceptualisation.</p> <p>Explore strengths and values and use these to discover what is important to the participant as a person and how they see THP supporting their goals and values.</p> <p>Explore personal project and consider HAVE/DO/BE goals.</p>

(continues)

Table C.1 (continued)

Adventure Programs and Coaching Only Program: Week-by-Week Comparison

Week	Adventure Programs	Coaching Only Program
3 Team Coaching for A/C Program	Theme: Hope is about developing a bigger perspective, setting goals, creating pathways towards them, and developing the agency to start and keep going.	Theme: Hope is about developing a bigger perspective, setting goals, creating pathways towards them, and developing the agency to start and keep going.
Individual Coaching for CO Program	Elicit stories about life challenges and anticipated challenges for the adventure experience.	Elicit stories about life challenges and anticipated challenges for the personal project.
	Normalise adventure, uncertainty, risk, and fun.	Normalise adventure, uncertainty, risk, and fun.
	Focus on Dweck's growth (vs fixed) mindset.	Focus on Dweck's growth (vs. fixed) mindset.
		Begin to record goals and actions for the personal project.
4 Adventure for A/C Program	Theme: learning about myself in unfamiliar environments.	Theme: learning about myself in unfamiliar environments.
Workshop for CO Program	Training course to provide the necessary skills for the second adventure.	Workshop to provide necessary skills for presentation of personal project.
	"Coaching huddles" drawing on Kolb's experiential learning cycle.	"Coaching huddles" drawing on Kolb's experiential learning cycle.
5 Individual Coaching for A/C Program	Theme: identifying strengths in ourselves and others.	Theme: developing confidence in presenting, identifying strengths in others and providing support.
Preliminary Presentation for CO Program	Reflect on the experiences of the adventure and begin to identify own strengths and those of others to generalise into other aspects of their lives.	Give preliminary presentation on personal project to cement skills learned the previous week.
		Coaches facilitate curiosity about other's projects, and help participants to see strengths, make suggestions for improvements, and support each other's work.

(continues)

Table C.1 (continued)

Adventure Programs and Coaching Only Program: Week-by-Week Comparison

Week	Adventure Programs	Coaching Only Program
6 Team coaching for both	<p>No theme.</p> <p>Assist participants to reflect on learning and experiences to date and extent to which DO goals have been achieved.</p> <p>Draw out the hard work (self-regulation) required in goal striving and focus on techniques that help us self-regulate.</p> <p>Look forward to next adventure and setting HAVE and DO goals for it, and encourage participants to begin to widen their focus to WOW goal (community project).</p>	<p>Key transition point.</p> <p>Assist participants to reflect on learning and experiences to date and begin to look for ways in which the learning can be applied to life in the future.</p> <p>Begin to widen their focus to WOW goals and the community project.</p> <p>Check in about expectations for second workshop.</p>
7 Adventure for A/C Program Workshop for CO Program	<p>Theme: Learning about myself in unfamiliar environments.</p> <p>Further experiential learning opportunities to enrich awareness and development of hope, self-regulation, and resilience.</p> <p>Stretch sailing skills further, taking participants into unfamiliar and challenging territory.</p> <p>“Coaching huddles” drawing on Kolb’s experiential learning cycle, and longer periods at the end of the day for reflection and discussion.</p>	<p>Theme: What do I bring to a group/team/family/community that helps it to thrive? [same as week 8 for A/C Program]</p> <p>Further development of skills and confidence in working in collaboration with others, making plans and presenting in groups – with a focus on the community project.</p> <p>“Coaching huddles” drawing on Kolb’s experiential learning cycle, but only if appropriate and time permits (given the short duration of the workshop – 3 hours).</p>

(continues)

Table C.1 (continued)

Adventure Programs and Coaching Only Program: Week-by-Week Comparison

Week	Adventure Programs	Coaching Only Program
8 Individual coaching for A/C Program	Theme: What do I bring to a group/team/family/community that helps it to thrive? Key transition point.	Theme: How do we stretch ourselves further? What will our community project look like? [same as week 9 for A/C Program]
Team coaching for CO Program	Guide participants to shift focus from adventure to community project. Assist participants to reflect on learning and experiences to date and begin to look for ways in which the learning can be applied to life in the future and the community project. Stimulate WOW goal thinking and community project.	Facilitate and support participants to continue to develop community project idea and begin to prepare to present it at a pitch event in week 12. Ensure BE and DO goals are set. Help participants to self-organise and develop effective meeting behaviour. Ensure principal approves community project idea. Transition role from facilitator to observer and accountability partner.
9 Team coaching for both	Theme: How do we stretch ourselves further? What will our community project look like? Facilitate and support participants to continue to develop community project idea and begin to prepare to present it at a pitch event in week 12. Ensure BE and DO goals are set. Help participants to self-organise and develop effective meeting behaviour. Ensure principal approves community project idea. Transition role from facilitator to observer and accountability partner.	As per week 8

(continues)

Table C.1 (continued)

Adventure Programs and Coaching Only Program: Week-by-Week Comparison

Week	Adventure Programs	Coaching Only Program
10 Team coaching for both	As per week 9	<p>Theme: Developing confidence in presenting, identifying strengths in other and providing support.</p> <p>Participants each give a final presentation on their personal project (aim for cohort teacher to be present).</p> <p>Coaches facilitate curiosity about other's projects, and help participants to see strengths, make suggestions for improvements, and support each other's work.</p> <p>Lead group discussion and debrief at the end.</p>
11 Team coaching for both	As per week 9	<p>No theme.</p> <p>Group prepares for its pitch.</p> <p>Help students link learning and experience from personal project to pitch.</p> <p>Same approach as weeks 8 and 9.</p>
12 Pitch for both [sometimes weeks 12 and 13 are reversed]	<p>Theme: The pitch.</p> <p>Potential funding of AUD 1,000 for successful pitch.</p> <p>For the experience to be authentic, there must be a possibility of unsuccessful outcome, but this must be matched with a pathway to success.</p> <p>Pitch, panel questions, coaches huddle with participants to plan how to address questions, students present new thinking, panel feedback and vote, coach debrief and reflection.</p>	<p>Theme: The pitch.</p> <p>Potential funding of AUD 1,000 for successful pitch.</p> <p>For the experience to be authentic, there must be a possibility of unsuccessful outcome, but this must be matched with a pathway to success.</p> <p>Pitch, panel questions, coaches huddle with participants to plan how to address questions, students present new thinking, panel feedback and vote, coach debrief and reflection.</p>

(continues)

Table C.1 (continued)

Adventure Programs and Coaching Only Program: Week-by-Week Comparison

Week	Adventure Programs	Coaching Only Program
B Individual coaching for both	<p>Theme: Emphasis on continued learning and growing taking place. THP adventure may be over but journey just beginning.</p> <p>Opportunity for participants to reflect on the experience of the group work developing and pitching the community project, on the experience of the pitch, and on the program more broadly.</p> <p>Explore how participants can maintain motivation to complete the project.</p> <p>Help them make meaning of their experiences towards enhancing their future.</p>	<p>Theme: Emphasis on continued learning and growing taking place. THP adventure may be over but journey just beginning.</p> <p>Opportunity for participants to reflect on the experience of the group work developing and pitching the community project, on the experience of the pitch, and on the program more broadly.</p> <p>Explore how participants can maintain motivation to complete the project.</p> <p>Help them make meaning of their experiences towards enhancing their future.</p>

APPENDIX D RECRUITMENT PROTOCOL

THP School Selection

- 1 THP determines schools of interest (Schools) through the *My School* database (<https://www.myschool.edu.au/>), focusing on schools with an *Index of Community Socio-Educational Advantage* (ICSEA) rating less than the average. THP is initially focusing on the area of Western Sydney but could spread to other areas of socioeconomic disadvantage. THP then markets their Program to these Schools.
- 2 Prospective School completes and submits Principal's Expression of Interest. This is an online form which asks the school to consider questions such as:
 - a "What has attracted you to The Helmsman Project?"
 - b "What do you hope The Helmsman Project will deliver for your school, your teachers and your students?"
 - c "Please give examples of some of the cultural/socioeconomic challenges faced by students at your school."and to indicate the approximate proportion of students detrimentally affected by disadvantage in the local community. The form can be found at this link:
https://docs.google.com/forms/d/e/1FAIpQLScZUzhIuZhiNjF7HZX4v-LAMndbqQ2LVd2xmeyf_gwzRjfwBg/viewform
- 3 THP determines whether to proceed with the prospective School and advises the school in writing. The first three schools were selected to comply with the Better Futures, Local Solutions grant provided to the Bankstown local government area. This program focuses on 10 regional centres across Australia identified as experiencing entrenched disadvantage. It has been implemented to provide opportunities for community members to gain skills and training, access new work opportunities, and build better life outcomes for themselves and their children. Those three schools were located in lower than average socioeconomic suburbs and achieved lower than average HSC attainment.

- 4 Following the selection of the initial schools, new schools are considered on the basis of the following criteria in the listed order of priority:
 - a Does THP have sufficient funding to offer the program to a new school?
 - b Does the THP Board approve the expansion of the program to new schools?
 - c Has the school principal expressed interested in running the program?
 - d Has the principal provided sufficient evidence that their school has a reasonably significant proportion of students affected by disadvantage – noting that a school’s catchment could capture pockets of both relative affluence and disadvantage that might produce an average for the school that hides the need of those children affected by disadvantage?
- 5 THP provides School with the School Delivery Guide and an Introduction Letter describing the particular program THP will provide to that school, and related research, and attaching the Principal Consent Form. The intention is to provide the same program mode to a school over multiple years. Further, it is intended to provide the Coaching Only program mode to schools that do not also run an adventure program mode.
- 6 Principal completes and signs the Principal Consent Form if they would like to proceed, and returns the form to THP.

THP School Set-Up

- 1 School selects cohort teacher to be responsible for coordinating and overseeing weekly coaching sessions and to join the student participants on both of their adventure trips (unless they are in a coaching only program). Cohort teacher completes the cohort teacher online particulars and consent form. A link to the form can be found here:
<https://docs.google.com/forms/d/e/1FAIpQLScM9yHhZ6dMkqLEc7yxNIF3A7gs8xrl6VImOirRlkqzl74ckg/viewform?c=0&w=1>
- 2 School appoints other staff to champion the program.
- 3 School to make appropriate space available for weekly meetings with participants. This space should be private space for one-on-one and group

coaching sessions. Note that open space in the library is not deemed to be sufficiently private.

THP Participant Selection for Program

- 1 School and THP agree a date for an information and application session for all year nine students, and school invites parents and students of the whole year group to attend.
- 2 School and THP run the information and application session and provide applications to interested students at the end of the session. The online part of the application can be found here: <http://bit.ly/HelmsmanApplication>.
- 3 Interested students return a completed application (including parental consent form) to the school by the deadline. The program is voluntary, however, the School may encourage the types of students they believe would benefit from the program to apply and where necessary, help them in writing their application, but remembering that it is always the student's choice whether to participate.
- 4 From the pool of applicants, the School may create a list of eligible applicants ("Eligible Applicants"), being those applicants the School determines are most suited to participate in the Program and likely to benefit most from the Program. Schools are encouraged to consider relative effort in completing the application (e.g., a (seemingly) poorly completed application by a student who never completed homework may be more effortful than a higher performing student with access to parental or other adult assistance). It is important that the program remain an aspirational program and never a remedial program for bad or poor performing students.
- 5 THP reviews the pool of Eligible Applicants and paperwork to confirm participation ("Selected Applicants"). If the number of Eligible Applicants exceeds the number of available places, Eligible Applicants will be allocated to places in the Program based on a draw. To date, this has not been required.

THP Participant Program Allocation

- 1 Selected Applicants are separated into groups through the following steps:

- a For **single gender schools**, Selected Applicants are randomly assigned to intervention or control by the school or a THP employee with no knowledge of the Selected Applicants.
 - b For **mixed gender schools**, Selected Applicants are first separated into gender groups (owing to the age of the students and other cultural sensitivities, there are no mixed gender groups), and then one gender group is assigned to intervention and the other to control. This group assignment is generally random, however, it may depend on the gender of the cohort teacher and coachees, and their availability (female groups need to have female coaches and a female cohort teacher), and also where we have had to rebalance gender vis-à-vis the various program modes and between intervention/control. Where there are four groups – two for each gender, then the two groups for each gender are randomly allocated between intervention and control by the school or a THP employee.
- 2 Intervention groups and control groups (for when they become intervention) are assigned to the same program mode based upon the program mode being offered.

THP Participant Selection for Qualitative Interview

- 1 Schools are requested to invite previous participants to participate in an interview, describing the interview and asking for anyone interested to participate.
- 2 If more than the required number of interviewees respond, then the respondents will be placed in categories based on the program mode in which they participated. Two names will be drawn for each program mode and an additional name will be drawn for each, as a back-up interviewee.
- 3 School to schedule interviews to take place at the school on the same day.
- 4 Interviewer to have the student sign a consent form prior to conducting the interview.

APPENDIX E
SELF-REPORTED DEMOGRAPHIC DETAILS OF THP PROGRAM PARTICIPANTS

Table E.1
Breakdown of Participants across THP Programs and Intervention/Control

	Arctos	James Craig	Outward Bound	Coaching Only	Adv. (Total)	Total
Total Participants	110	101	89	62	300	362
Male	33	68	44	30	145	175
Female	77	33	45	32	155	187
Intervention	59	54	45	28	158	186
Male	6	38	16	12	60	72
Female	53	16	29	16	98	114
Control	51	47	44	34	142	176
Male	27	30	28	18	85	101
Female	24	17	16	16	57	73

Note. THP = The Helmsman Project; Arctos = Arctos Adventure Program; James Craig = James Craig Adventure Program; Outward Bound = Outward Bound Adventure Program; Adv = Adventure Programs (taken together).

Table E.2
Breakdown of Participants across Schools

Schools	Cohorts	Total Participants	Participants by Group	
			Intervention	Control
A	14	93	48	45
B	7	46	27	19
C	5	32	21	11
D	2	12	6	6
E	4	16	6	10
F	8	63	30	33
G	4	30	15	15
H	2	13	8	5
I	2	16	7	9
J	2	14	6	8
K	4	27	12	15

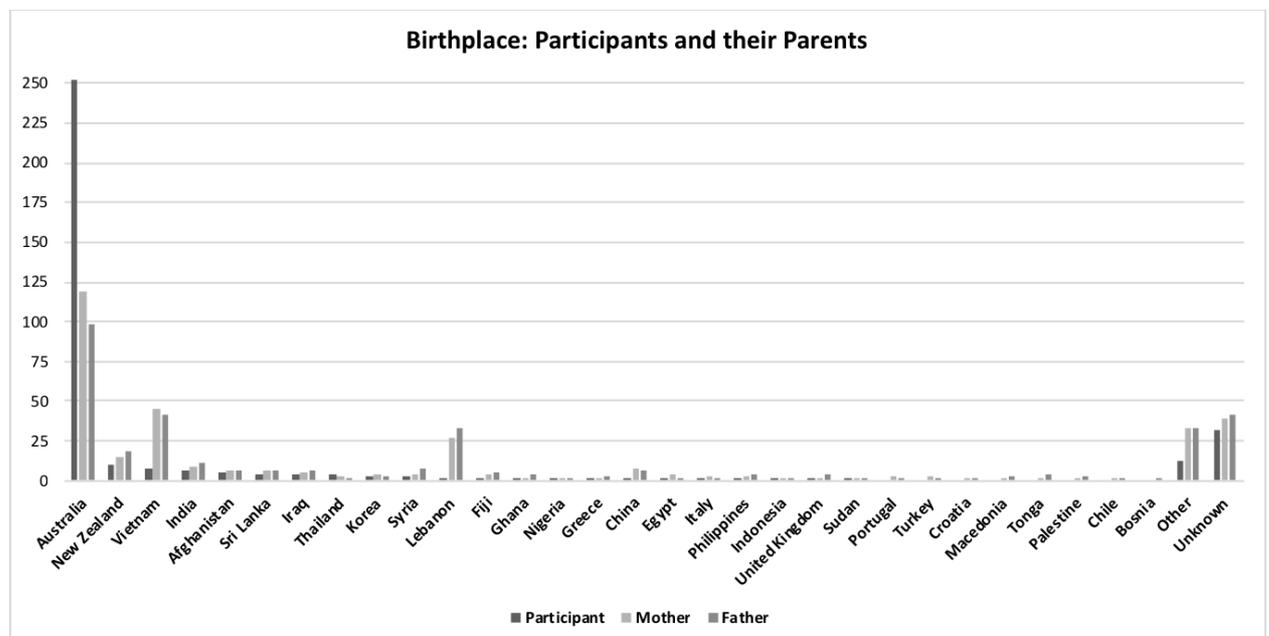
Note. School names are replaced by letters for confidentiality.

Table E.3
Breakdown of Participants by Other Factors

Does participant have...?	Intervention			Control		
	Yes	No	NR	Yes	No	NR
Dishwasher (percentage)	87 46.77%	93 50.00%	6 3.23%	80 45.45%	67 38.07%	29 16.48%
Dictionary (percentage)	159 85.48%	19 10.22%	8 4.30%	125 71.02%	22 12.50%	29 16.48%
Internet (percentage)	148 79.57%	33 17.74%	5 2.69%	121 68.75%	26 14.77%	29 16.48%
Own Room (percentage)	118 63.44%	63 33.87%	5 2.69%	105 59.66%	45 25.57%	26 14.77%
Study Space (percentage)	140 75.27%	41 22.04%	5 2.69%	117 66.48%	31 17.61%	28 15.91%
Previous Experience OAE (percentage)	146 78.49%	34 18.28%	6 3.23%	116 65.91%	35 19.89%	25 14.20%

Note. NR = no response.

Table E.4
Breakdown of Participants by Birthplace



APPENDIX F
SURVEY: PARTICIPANT CONSENT AND DEMOGRAPHIC QUESTIONS
(QUALTRICS ONLINE VERSION)



PARTICIPANT CONSENT FORM

Purpose

The purpose of this survey is to help to find out what you think about your schoolwork, yourself, your school and others. Your participation in the study is voluntary and you can withdraw from the study at any time. Not participating in the study will not affect your relationship with your school. This is not a test. There are no right or wrong answers and everybody will have different answers. Just make sure that your answers show what you really think about yourself. I will read the questions aloud to you and explain how to answer each one. There are some questions that seem the same. This is not a trick. It is just that this type of survey needs to ask questions in slightly different ways. Just answer them in a way that shows what you really think about yourself. Your answers will only be seen by the researchers and will not be shown to anyone in your school or your community. The researchers will remove the consent form you sign below and store this separately. The research team will not report the names of students or schools that participate in the study.

Do you agree to participate in the study?

- Yes
- No

Student's Name (First name and Surname):

Date of Birth (DD/MM/YYYY) for example: 12th December 2002 = 12/12/2002

Today's Date (DD/MM/YYYY):

Your Background

What is the name of your school?

What year/grade are you currently in at school? (e.g. Year 9)

What is the month that you were born?

What is the year that you were born?

Have you ever been on an outdoor or adventure camp?

- Yes
- No

Are you male or female?

- Male
- Female

Where were YOU born?

- Australia
- China
- Croatia
- Fiji
- Greece
- Hong Kong
- India
- Indonesia
- Italy
- Korea
- Lebanon
- Macedonia
- Malaysia
- New Zealand
- Philippines
- Singapore
- Sri Lanka
- United Kingdom
- United States
- Other: name country _____

If you did not live in Australia when you were born, in what YEAR did you move to Australia?

In which country was your MOTHER (or STEPMOTHER) born?

- Australia
- China
- Croatia
- Fiji
- Greece
- Hong Kong
- India
- Indonesia
- Italy
- Korea
- Lebanon
- Macedonia
- Malaysia
- New Zealand
- Philippines
- Singapore
- Sri Lanka
- United Kingdom
- United States
- Other: name country _____

In which country was your FATHER (or STEPFATHER) born?

- Australia
- China
- Croatia
- Fiji
- Greece
- Hong Kong
- India
- Indonesia
- Italy
- Korea
- Lebanon
- Macedonia
- Malaysia
- New Zealand
- Philippines
- Singapore
- Sri Lanka
- United Kingdom
- United States
- Other: name country _____

What language does your family speak the most at home?

- English
- Croatian
- Italian
- Mandarin
- Thai
- Arabic/Lebanese
- Greek
- Japanese
- Samoan
- Tongan
- Cantonese
- Indian
- Korean
- Serbian
- Vietnamese
- Chinese
- Indonesian
- Macedonian
- Sri Lankan
- Other: Please specify _____

What is the highest level of education that your parents/guardians have?

	None	Primary School	Some High School	Completed High School	TAFE	University
Mother/female guardian	<input type="radio"/>					
Father/male guardian	<input type="radio"/>					

Which of the following are in your home?

	Yes	No
A room of your own	<input type="radio"/>	<input type="radio"/>
A quiet place to study	<input type="radio"/>	<input type="radio"/>
A computer you can use for school work	<input type="radio"/>	<input type="radio"/>
A high speed internet connection	<input type="radio"/>	<input type="radio"/>
A dictionary	<input type="radio"/>	<input type="radio"/>
A dishwasher	<input type="radio"/>	<input type="radio"/>

APPENDIX G
THP SURVEY: QUALTRICS ONLINE VERSION

PARTICIPANT QUESTIONNAIRE



Student's Name:

School:

Please choose one response for each of the following questions:

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
I think I'm doing pretty well	<input type="radio"/>				
When I have spare time I always use it to paint	<input type="radio"/>				
I like cooperating in a team	<input type="radio"/>				
I forget about whatever else I need to do when I'm doing something really fun	<input type="radio"/>				
I believe I am mentally tough when it comes to overcoming life challenges	<input type="radio"/>				
When I'm sad, I can usually start doing something that will	<input type="radio"/>				

make me feel
better

No matter
what the
situation is, I
can handle it

I believe I'm
mentally
tough when
it comes to
exams

Little
problems
detract me
from my
long-term
plans

I can be a
good leader

I can resist
doing
something
when I know
I shouldn't
do it

I can think of
many ways to
get things in
life that are
most
important to
me

My own
efforts and
actions are
what will
determine
my future

I don't let
study stress
get on top of
me

I can stay
focused on
my work
even when
it's dull

I prefer to be actively involved in things	<input type="radio"/>				
It's hard for me to notice when I have had enough (sweets, food, etc.)	<input type="radio"/>				
I've been feeling relaxed	<input type="radio"/>				
If I really want something, I have it right away	<input type="radio"/>				
I'm good at bouncing back from a poor mark in my schoolwork	<input type="radio"/>				
I am open to different thinking if there is a better idea	<input type="radio"/>				
I am doing just as well as other kids my age	<input type="radio"/>				
If something isn't going according to my plans, I change my actions to try and reach my goal	<input type="radio"/>				
I don't usually let life stresses get on top of me	<input type="radio"/>				
In everything I do I try my	<input type="radio"/>				

best to get the details right					
I lose control whenever I don't get my own way	<input type="radio"/>				

Please choose one response for each of the following questions:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
During a dull task, I have trouble forcing myself to start paying attention	<input type="radio"/>				
I think I'm good at dealing with schoolwork pressures	<input type="radio"/>				
Luck, other people and events control most of my life	<input type="radio"/>				
When I have a problem, I can come up with lots of ways to solve it	<input type="radio"/>				
I'm good at bouncing back from disappointments in my life	<input type="radio"/>				
I can find ways to make myself study, even when my friends want to go out	<input type="radio"/>				
I am confident that I have the ability to succeed in anything I want to do	<input type="radio"/>				
After I'm interrupted or	<input type="radio"/>				

distracted, I can easily continue working where I left off					
I often feel like giving up in my schoolwork	<input type="radio"/>				
I think the things I have done in the past will help me in the future	<input type="radio"/>				
I am effective in social situations	<input type="radio"/>				
When I'm bored, I fidget or can't sit still	<input type="radio"/>				
I think I'm good at dealing with sources of pressure in my life	<input type="radio"/>				
English is one of my best subjects	<input type="radio"/>				
Overall I expect more good things	<input type="radio"/>				

Please choose one response for each of the following questions:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I am awkward at things like sports, gym and dance	<input type="radio"/>				
I am a hard worker	<input type="radio"/>				
I lose track of the time when I'm doing something fun	<input type="radio"/>				
In my schoolwork I sometimes reduce my chances of doing	<input type="radio"/>				

well (e.g. waste time, not study, disrupt others, procrastinate)

I am calm in stressful situations

If there are other things going on around me, I find it hard to keep my attention focused on whatever I'm doing

Even when others want to quit, I know that I can find ways to solve the problem

I sometimes cheat

I have trouble getting excited about something that's really special when I am tired

My overall effectiveness in life is very high

It's hard for me to get started on big projects that require planning in advance

I don't let difficulties and disappointments in life affect my confidence

I rarely count on good things happening to me

I've been feeling useful	<input type="radio"/>				
It's hard for me to keep focused on something I find unpleasant or upsetting	<input type="radio"/>				
Overall most things I do turn out well	<input type="radio"/>				
I can usually act normal around everybody if I am upset with someone	<input type="radio"/>				
I do not get along very well with boys	<input type="radio"/>				
I'm good at dealing with Setbacks (e.g. negative feedback on what I do, disappointing outcomes)	<input type="radio"/>				
I plan and use my time efficiently	<input type="radio"/>				
I lose control whenever I don't get my own way	<input type="radio"/>				
I often set a goal but later choose to pursue a different one	<input type="radio"/>				

Please choose one response for each of the following questions:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I don't think I have much control over how well I do in my schoolwork	<input type="radio"/>				

I never know how much more work I have to do	<input type="radio"/>				
I cope well with changing situations	<input type="radio"/>				
I get carried away emotionally when I get excited about something	<input type="radio"/>				
I am good at most school subjects	<input type="radio"/>				
I don't get upset too easily	<input type="radio"/>				
New ideas and projects sometimes distract me from previous ones	<input type="radio"/>				
I cooperate well when working in a team	<input type="radio"/>				
I am good at keeping track of lots of things going on around me, even when I'm feeling stressed	<input type="radio"/>				
Other people think I am good looking	<input type="radio"/>				
When I have a serious disagreement with someone, I can talk	<input type="radio"/>				

calmly about it without losing control					
Setbacks don't discourage me	<input type="radio"/>				
I prefer things that taste sweet instead of bitter	<input type="radio"/>				
I can start a new task, even if I'm already tired	<input type="radio"/>				
I have been obsessed with a certain idea or project for a short time but later lost interest	<input type="radio"/>				
I mainly do my schoolwork to avoid failing or disapproval from parents or the teacher/s	<input type="radio"/>				
I hardly ever expect things to go my way	<input type="radio"/>				
No matter what happens I can handle it	<input type="radio"/>				
I am diligent	<input type="radio"/>				
I have always done well in mathematics	<input type="radio"/>				
It's hard to start making plans to deal with a big project or	<input type="radio"/>				

problem, especially when I'm feeling stressed					
I do not like my parents very much	<input type="radio"/>				
I get quite anxious about schoolwork and tests	<input type="radio"/>				
I am capable of being a good leader	<input type="radio"/>				

Please choose one response for each of the following questions:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
When I'm having a tough day, I stop myself from whining about it to my family and friends	<input type="radio"/>				
In most ways my life is close to ideal	<input type="radio"/>				
I can calm myself down when I'm excited or all wound up	<input type="radio"/>				
I like being active and energetic	<input type="radio"/>				
I don't let a bad mark affect my confidence	<input type="radio"/>				
I've been feeling optimistic about the future	<input type="radio"/>				

It's important for me to keep busy	<input type="radio"/>				
I persist at schoolwork even when it is challenging or difficult	<input type="radio"/>				
Mathematics is one of my best subjects	<input type="radio"/>				
What I do and how I do it will determine my successes in life	<input type="radio"/>				
I usually know when I'm going to start crying	<input type="radio"/>				
I have a nice looking face	<input type="radio"/>				
I get upset easily	<input type="radio"/>				
When I'm excited about reaching a goal (e.g. getting my driver's license, going to college), it's easy to start working on it	<input type="radio"/>				
Overall, I have a lot to be proud of	<input type="radio"/>				
I'm good at dealing with setbacks at school (e.g. bad mark, negative feedback on my work)	<input type="radio"/>				

I am open to new thoughts and ideas	<input type="radio"/>				
I can stop myself from doing things like throwing objects when I'm mad	<input type="radio"/>				
I use my study / homework time well and try to study and do homework under conditions that will bring out my best	<input type="radio"/>				
It is difficult to make friends with members of my own sex	<input type="radio"/>				
I have difficulty maintaining my focus on projects that take more than a few months to complete	<input type="radio"/>				
I've been feeling good about myself	<input type="radio"/>				
I get good marks in mathematics	<input type="radio"/>				
I enjoy my friends a lot	<input type="radio"/>				
I am honest	<input type="radio"/>				

Please choose one response for each of the following questions:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I try to get the best possible results when I do things	<input type="radio"/>				
I can find a way to stick with my plans and goals, even when it's tough	<input type="radio"/>				
I am confident in my ability to be successful	<input type="radio"/>				
I finish whatever I begin	<input type="radio"/>				
When I apply myself to something I am confident I will succeed	<input type="radio"/>				
I work carefully when I know something will be tricky	<input type="radio"/>				
Overall I am a failure	<input type="radio"/>				
I enjoy things like sports, gym and dance	<input type="radio"/>				
I do not get along very well with girls	<input type="radio"/>				
When I have a big project, I can keep working on it	<input type="radio"/>				
I communicate successfully	<input type="radio"/>				

in social situations					
Work in English classes is easy for me	<input type="radio"/>				
I plan out how I will do my schoolwork and study	<input type="radio"/>				
I've been feeling cheerful	<input type="radio"/>				
My future is mostly in the hands of other people	<input type="radio"/>				
I'm usually aware of my feelings before I let them out	<input type="radio"/>				
I am hopeless in English classes	<input type="radio"/>				
The conditions of my life are excellent	<input type="radio"/>				
I get good marks in English	<input type="radio"/>				
I'm always optimistic about my future	<input type="radio"/>				
I am better than most of my friends at sports, gym and dance	<input type="radio"/>				
I am satisfied with life	<input type="radio"/>				
I worry more than I need to	<input type="radio"/>				

I can usually tell when I'm getting tired or frustrated	<input type="radio"/>				
I am confident and effective in social situations	<input type="radio"/>				
I get along well with my parents	<input type="radio"/>				

Please choose one response for each of the following questions:

	Strongly Agree	Agree	Neutral	Disagree	Strongly disagree
In class, I can concentrate on my work even if my friends are talking	<input type="radio"/>				
I make friends easily with girls	<input type="radio"/>				
I've been interested in new things	<input type="radio"/>				
I get bad marks in most school subjects	<input type="radio"/>				
In my schoolwork, I am focused on learning and improving more than competing and being the best	<input type="radio"/>				
I can stay calm and overcome anxiety in almost all situations	<input type="radio"/>				

My life is mostly controlled by external things	<input type="radio"/>				
I've been able to make up my own mind on things	<input type="radio"/>				
I am efficient and do not waste time	<input type="radio"/>				
If something can go wrong for me, it will	<input type="radio"/>				
I am not very popular with members of the opposite sex	<input type="radio"/>				
I've been feeling loved	<input type="radio"/>				
Overall, in all things in life, I am effective	<input type="radio"/>				
So far I have gotten the important things I want in life	<input type="radio"/>				
I have lots of friends of the opposite sex	<input type="radio"/>				
When things around me change, I cope well	<input type="radio"/>				
I've been feeling interested in other people	<input type="radio"/>				
If I really try I can do almost anything I want to	<input type="radio"/>				

I am good looking	<input type="radio"/>				
I am calm when things go wrong	<input type="radio"/>				
What I learn in my schoolwork is important and useful	<input type="radio"/>				
I often feel confused and mixed up	<input type="radio"/>				
I've had energy to spare	<input type="radio"/>				
I sometimes take things that belong to other people	<input type="radio"/>				
It's easy for me to relax	<input type="radio"/>				
I am good at cooperating with team members	<input type="radio"/>				

Please choose one response for each of the following questions:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Most things I do, I do well	<input type="radio"/>				
I've been feeling close to other people	<input type="radio"/>				
I make friends easily with boys	<input type="radio"/>				
I can handle things, no matter what happens	<input type="radio"/>				
If I could lead my life over, I	<input type="radio"/>				

would change almost nothing					
My parents treat me fairly	<input type="radio"/>				
I often tell lies	<input type="radio"/>				
I've been dealing with problems well	<input type="radio"/>				
My parents understand me	<input type="radio"/>				
I solve all my mathematics problems easily	<input type="radio"/>				
I believe I can do a good job in my schoolwork	<input type="radio"/>				
I am efficient in the way I use my time	<input type="radio"/>				
I learn things quickly in English classes	<input type="radio"/>				
I am good at things like sports, gym and dance	<input type="radio"/>				
I worry about a lot of things	<input type="radio"/>				
I've been thinking clearly	<input type="radio"/>				
I am seen as a capable leader	<input type="radio"/>				
I learn things quickly in most school subjects	<input type="radio"/>				

I've been feeling confident	<input type="radio"/>				
I sometimes tell lies to stay out of trouble	<input type="radio"/>				
In uncertain times, I usually expect the best	<input type="radio"/>				
I make friends easily with members of my own sex	<input type="radio"/>				
I am a nervous person	<input type="radio"/>				
I can cope well when things change	<input type="radio"/>				
I do things as well as most people	<input type="radio"/>				
I like to get into things and make action	<input type="radio"/>				
I can adapt my thinking and ideas	<input type="radio"/>				

Please choose one response for each of the following questions:

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Overall, in my life I am a very effective person	<input type="radio"/>				
If I succeed in life it will be because of my efforts	<input type="radio"/>				

I try to get the very best results in everything I do	<input type="radio"/>				
I have a good looking body	<input type="radio"/>				
I do badly in tests in mathematics	<input type="radio"/>				
Not many people of my own sex like me	<input type="radio"/>				
I do well in tests in most school subjects	<input type="radio"/>				
I always tell the truth	<input type="radio"/>				
I completed all of the questions in this survey carefully and honestly	<input type="radio"/>				

APPENDIX H ETHICS APPROVAL REPORT

PERSON ETHICS APPROVAL REPORT

Gwyn, Wendy (S00214727)

Ethics Approvals

2 ethics applications found

ECODE	ETHICS APPLICATION TITLE	STATUS	APPLIED DATE	APPROVED DATE	START DATE	END DATE	RISK
2014-137Q	<input checked="" type="checkbox"/> The Helmsman Project: Evaluating the efficacy of a combined outdoor education and coaching program on the life-effectiveness skills of at-risk adolescents	Completed	28/04/2014	2/05/2014	3/05/2014	30/06/2018	Low Risk
2017-252HE	<input checked="" type="checkbox"/> The Helmsman Project: Evaluating the efficacy of a combined outdoor education and coaching program on the life-effectiveness skills of at-risk adolescents *please note that this is a resubmitted protocol from approved 2014-137Q. The Helmsman Project: Giving at-risk adolescents skills to navigate life's journey and make a difference (LP140100041) ending 12.12.18	Approved	18/10/2017	3/11/2017	4/11/2017	30/06/2020	Low Risk

APPENDIX I
THP APPLICATION AND CONSENT FORM

THE HELMSMAN
PROJECT

Young people. Making a difference.

Application invitation

Please type this URL into a web browser to complete your application.
This will work on desktops and smartphones

<http://bit.ly/HelmsmanApplication>



What's your potential?

The Helmsman Project is a combined coaching and adventure education program designed to help you develop resilience, self-regulation, hope, determination and many more skills that can help you give and get the maximum from life.

The program goes for one year and is run during school terms.

You are invited to apply to the program to gain one of the limited spaces. Applications will be assessed on the effort you put into the application. So do your best to answer the questions fully, and give yourself the best chance at a spot!

Discover it, become it!

(continues)

What's involved?

For the first two terms of the program, you will participate in an individual and group coaching process. This involves conversations with trained volunteer coaches at your school as well as two adventure experiences designed to get you thinking way outside your box!

Towards the end of these two terms, you will work with your group at designing a community project. This is something that you should all feel excited about and can bring benefits to many. You will then pitch your idea of community project to a panel of judges. If your idea receives a green light, there will be up to **\$1,000** available to your group to make your project happen. In the following two terms, you will work with your group to make your community project come to life. When you achieve it, you will become a graduate of The Helmsman Project and have an incredible story to tell in job interviews, further education applications and to friends and family!

Adventure education?

A key part of The Helmsman Project is your participation in an adventure experience. This is essential to getting you thinking beyond what you currently think is possible for you. Adventure experiences take on many forms but usually involve going to a wilderness environment and experiencing nature, working together as groups and getting out of your comfort zone. On The Helmsman Project, you might be on a course with your group learning to sail a yacht or tall ship, then taking it on a voyage. Or you might learn bush survival techniques and navigation, and take an expedition through the bush - Bear Grylls style, but with better food! Your school will know which adventure experience your group will be going on.

Application questions

Below are the questions you will be asked to answer on your application form. It's important that you have a good think about your answers before you start filling out the online form.

- If you have been on an adventure education program or camp, please briefly outline what activities it involved and what you enjoyed about it.
- What is it about The Helmsman Project that interests you most?
- What is it about The Helmsman Project that you are most apprehensive about?
- If you are selected to participate in The Helmsman Project, what do you hope to get out of the program?
- If you are selected to participate in The Helmsman Project, what do you think you can bring to the program - what are you good at?
- What is the most significant challenge you have faced in your life so far and how have you managed to address it?
- If there was one thing you could change in the world, what would it be? Please explain your answer.
- What is the most important thing you would like to achieve in your life?
- Participating on The Helmsman Project is going to require you to be really engaged. You will learn a lot and have fun, but it will be challenging.
- What will you do to get the most out of the program?

(continues)

Frequently Asked Questions

Q. What do participants gain from the program?

A. The Helmsman Project is designed to be both fun and challenging. Participants receive individual and group coaching and take part in two adventure education expeditions to help complement their learning process. They work on goal setting skills, self-belief, resilience, team-working skills, accountability and much more. The community project phase of the program helps participants gain job-readiness skills, such as communication and conflict resolution skills, leadership, project management, teamwork and presentation skills – all attributes that will help participants navigate their way to success!

Q. How much does it cost to participate in the program?

A. The Helmsman Project is generously supported by corporate sponsors, individuals and volunteers, meaning the program is provided free of charge for individual participants. [On a fully costed basis, the value of the program offered by The Helmsman Project to participants is around \$10,000 per participant.]

Q. How challenging is it to participate in the program?

A. The program includes a series of goals and challenges, which are challenging enough for participants to feel stretched - that is the reason for incorporating them in the program. It is critical that they are not so stretching that participants feel they cannot complete the challenges, so we have been careful in designing them. Participants will receive lots of guidance and encouragement as they tackle challenges. In the end, no one will be forced to do anything that they decide that they do not wish to do.

Q. What is coaching?

A. Coaching is not teaching, but it is about learning. The coach's role is to help participants think about what they want to achieve and how they can make it happen. The coach supports participants in their journey, challenges them to think more broadly and helps them to apply their learning to different areas of their life.

Q. Is the adventure education part of the program compulsory?

A. Adventure education is a core component of The Helmsman Project, which must be completed in order to continue the journey and graduate from the program. Only participants who agree to take part in the adventure education trips are selected to take part in the program. Not attending these trips will lead to withdrawal from The Helmsman Project.

Q. How will the safety of participants be looked after while on the adventure education blocks?

A. The safety of all involved in the adventure is our number one priority.

The Helmsman Project undertakes regular comprehensive risk assessments with regard to the risks to individuals participating in the adventure education programs. The providers are expert operators in their fields and use their own risk management systems and operating procedures, which have been closely audited by The Helmsman Project and other relevant peak bodies. Student participants, supporting teachers and coaches will be briefed on key safety procedures or learn them as part of their training course.

(continues)

(THP Application and Consent Form continued)



Frequently Asked Questions (*cont*)

Q. What does the adventure education component of the program entail?

A. Participants experience two adventure education trips: a two-day (one-night) trip will occur during week 4 of the program and a five-day (four-night) trip during week 7. Participants, their cohort teachers, coaches and highly experienced adventure education instructor will embark on the adventure education trip.

During the first two-day experience, participants will engage in an outdoor adventure course that will teach them the skills required to make the most of their five-day experience. Students will be with their school team (single-gender groups only) and it is critical they take part in both adventures.

Although this adventure is fun and rewarding, it is also a challenging part of the program where participants' learnings from their coaching sessions are put to the test.

The Helmsman Project, along with the adventure education providers, will ensure participants have everything they need for their safety. Students are provided with a list of things they will need to bring to ensure they are well prepared and adequately dressed for the expedition.

Q. How will participants be able to catch up on the time missed at school?

A. It is the participants' responsibility to catch up on lessons and homework missed while attending the coaching sessions or adventure education trips. In our experience, schools have been very supportive of their students participating in the program. Teachers will be able to provide guidance on how to best manage ongoing studies and achieve academic goals whilst participating in the program.

For any further information, please contact:

The Helmsman Project
General Manager
Kim Larochelle: kim.larochelle@thehelmsmanproject.org.au

(continues)

(THP Application and Consent Form continued)



THE HELMSMAN PROJECT – INDEMNITY, RELEASE & WAIVER OF LIABILITY

This form must be returned to your school for your online application to be accepted

I am the parent or legal guardian of _____

This document is an important legal document. You must read it and sign it before your child can take part in the Helmsman Project activities.

We are pleased that your child is participating in the Helmsman Project. Your child will be involved in individual and group conversations facilitated by trained volunteer coaches taking place on school premises, during school hours. Your child may also involve themselves in activities off school premises and out of school hours as a result of undertaking their group community project. These projects will have been approved by the school for their safety to students and third parties but as they are conceived by the students the precise nature of the activities cannot be determined at this stage. By signing this document, you agree that you and your child are aware of these risks and your child is voluntarily participating in The Helmsman Project activities at his or her own risk.

Acknowledgement

I, the undersigned, am the parent/guardian [*delete as appropriate*] of _____
and hereby declare that both my child and I have read and agree to the terms of the above Indemnity, Release & Waiver of Liability.

Date _____

Signature _____

I understand that video or photographic recordings of the program may be made from time to time during the program. A small number of these recordings / images may be used to promote the program to other potential participants from other schools. Participants' names will not be identified, nor will the children's school be directly linked to photographs. Please put a line through this paragraph if you do not consent to your child's face being identifiable in video or photographic recordings.

(continues)

(THP Application and Consent Form continued)


Institute for Positive Psychology & Education


INFORMATION FOR PARENTS

The Helmsman Project is a new type of student development program and, as a result, is undertaking research to demonstrate the benefits of the program for students. Please read below for more detailed information about the program and the research efforts.

Dear Parent / Guardian,

The Helmsman Project: Research study to evaluate the efficacy of a combined outdoor education and coaching program designed to build the life effectiveness skills of adolescents

We wish to invite your child to participate in the above research study conducted by the Australian Catholic University (ACU), application 2014-137Q.

Research suggests that both personal coaching programs and adventure education activities have the potential to have a positive developmental impact on adolescents. Building upon research in the field, an innovative program has been developed to create a combined personal coaching and adventure education program targeting Year 9 students: The Helmsman Project (THP). THP is a 13-week program, delivered on the school premises, with the exception of two adventure education trips (the first being a two-day trip and the second a five-day trip) and a graduation ceremony at the end of the year.

This study seeks to gather and evaluate data from you and your child, the school, and the teachers and coaches involved in the program, with a view to evaluating the efficacy of the above developmental program.

This research project has been approved by the State Education Research Approval Process (SERAP No. 2013134). Submitting an application to participate in The Helmsman Project is voluntary and places will be limited.

If you would like your child to apply to participate in The Helmsman Project and the associated research, please fill out the consent form attached hereto and return it to your child's year group teacher.

Yours sincerely,

Distinguished Professor Herb Marsh
 Centre for Positive Psychology in Education, Australian Catholic University
 25A Barker Road, Strathfield NSW 2135, Australia.
 email: herb.marsh@acu.edu.au

Institute for Positive Psychology & Education
 Australian Catholic University 25A Barker Road Strathfield NSW 2135 Australia
 T: +61 (0)2 9701 4626 E: ippe@acu.edu.au W: ippe.acu.edu.au

Australian Catholic University and the courses offered by the University are registered on the Commonwealth Register of Institutions and Courses for Overseas Students (CRICOS), Provider registration codes: 00004G, ABN: 15 050 192 660

**THE HELMSMAN
 PROJECT**
Young people. Making a difference.

(continues)

(THP Application and Consent Form continued)


Institute for Positive Psychology & Education


PARTICIPANT INFORMATION LETTER

PROJECT TITLE: The Helmsman Project: Giving at-risk adolescents skills to navigate life's journey and fulfil their potential.
APPLICATION NUMBER: 2014-137Q
PRINCIPAL INVESTIGATOR: Distinguished Professor Herb Marsh
STUDENT RESEARCHER: Wendy Gwyn

What is the project about?

The Helmsman Project (THP) seeks to gather and evaluate data with a view to evaluating the efficacy of combined personal coaching and adventure education program targeting Year 9 students: The Helmsman Project (THP). It is our hope that the findings resulting from this research will have a meaningful impact on the way in which adolescents are encouraged to remain engaged at school through year 12 and are prepared for a successful transition to further education, training or employment thereafter.

THP is a 13-week program, delivered on the school premises, with the exception of two adventure education trips (the first being a two-day trip and the second a 5-day trip) and a graduation ceremony at the end of the year. The school-based component, running over 13 weeks will comprise a launch event and ten weekly group or individual coaching sessions with professional personal coaches. Adventure experiences take on many forms but usually involve going to a wilderness environment and experiencing nature, working together as groups. On The Helmsman Project, you might be on a course with your group learning to sail a yacht or tall ship, then taking it on a voyage. Or you might learn bush survival techniques and navigation, and take an expedition through the bush with experienced instructors, a teacher from your child's school and the coaches who your child will have got to know through the coaching in the lead up to the first adventure education event.

There are limited spaces available for participants on The Helmsman Project research program and we do not expect that everyone who applies will be able to be accommodated on the program. As a result, based on the quality of the submitted applications, the school will make recommendations to the Helmsman Project team on which students are most suited to participate in The Helmsman Project research program. If the number of qualified applicants exceeds the number of available places, qualified applicants will be allocated places on the research program based on a draw.

As a result of the structure of the research, it is possible that groups of children will participate in certain parts of The Helmsman Project program in a different sequence to that experienced by other children.

Who is undertaking the project?

This project is being conducted by Wendy Gwyn at Australian Catholic University under the supervision of Distinguished Professor Herb Marsh who is one of the most productive educational psychologist in the world.

Are there any risks associated with participating in this project?

The safety of all involved in the adventure is our number one priority. The Helmsman Project undertakes regular comprehensive risk assessments with regard to the risks to individuals participating in the adventure education programs. The providers are expert operators in their fields and use their own risk management systems and operating procedures, which have been closely audited by The Helmsman Project and other

Institute for Positive Psychology & Education
 Australian Catholic University 25A Barker Road Strathfield NSW 2135 Australia
 T: +61 (0)2 9701 4626 E: ippe@acu.edu.au W: ippe.acu.edu.au

Australian Catholic University and the courses offered by the University are registered on the Commonwealth Register of Institutions and Courses for Overseas Students (CRICOS), Provider registration codes: 00004G, ABN: 15 050 192 660

**THE HELMSMAN
 PROJECT**
Young people. Making a difference.

(continues)



relevant peak bodies. Student participants, supporting teachers and coaches will be briefed on key safety procedures or learn them as part of their training course.

In the unlikely event that a student should feel any sort of concern as a result of participation in The Helmsman Project research program, they can meet with the school counsellor or the teacher who is accompanying the children on The Helmsman Project journey. If you or your child should experience any discomfort while participating in this project and you feel you would benefit from further support, assistance is available:

Lifeline - 13 11 14

Kids helpline - 1800 55 1800

What will I be asked to do?

Your child will be asked to voluntarily complete a number of surveys prior to the program, throughout the course of the program, and three months after program completion. Completion of these surveys will be coordinated by the school and The Helmsman Project. Additionally, your child may also be asked if he or she would like to voluntarily participate in a follow-up study, a tape-recorded interview to be conducted after completion of the program.

How much time will the project take?

The Helmsman Project is a 13-week program, delivered on the school premises, with the exception of two adventure education trips (the first being a two-day trip and the second a 5-day trip) and a graduation ceremony at the end of the year. The school-based component, running over 13 weeks will comprise a launch event and ten weekly group or individual coaching sessions with professional personal coaches.

What are the benefits of the research project?

We hope that your child benefits from participating in THP and it is anticipated that the information we collect from this research study will benefit future program participants.

Can I withdraw from the study?

Participation in this project is completely voluntary. Your child does not have to participate. If your child agrees to participate, he/she can withdraw from the study at any time without negative consequences. If you decide to withdraw you will be given the choice of having the information you and your child have provided deleted or kept on file for the rest of the study.

Will anyone else know the results of the project?

The data collected in this study is completely confidential and cannot lead to identification. Information about the program and your child may also be collected from the school, the teachers and the coaches involved in program. Information collected during the course of this study will not be given to others. The details of this undertaking have been agreed with your child's school. However, researchers are required by the Department of Education and Communities to ensure that *"when studies have the potential to identify students as being at risk of harm from themselves or others, then the names of such students will need to be disclosed to the relevant school principal(s) to enable further action to be taken as may be appropriate"* (NSW Department of Education and Training, State Education Research Approval Process, Guidelines for Approving Applications from External Agencies to Conduct Research in NSW Government Schools, 2006, p. 14).

Will I be able to find out the results of the project?

Findings published in research reports will be presented in an aggregate format, without identifying either individuals or schools. The data will be kept in a locked file, accessible only to the university researchers in this study, although the unidentified data may be further analysed by other university researchers. The results of the study will be reported back to your school. Group findings from this study will be published in

Institute for Positive Psychology & Education

Australian Catholic University 25A Barker Road Strathfield NSW 2135 Australia
T: +61 (0)2 9701 4626 E: ippe@acu.edu.au W: ippe.acu.edu.au

**THE HELMSMAN
PROJECT**
Young people. Making a difference.

Australian Catholic University and the courses offered by the University are registered on the Commonwealth Register of Institutions and Courses for Overseas Students (CRICOS). Provider registration codes: 00004G, ABN: 15 050 192 660

(continues)

(THP Application and Consent Form continued)


Institute for Positive Psychology & Education


journals, presented at conferences and research symposia, in PhD theses and, provided to the participating schools in a report. However, all disseminated information will be in a non-identifiable format. That is, the results of the research will not identify you, your family or your child.

What if I have a complaint or any concerns?

The study has been reviewed by the Human Research Ethics Committee at Australian Catholic University (review number 2013-137Q). If you have any complaints or concerns about the conduct of the project, you may write to the Manager of the Human Research Ethics Committee care of the Office of the Deputy Vice Chancellor (Research). Any complaint or concern will be treated in confidence and fully investigated. You will be informed of the outcome.

Manager, Ethics
 c/o Office of the Deputy Vice Chancellor (Research)
 Australian Catholic University
 North Sydney Campus
 PO Box 968
 NORTH SYDNEY, NSW 2059
 Ph.: 02 9739 2519
 Fax: 02 9739 2870
 Email: resethics.manager@acu.edu.au

For any further information, please contact:

The Helmsman Project
 General Manager
 Kim Larochelle: kim.larochelle@thehelmsmanproject.org.au

I want to participate! How do I sign up?

Please complete the online application questions at the link provided in the Application Invitation below. You will also need to complete the consent forms and return a copy signed by your parent and/or legal guardian and return to your teacher.

Distinguished Professor Herb Marsh
 Centre for Positive Psychology in Education, Australian Catholic University
 25A Barker Road, Strathfield NSW 2135, Australia.
 email: herb.marsh@acu.edu.au

Institute for Positive Psychology & Education
 Australian Catholic University 25A Barker Road Strathfield NSW 2135 Australia
 T: +61 (0)2 9701 4626 E: ippe@acu.edu.au W: ippe.acu.edu.au

Australian Catholic University and the courses offered by the University are registered on the Commonwealth Register of Institutions and Courses for Overseas Students (CRICOS), Provider registration codes: 00004G, ABN: 15 050 192 660

**THE HELMSMAN
 PROJECT**
Young people. Making a difference.

(continues)

(THP Application and Consent Form continued)


Institute for Positive Psychology & Education

PARENT/GUARDIAN CONSENT FORM

PROJECT TITLE: The Helmsman Project: Giving at-risk adolescents skills to navigate life's journey and fulfil their potential.
APPLICATION NUMBER: 2014- 137Q
PRINCIPAL INVESTIGATOR: Distinguished Professor Herb Marsh
STUDENT RESEARCHER: Wendy Gwyn

I acknowledge that I have read the "The Helmsman Project: Research study to evaluate the efficacy of a combined outdoor education and coaching program designed to build the life effectiveness skills of adolescents – Participant Information Letter" and that I have discussed the project with my child. I give consent for my child to apply to participate in The Helmsman Project (THP) and the associated research and, if selected for her / him to participate in the program.

I have had the opportunity to raise any questions with respect to my child's application and possible participation in THP. To the extent that I have raised questions, these have been satisfactorily addressed.

I understand that my participation and my child's participation in this research program is voluntary. A decision not to participate will in no way affect their academic standing or relationship with the school and they are free to withdraw their participation at any time.

I am aware that, according to the Department of Education and Communities (DEC) procedures, researchers are required to inform the Principal if my child is identified as being at risk of harm from her / himself or others during participation in this research project and give consent for the researchers to notify the Principal of this.

Should you consent to participate in the research study, the researchers will be given access to your prior year academic results and days absent.

I agree that research data collected for the study may be published or may be provided to other researchers in a form that does not identify my child in any way.

PARENT/GUARDIAN NAME: (please print) _____

PARENT/GUARDIAN SIGNATURE: _____

NAME OF CHILD: (family name) _____ (first name) _____

SCHOOL: _____

ASSENT OF PARTICIPANTS AGED UNDER 18 YEARS

I _____ (*the participant aged under 18 years*) understand what this research project is designed to explore. What I will be asked to do has been explained to me. I agree to take part in THP realising that I can withdraw at any time without having to give a reason for my decision.

NAME OF PARTICIPANT AGED UNDER 18: _____

SIGNATURE: _____ DATE: _____

Institute for Positive Psychology & Education
 Australian Catholic University 25A Barker Road Strathfield NSW 2135 Australia
 T: +61 (0)2 9701 4626 E: ippe@acu.edu.au W: ippe.acu.edu.au

THE HELMSMAN PROJECT
 Young people. Making a difference.

Australian Catholic University and the courses offered by the University are registered on the Commonwealth Register of Institutions and Courses for Overseas Students (CRICOS), Provider registration codes: 00004G, ABN: 15 050 192 660

APPENDIX J RESEARCH INTERVIEW CONSENT FORM



Institute for Positive Psychology & Education



Past Participant Interview Consent

Further to my consent previously given for The Helmsman Project research, I agree to participate in a tape-recorded interview for a study about my experiences in The Helmsman Project Program.

I understand that I do not have to answer any questions I choose not to answer, and may withdraw from the interview at any time.

I understand that excerpts taken from this interview, written or spoken, will disguise all names of persons and places so as to preserve my anonymity and privacy.

I understand that I will not receive feedback on my interview.

We thank you for your generosity in making time available for our learning.

Date _____

Name of Participant _____

Signature of Participant _____

Institute for Positive Psychology & Education

Australian Catholic University 25A Barker Road Strathfield NSW 2135 Australia
T: +61 (0)2 9701 4626 E: ippe@acu.edu.au W: ippe.acu.edu.au

Australian Catholic University and the courses offered by the University are registered on the Commonwealth Register of Institutions and Courses for Overseas Students (CRICOS). Provider registration codes: 00004G, ABN: 15 050 192 660

**THE HELMSMAN
PROJECT**
Young people. Making a difference.

APPENDIX K

THP PROGRAM/SURVEY DATES AND COMMUNITY PROJECT DETAILS

Table K.1

THP Group Program and Survey Dates and Community Project Details

Cohort Number (waitlist control group in [])	School	Gender	Group/ Program	Wave	Program Launch and Graduation Dates	Survey Administration Dates	Notes	Community Project
1 [4]	A	F [M]	Intervention Arctos	T1 T2 T3	Launch: 22-Jul-13 Graduation: 6-Nov-13	22-Jul-13 07-Nov-13 04-Mar-14	Wk 12: 21-Oct-13	No Community Project Offered
2 [5]	B	F [M]	Intervention Arctos	T1 T2 T3	Launch: 5-Aug-13 Graduation: 6-Nov-13	29-Jul-13 07-Nov-13 14-Mar-14	Wk 12: 4-Nov-13	No Community Project Offered
3 [6]	C Girls only	F [F]	Intervention Arctos	T1 T2 T3	Launch: 29-Jul-13 Graduation: 6-Nov-13	29-Jul-13 07-Nov-13 11-Mar-14	Wk 12: 30-Oct-13	No Community Project Offered
4	A	M	Control Arctos	T1 T2 T3 T4 T5	Launch: 3-Mar-14 Graduation: 23-Jun-14	22-Jul-13 07-Nov-13 04-Mar-14 12-Jun-14 23-Oct-14	Wk 12: 2-Jun-14	No Community Project Offered
5	B	M	Control Arctos	T1 T2 T3 T4 T5	Launch: 10-Mar-14 Graduation: 23-Jun-14	29-Jul-13 07-Nov-13 14-Mar-14 26-Jun-14 13-Nov-14	Wk 12: 9-Jun-14	No Community Project Offered
6	C Girls only	F	Control Arctos	T1 T2 T3 T4 T5	Launch: 10-Mar-14 Graduation: 23-Jun-14	29-Jul-13 07-Nov-13 11-Mar-14 18-Jun-14 22-Oct-14	Wk 12: 9-Jun-14	No Community Project Offered
7 [9]	A	M [M]	Intervention Arctos	T1 T2 T3	Launch: 3-Feb-14 (wk of) Graduation: 23-Jun-14	05-Feb-14 27-May-14 04-Aug-14	Wk 12: 12-May-14	No Community Project Offered
8 [10]	A	F [F]	Intervention Arctos	T1 T2 T3	Launch: 3-Feb-14 (wk of) Graduation: 23-Jun-14	05-Feb-14 27-May-14 04-Aug-14	Wk 12: 12-May-14	No Community Project Offered
9	A	M	Control James Craig	T1 T2 T3 T4 T5	Launch: 14-Aug-14 Graduation: 1-Dec-14	05-Feb-14 27-May-14 04-Aug-14 27-Nov-14 13-Mar-15	Wk 12: 13-Nov-14	No Community Project Offered
10	A	F	Control Arctos	T1 T2 T3 T4 T5	Launch: 14-Aug-14 Graduation: 1-Dec-14	05-Feb-14 27-May-14 04-Aug-14 27-Nov-14 13-Mar-15	Wk 12: 20-Nov-14	No Community Project Offered
11-11 [16]	B	M [F]	Intervention James Craig	T1 T2 T3	Launch: 13-Aug-14 Graduation: 1-Dec-14	12-Aug-14 27-Nov-14 19-Feb-15	Wk 12: 14-Nov-14	No Community Project Offered
12-12 no control	B	M	Intervention James Craig	T1 T2 T3	Launch: 13-Aug-14 Graduation: 1-Dec-14	12-Aug-14 27-Nov-14 19-Feb-15	Wk 12: 14-Nov-14	No Community Project Offered
13 [17]	C Girls only	F [F]	Intervention Arctos	T1 T2 T3	Launch: 18-Aug-14 Graduation: 1-Dec-14	13-Aug-14 27-Nov-14 18-Feb-15	Wk 12: 26-Nov-14	No Community Project Offered
14 [18]	D Boys only	M [M]	Intervention James Craig	T1 T2 T3	Launch: 14-Aug-14 Graduation: 1-Dec-14	04-Aug-14 27-Nov-14 18-Feb-15	Wk 12: 14-Nov-14	No Community Project Offered
15 [19]	E	M [F]	Intervention James Craig	T1 T2 T3	Launch: 11-Aug-14 Graduation: 1-Dec-14	04-Aug-14 12-Dec-14 19-Feb-15	Wk 12: 10-Nov-14	No Community Project Offered
16	B	F	Control James Craig	T1 T2 T3 T4 T5	Launch: 25-Feb-15 Graduation: 16-Jun-15	12-Aug-14 27-Nov-14 19-Feb-15 09-Jun-15 26-Oct-15	Wk 12: 27-May-15	No Community Project Offered
17	C Girls only	F	Control James Craig	T1 T2 T3 T4 T5	Launch: 25-Feb-15 Graduation: 16-Jun-15	13-Aug-14 27-Nov-14 18-Feb-15 01-Jun-15 26-Oct-15	Wk 12: 27-May-15	No Community Project Offered
18	D Boys only	M	Control DNP	T1 T2 T3 T4 T5	DNP	04-Aug-14 27-Nov-14 18-Feb-15 DNP DNP	Control only. No subsequent	No Community Project Offered
19	E	F	Control James Craig	T1 T2 T3 T4 T5	Launch: 23-Feb-15 Graduation: 16-Jun-15	04-Aug-14 12-Dec-14 19-Feb-15 04-Jun-15 UTC	Wk 12: 25-May-15	No Community Project Offered

Note. DNP = did not proceed; UTC = unable to collect; DNI = data not included in research. Bold reflects data collection dates that were either less than two months or more than four months after the previous data collection date or pre-test data that was collected after program launch. Waitlist control group cohorts shaded gray for ease of reference.

(continues)

Table K.1 (continued)
 THP Group Program and Survey Dates and Community Project Details

Cohort Number (waitlist control group in [])	School	Gender	Group/ Program	Wave	Program Launch and Graduation Dates	Survey Administration Dates	Notes	Community Project
20 [22]	F	F [M]	Intervention James Craig	T1 T2 T3	Launch: 27-Feb-15 Graduation: 05-Dec-16	19-Feb-15 04-Jun-15 27-Aug-15	Wk 12: 29-May-15 (pitch date)	Completed (pilot group) Spooky science show at childrens' hospital
21 [23]	A	F [M]	Intervention James Craig	T1 T2 T3	Launch: 26-Feb-15 Graduation: 16-Jun-15	19-Feb-15 05-Jun-15 19-Aug-15	Wk 12: 28-May-15 (no pitch)	No Community Project Offered
22	F	M	Control James Craig	T1 T2 T3 T4 T5	Launch: 21-Aug-15 Graduation: Not yet	19-Feb-15 04-Jun-15 27-Aug-15 11-Dec-15 06-May-16	Wk 12: 27-Nov-15; Pitch: 4-Dec-15	*CP Pitch introduced after Wk 12 Not Yet Completed (still making efforts) Magic & comedy show for community center
23	A	M	Control James Craig	T1 T2 T3 T4 T5	Launch: 20-Aug-15 Graduation: No	19-Feb-15 05-Jun-15 19-Aug-15 11-Dec-15 06-May-16	Wk 12: 26-Nov-15; Pitch: 10-Dec-	Not Completed Disco to Welcome Year 7
24 [30]	G 7-10 only	F [M]	Intervention Arctos	T1 T2 T3	Launch: 11-Aug-15 Graduation: 05-Dec-16	10-Aug-15 09-Dec-15 23-Feb-16	Wk 12: 17-Nov-15; Pitch 24-Nov-15	Completed Purple week to raise money for cystic fibrosis
25 no control	C Girls only	F	Intervention Outward Bound	T1 T2 T3	Launch: 25-Aug-15 Graduation: 05-Dec-16	21-Aug-15 07-Dec-15 05-May-16	Wk 12: 2-Dec-15; Pitch 7-Dec-15	Completed Post New Year party for childrens' hospital
26 [31]	H	M [F]	Intervention James Craig	T1 T2 T3	Launch: 18-Aug-15 No Graduation	17-Aug-15 09-Dec-15 04-Feb-16	Wk 12: 24-Nov-15; Pitch: 1-Dec-15	Not Completed Wheelchair accessible path
27 [32]	I	F [M]	Intervention Outward Bound	T1 T2 T3	Launch: 24-Aug-15 Graduation: 05-Dec-16	20-Aug-15 11-Dec-15 08-Feb-16	Wk 12: 4-Dec-15; Pitch 11-Dec-15	Completed Restored nursing home garden
28			Not used					
29 [33]	B	M [F]	Intervention James Craig	T1 T2 T3	Launch: 17-Aug-15 No Graduation	17-Aug-15 07-Dec-15 10-Feb-16	Wk 12: 23-11-15; Pitch: 14-Dec-15	Not Completed Support charity that helps homeless
30	G 7-10 only	M	Control Outward Bound	T1 T2 T3 T4 T5	Launch: 01-Mar-16 Graduation: 05-Dec-16	10-Aug-15 09-Dec-15 23-Feb-16 29-Jun-16 23-Nov-16	Wk 12: 07-Jun-16; Pitch 14-Jun-16	Completed Film/Posters to help year 7 students transition to high school (anti-bullying)
31	H	F	Control Arctos	T1 T2 T3 T4 T5	Launch: 8-Feb-16 DNP	17-Aug-15 09-Dec-15 04-Feb-16 DNP DNP		Program Not Completed
32	I	M	Control Outward Bound	T1 T2 T3 T4 T5	Launch: 10-Feb-16 Graduation: 04-Dec-17	20-Aug-15 11-Dec-15 08-Feb-16 30-Jun-16 UTC	Wk 12: 31-May-16; Pitch 20-Jun-16	Completed Community garden project
33	B	F	Control Arctos	T1 T2 T3 T4 T5	Launch: 10-Feb-16 No Graduation	17-Aug-15 07-Dec-15 10-Feb-16 UTC 13-Oct-16	Wk 12: 20-May-16; Pitch 5-Jun-16	No Pitch
34 [38]	F	F [M]	Intervention Arctos	T1 T2 T3	Launch: 04-Feb-16 Graduation: 05-Dec-16	03-Feb-16 29-Jun-16 23-Sep-16	Wk 12: 13-May-16; Pitch 3-Jun-16	Completed Video for year 5 & 6 kids about discrimination and bullying
35 [39]	A	F [M]	Intervention Arctos	T1 T2 T3	Launch: 04-Feb-16 Graduation: 04-Dec-17	02-Feb-16 30-Jun-16 12-Aug-16	Wk 12: 12-May-16; Pitch 26-May-	Completed Dance/Presentation/Workshop at local public school about being yourself
36 [40]	J	M [F]	Intervention Coaching Only	T1 T2 T3	Launch: 08-Feb-16 No Graduation	05-Feb-16 01-Jul-16 30-Nov-16	Wk 12: 08-Jun-16; Pitch 21-Jun-16	Not Completed Cleaning the school and updating bathrooms
37 [41]	K	M [F]	Intervention Coaching Only	T1 T2 T3	Launch: 26-Feb-16 Graduation: Not yet	18-Mar-16 01-Jul-16 15-Aug-16	T1 collected after launch Wk 12: 02-Jun-16; Pitch 16-Jun-16	Not Yet Completed (still making efforts) Music, art, science workshops for school with special needs

Note. DNP = did not proceed; UTC = unable to collect; DNI = data not included in research. Bold reflects data collection dates that were either less than two months or more than four months after the previous data collection date or pre-test data that was collected after program launch. Waitlist control group cohorts shaded gray for ease of reference.

(continues)

Table K.1 (continued)
 THP Group Program and Survey Dates and Community Project Details

Cohort Number (waitlist control group in [])	School	Gender	Group/ Program	Wave	Program Launch and Graduation Dates	Survey Administration Dates	Notes	Community Project	
38	F	M	Control Arctos	T1 T2 T3 T4 T5	Launch: 19-Aug-16 Graduation: 05-Dec-16	03-Feb-16 29-Jun-16 12-Aug-16 02-Mar-17 27-Jun-17	Wk 12 (pitch): 25-Nov-16; Wk 13: 02-Dec-16	*CP pitch moved to Wk 12 (and new Wk 13 introduced) Not Yet Completed (still making efforts) Day to remember for aged care home	
39	A	M	Control Arctos	T1 T2 T3 T4 T5	Launch: 16-Aug-16 No Graduation	02-Feb-16 30-Jun-16 12-Aug-16 14-Dec-16 07-Aug-17	Wk 12 (pitch): 29-Nov-16; Wk 13: 06-Dec-16	Not Completed Create chill zone at school - safe space for all - break down racism	
40	J	F	Control Coaching Only	T1 T2 T3 T4 T5	Launch: 15-Aug-16 No Graduation	05-Feb-16 01-Jul-16 12-Aug-16 09-Dec-16 UTC	Wk 12 (pitch): 28-Nov-16; Wk 13: 30-Nov-16	Not Completed Fundraising for Starlight Foundation	
41	K	F	Control Coaching Only	T1 T2 T3 T4 T5	Launch: 18-Aug-16 Graduation: 04-Dec-17	18-Mar-16 01-Jul-16 15-Aug-16 14-Dec-16 UTC	Wk 12 (pitch): 30-Nov-16; Wk 13: 01-Dec-16	Completed Easter party at childcare facility	
42 [46]	F	M [F]	Intervention Outward Bound	T1 T2 T3	Launch: 9-Mar-17 Graduation: 04-Dec-17	02-Mar-17 27-Jun-17 01-Aug-17	Wk 12 (pitch): 08-Jun-17; Wk 13: 22-Jun-17 affected by control group launch	Completed Fundraise to create 50 care packages for local homeless shelter	
43 [47]	A	M [F]	Intervention Outward Bound	T1 T2 T3	Launch: 2-Mar-17 Graduation: 30-Nov-18	01-Mar-17 29-Jun-17 07-Aug-17	Wk 12 (pitch): 05-Jun-17; Wk 13: 07-Jun-17 affected by control group launch	Not Yet Completed (still making efforts) Introduce recycling at school	
44 [48]	G 7-10 only	F [M]	Intervention Coaching Only	T1 T2 T3	Launch: 28-Feb-17 Graduation: 04-Dec-17	21-Feb-17 29-Jun-17 22-Aug-17	Wk 12 (pitch): 06-Jun-17; Wk 13: 13-Jun-17 affected by control group launch	Completed Provide 100 care packages to homeless shelter and help cook for homeless	
45			Did Not Proceed						
46	F	F	Control Outward Bound	T1 T2 T3 T4 T5	Launch: 2-Aug-17 Graduation: 30-Nov-18	02-Mar-17 27-Jun-17 01-Aug-17 11-Dec-17 UTC	affected by need to launch Wk 12 (pitch): 29-Nov-17; Wk 13: 06-Dec-17	Not Yet Completed (still making efforts) Fundraise for childrens' hospital	
47	A	F	Control Outward Bound	T1 T2 T3 T4 T5	Launch: 7-Aug-17 Graduation: 30-Nov-18	01-Mar-17 29-Jun-17 07-Aug-17 30-Nov-17 UTC	affected by need to launch Wk 12 (pitch): 20-Nov-17; Wk 13: 27-Nov-17	Not Yet Completed (still making efforts) Make toys and do show at childrens' hospital	
48	G 7-10 only	M	Control Coaching Only	T1 T2 T3 T4 T5	Launch: 22-Aug-17 Graduation: 30-Nov-18	07-Mar-17 29-Jun-17 22-Aug-17 28-Nov-17 UTC	affected by need to launch Wk 12 (pitch): 21-Nov-17; Wk 13: 28-Nov-17	Completed Active day against childhood obesity for Year 7	
49 [50]	F	F [M]	Intervention Outward Bound	T1 T2 T3	Launch: 26-Feb-18 Graduation: Not yet	26-Feb-18 04-Jul-18 30-Jul-18	Wk 12 (pitch): 18-Jun-18; Wk 13: 25-Jun-18	Not Yet Completed (still making efforts) Reduce use of plastic water bottles and install bottle refill station	
50	F	M	Control Outward Bound	T1 T2 T3 T4 T5	Launch: 30-Jul-18 Graduation: 30-Nov-18	26-Feb-18 04-Jul-18 30-Jul-18 DNI DNI		Completed Improve school yard garden beds (laying tan bark, planting)	
51 [52]	A	F [M]	Intervention Outward Bound	T1 T2 T3	Launch: 23-Feb-18 Graduation: 30-Nov-18	23-Feb-18 03-Jul-18 30-Jul-18	Wk 12 (pitch): 18-Jun-18; Wk 13: 25-Jun-18	Completed Fundraise to make care packages for homeless	
52	A	M	Control Outward Bound	T1 T2 T3 T4 T5	Launch: 6-Aug-18 Graduation: 30-Nov-18	23-Feb-18 03-Jul-18 30-Jul-18 DNI DNI		Completed Improve schoolyard amenities (BBQ, garden beds, etc)	
53 [54]	K	F [M]	Intervention Coaching Only	T1 T2 T3	Launch: 09-Mar-18 Graduation: 30-Nov-18	09-Mar-18 UTC 10-Aug-18	Week 12: 22-Jun-18; Wk 13 (pitch): 29-Jun-18	Completed Make care packages for homeless youth and raise awareness	
54	K	M	Control Coaching Only	T1 T2 T3 T4 T5	Launch: 6-Aug-18 Graduation: 30-Nov-18	09-Mar-18 UTC 10-Aug-18 DNI DNI	Unclear of exact launch date and whether data collected before or after	Completed Xmas focused care packages for disadvantaged youth	
55 [56]	E	F [M]	Intervention Coaching Only	T1 T2 T3	Launch: 6-Mar-18 No Graduation	06-Mar-18 24-Jul-18 DNP	Data for T1 and T2 for 2 students Wk 12 (pitch): 19-Jun-18; Wk 13: 26-Jun-18	No Pitch	
56	E	M	Control No Program	T1 T2 T3 T4 T5	DNP	06-Mar-18 24-Jul-18 DNP DNP DNP	Data for T1 and T2 for 3-4 students only	DNP	

Note. DNP = did not proceed; UTC = unable to collect; DNI = data not included in research. Bold reflects data collection dates that were either less than two months or more than four months after the previous data collection date or pre-test data that was collected after program launch. Waitlist control group cohorts shaded gray for ease of reference.

APPENDIX L

LIST OF STATISTICAL R PACKAGES

The following statistical packages were used within R (in addition to the packages that are part of R) for data cleaning and analyses in this thesis, as indicated:

- `car` (Fox & Weisberg, 2011): recoding data;
- `dplyr` (Wickham, Francois, Henry, & Müller, 2017): data manipulation;
- `ggplot2` (Wickham, 2016): creating graphs;
- `magrittr` (Bache & Wickham, 2014): simplifying code;
- `MBESS` (Kelley, 2018): calculating omega.
- `mice` (van Buuren & Groothuis-Oudshoorn, 2011): missing data;
- `MplusAutomation` (Hallquist & Wiley, 2017): using data between R and Mplus;
- `psych` (Revelle, 2018): basic descriptive statistics, psychometric analysis, including reliability and scale structure, and correlation analysis;
- `semPlot` (Epskamp & Stuber, 2017): path diagrams and visual analysis of structural models;
- `semTools` (semTools Contributors, 2016): psychometric analysis of lavaan models, including reliability;
- `stringr` (Wickham, 2017): character manipulation; and
- `tidyr` (Wickham & Henry, 2017): data manipulation.

APPENDIX M

COEFFICIENT ALPHA FOR EACH SCALE

Table M.1
Coefficient Alpha for Each Scale using Long Form Data

Scale/Subscale	Alpha Coefficient	Scale/Subscale	Alpha Coefficient
CHS		SDQII-S	
Agency (3 items)	.57	<i>Non-Academic SC</i>	
Pathways Thinking (3 items)	.60	Physical Abilities SC ^a (4 items)	.88
Hope (6 items)	.70	Physical Appearance SC (4 items)	.90
LOT-R		Opp-Sex Rel'ships SC ^a (4 items)	.83
Optimism (3 items)	.53	Same-Sex Rel'ships SC ^a (5 items)	.83
Pessimism ^a (3 items)	.66	Parent Rel'ships SC ^a (4 items)	.86
LOT-R (6 items)	.60	Honesty-Trust SC ^a (6 items)	.79
ASRI		Emotional Stability SC ^a (5 items)	.78
Long-Term Self-Regulation ^a (14 items)	.80	<i>Academic SC</i>	
Short-Term Self-Regulation ^a (15 items)	.74	Math SC ^a (4 items)	.89
ASRI (29 items)	.86	Verbal SC ^a (5 items)	.89
ASRI-R		School SC ^a (4 items)	.79
Focus (5 items)	.77	<i>Global SC</i>	
Goal Self-Regulation (5 items)	.76	General Self-Esteem/SC ^a (6 items)	.81
Emotion Self-Regulation (5 items)	.63	ROPELOC	
Grit-S		<i>Personal Abilities and Beliefs</i>	
Consistency of Interest ^a (4 items)	.59	Self-Confidence (3 items)	.79
Perseverance of Interest (4 items)	.68	Self-Efficacy (3 items)	.77
Grit (8 items)	.64	Stress Management (3 items)	.77
MES-S		Open Thinking (3 items)	.71
Booster Thoughts (3 items)	.74	<i>Social Skills</i>	
Booster Behaviours (3 items)	.64	Social Effectiveness (3 items)	.82
Mufflers ^a (3 items)	.36	Cooperative Teamwork (3 items)	.79
Guzzlers ^a (2 items)	.53	Leadership Ability (3 items)	.88
ARS		<i>Organisational Skills</i>	
Academic Resilience (6 items)	.77	Time Efficiency (3 items)	.77
LRS		Quality Seeking (3 items)	.72
Life Resilience (6 items)	.78	Coping with Change (3 items)	.81
SWLS		<i>Energy</i>	
Satisfaction with Life (5 items)	.79	Active Involvement (3 items)	.65
WEMWBS		<i>Overall Effectiveness</i>	
Wellbeing (14 items)	.88	Overall Effectiveness (3 items)	.80
		<i>Locus of Control</i>	
		Internal LOC (3 items)	.71
		External LOC ^a (3 items)	.71

Note. CHS = Children's Hope Scale; LOT-R = Life Orientation Test, Revised; ASRI = Adolescent Self-Regulatory Inventory; ASRI-R = Adolescent Self-Regulatory Inventory, Revised; Grit-S = Short Grit Scale; MES-S = Motivation and Engagement Scale-Short; ARS = Academic Resilience Scale; LRS = Life Resilience Scale; SWLS = Satisfaction with Life Scale; WEMWBS = Warwick-Edinburgh Mental Well-Being Scale; SDQII-S = Self-Description Questionnaire II-Short; Opp = Opposite; Rel'ships = Relationships; Trust = Trustworthiness; ROPELOC = Review of Personal Effectiveness with Locus of Control; LOC = Locus of Control.

^aAll or some of the items for this scale are negatively worded and were reverse-scored prior to analysis.

APPENDIX N

SUPPLEMENTAL RESEARCH HYPOTHESIS ON GENDER INVARIANCE

Supplemental Research Hypothesis: Gender Invariance

The factor structure for each instrument will be similar for male and female participants, as demonstrated by tests of configural, metric, and scalar invariance across male and female gender groups.

Methodology for Gender Invariance Testing

Invariance across genders was tested for each measurement instrument using long form data grouped by gender. The best-fitting factor structure from the factor analysis in Chapter Five was initially modelled for each gender group simultaneously with no constraints (the configural model). Similar to the factor analysis, no indicator was fixed, but factor variance was fixed to one for model identification. This baseline model was then compared with a similar model with constrained factor loadings (the metric model), and then with a similar model with both constrained factor loadings and constrained item uniquenesses (the scalar model). These analyses were performed in Mplus with gender as a grouping variable and MODEL equal to “configural metric scalar,” which produces all three models at once. The complex option and MLR estimation were used for all analyses.

Results of Gender Invariance Testing

Multigroup invariance tests whether different groups respond to a measurement instrument in the same way (Vandenberg & Lance, 2000). Testing that this is the case is important to establish before making between group comparisons. For each measurement instrument, findings of an excellent fit for the configural model for that instrument demonstrate that the latent factors for the instrument have the same pattern of free and fixed loadings of items across both genders. If the metric model for an instrument does not result in a reduction in fit when compared to the configural model, this finding suggests that the latent variables being measured are the same in males and females for the data. If the scalar model for an instrument does not result in a reduction in fit when compared to the metric model, this finding indicates that both genders have the same

expected item response for the instrument. It is suggested that based on such results for a measurement instrument, it would be reasonable to make gender group comparisons of the relevant measurement scales.

Results of the gender invariance tests for each measurement instrument modelled with the preferred factor structure from the factor analysis, are set out in Table N.1. Following this table is Table N.2, which includes the chi square test statistics (χ^2), degrees of freedom (df), scaling correlation factor (SCF), and p -value for the Satorra-Bentler scaled chi square difference tests. Each configural model had excellent fit statistics other than the ARS which had fit statistics just below the acceptable threshold. Other than as detailed below and certain other minor exceptions, each metric and scalar model resulted in changes in fit statistics within the acceptable range.

The metric model for the LOT-R fit well, but resulted in changes in fit statistics that were outside of the acceptable range for the TLI and RMSEA and borderline for the CFI. However, the Satorra-Bentler scaled chi square difference test was not significant ($\chi^2_{diff(6)} = 12.021, p = .08$), indicating that the model fit did not get significantly worse with the metric model. Accordingly, it is arguable that the same latent variables of Optimism and Pessimism are being measured in both males and females. The scalar model also fit well and resulted in changes in fit statistics within the acceptable range, indicating that both genders have the same expected item response. As a consequence, it is arguably reasonable to make gender group comparisons of the LOT-R scales.

The scalar model for the Grit-S resulted in changes in fit statistics outside of the acceptable range. A partial scalar model was also tested by freeing the intercept for POE2 ("Setbacks don't discourage me"), consistent with modification indices. This item has been problematic in other psychometric research, including in connection with gender invariance testing (Wyszyńska, Ponikiewska, Karaś, Najderska, & Rogoza, 2017). However, the changes in fit statistics from the configural to the partial scalar model were also outside of the acceptable range. Further freeing the intercept for POE3 ("I am diligent") also consistent with modification indices, failed to bring the changes in fit statistics within the acceptable range.

Table N.1
Multigroup Invariance Models by Gender: Change in Fit Statistics

Invariance Models	Factor Structure	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA
CHS Configural (1)	Two-Factor CFA-M	.983		.963		.037	
CHS Metric (2)		.984	-.001	.974	-.011	.030	-.007
CHS Scalar (3)		.984	.000	.979	-.005	.028	-.002
LOT-R Configural (1)	Two-Factor CFA	1.000		1.005		.000	
LOT-R Metric (2)		.989	.011	.985	.020	.020	.020
LOT-R Scalar (3)		.981	.008	.978	.007	.025	.005
ASRI-R Configural (1)	Three-Factor ESEM	.973		.955		.033	
ASRI-R Metric (1)		.969	.004	.960	-.005	.031	-.002
ASRI-R Scalar (1)		.967	.002	.961	-.001	.031	.000
Grit-S Configural (1)	Two-Factor ESEM	.994		.986		.018	
Grit-S Metric (2)		.995	-.001	.993	-.007	.012	-.006
Grit-S Scalar (3a)		.938	.057	.921	.072	.043	.031
Grit-S Partial Scalar (3b)		.970	.025	.961	.032	.030	.018
Grit-S Partial Scalar (3c)		.980	.015	.973	.020	.025	.013
MES-S Configural (1)	Three-Factor ESEM	.951		.892		.055	
MES-S Metric (2)		.940	.011	.911	-.019	.050	-.005
MES-S Scalar (3)		.936	.004	.915	-.004	.049	-.001
ARS Configural (1)	One-Factor CFA	.884		.807		.104	
ARS Metric (2)		.889	-.005	.855	-.048	.091	-.013
ARS Scalar (3)		.886	.003	.878	-.023	.083	-.008
LRS Configural (1)	One-Factor CFA	.996		.994		.018	
LRS Metric (2)		.999	-.003	.999	-.005	.007	-.011
LRS Scalar (3)		.999	.000	.999	.000	.007	.000
SWLS Configural (1)	One-Factor CFA	1.000		1.001		.000	
SWLS Metric (2)		.997	.003	.996	.005	.018	.018
SWLS Scalar (3)		.996	.001	.996	.000	.019	.001
WEMWBS Configural (1)	One-Factor CFA	.945		.935		.043	
WEMWBS Metric (2)		.935	.010	.929	.006	.045	.002
WEMWBS Scalar (3)		.921	.014	.920	.009	.048	.003
SDQ-IIS Configural (1)	ESEM/CFA	.945		.916		.036	
SDQ-IIS Metric (2)		.936	.009	.920	-.004	.035	-.001
SDQ-IIS Scalar (3)		.931	.005	.916	.004	.036	.001
ROPELOC Configural (1)	14-Factor CFA	.937		.925		.034	
ROPELOC Metric (2)		.936	.001	.926	-.001	.034	.000
ROPELOC Scalar (3)		.934	.002	.924	.002	.034	.000

Note. CFI = comparative fit index; Δ CFI = decrease in CFI from the prior less constrained model (with a negative number reflecting an increase); TLI = Tucker-Lewis fit index; Δ TLI = decrease in TLI from the prior less constrained model (with a negative number reflecting an increase); RMSEA = root mean square error of approximation; Δ RMSEA = increase in RMSEA from the prior less constrained model (with a negative number reflecting a decrease); CHS = Children's Hope Scale; CFA = confirmatory factor analysis; M = modified; LOT-R = Life Orientation Test, Revised; ASRI-R = Adolescent Self-Regulatory Inventory, Revised; ESEM = exploratory structural equation model; Grit-S = Short Grit Scale; MES-S = Motivation and Engagement Scale-Short; ARS = Academic Resilience Scale; LRS = Life Resilience Scale; SWLS = Satisfaction with Life Scale; WEMWBS = Warwick-Edinburgh Mental Well-Being Scale; SDQII-S = Self-Description Questionnaire II-Short; ROPELOC = Review of Personal Effectiveness with Locus of Control. Grit-S Partial Scalar (3b) model: the intercept for POE₂ (being the second item on the Perseverance of Effort factor) was free. Grit-S Partial Scalar (3c) model: the intercept for POE₃ (being the third item on the Perseverance of Effort factor) was free in addition to POE₂.

Table N.2
Chi-Square Test Statistics for the Multigroup Gender Invariance Models for each Scale

Models	χ^2	df	<i>p</i> -value	SCF	Dif Test <i>p</i> -value
CHS Configural (1)	24.083	14	.045	1.2721	
CHS Metric (2)	28.361	19	.077	1.2959	.481
CHS Scalar (3)	32.468	23	.091	1.3137	.377
LOT-R Configural (1)	14.732	16	.544	1.4306	
LOT-R Metric (2)	26.753	22	.221	1.5185	.084
LOT-R Scalar (3)	34.497	26	.123	1.5162	.100
ASRI-R Configural (1)	199.526	126	.000	1.2788	
ASRI-R Metric (1)	245.241	162	.000	1.3493	.096
ASRI-R Scalar (1)	262.426	174	.000	1.3456	.144
Grit-S Configural (1)	30.354	26	.253	1.3748	
Grit-S Metric (2)	41.148	38	.334	1.3976	.537
Grit-S Scalar (3a)	86.923	44	.000	1.3778	.000
Grit-S Partial Scalar (3b)	63.825	43	.021	1.3827	.000
Grit-S Partial Scalar (3c)	55.848	42	.075	1.3858	.004
MES-S Configural (1)	131.513	50	.000	1.0316	
MES-S Metric (2)	173.308	74	.000	1.2685	.003
MES-S Scalar (3)	187.154	82	.000	1.2973	.066
ARS Configural (1)	122.782	18	.000	1.4103	
ARS Metric (2)	123.981	23	.000	1.4648	.406
ARS Scalar (3)	130.849	28	.000	1.4788	.173
LRS Configural (1)	21.228	18	.268	1.409	
LRS Metric (2)	23.556	23	.429	1.4219	.785
LRS Scalar (3)	28.636	28	.431	1.4003	.407
SWLS Configural (1)	9.558	10	.480	1.4231	
SWLS Metric (2)	16.537	14	.282	1.445	.143
SWLS Scalar (3)	21.438	18	.258	1.4414	.297
WEMWBS Configural (1)	306.866	154	.000	1.4776	
WEMWBS Metric (2)	346.877	167	.000	1.4599	.000
WEMWBS Scalar (3)	399.278	180	.000	1.4555	.000
SDQ-IIS Configural (1)	2795.320	1670	.000	1.1927	
SDQ-IIS Metric (2)	3363.717	2043	.000	1.2410	.000
SDQ-IIS Scalar (3)	3486.783	2083	.000	1.2421	.000
ROPELOC Configural (1)	2372.943	1456	.000	1.2529	
ROPELOC Metric (2)	2410.525	1484	.000	1.2559	.090
ROPELOC Scalar (3)	2470.492	1512	.000	1.2559	.000

Note. χ^2 = chi square test statistic; *df* = degrees of freedom; SCF = scaling correlation factor; and the Dif Test *p* value represents the *p*-value of the Satorra-Bentler scaled chi square difference test. For the Metric model, the Dif Test is between the Configural model and the Metric model, and for the Scalar model or Partial Scalar model, the Dif Test is between the Metric model and the Scalar or Partial Scalar model. CHS = Children's Hope Scale; LOT-R = Life Orientation Test, Revised; ASRI-R = Adolescent Self-Regulatory Inventory, Revised; Grit-S = Short Grit Scale; MES-S = Motivation and Engagement Scale-Short; ARS = Academic Resilience Scale; LRS = Life Resilience Scale; SWLS = Satisfaction with Life Scale; WEMWBS = Warwick-Edinburgh Mental Well-being Scale; SDQII-S = Self-Description Questionnaire II-Short; ROPELOC = Review of Personal Effectiveness with Locus of Control. Grit-S Partial Scalar (3b) model: the intercept for POE₂ (being the second item on the Perseverance of Effort factor) was free. Grit-S Partial Scalar (3c) model: the intercept for POE₃ (being the third item on the Perseverance of Effort factor) was free in addition to POE₂.

Conclusion

With some limited exceptions, the preferred factor structure from the factor analysis for each measurement instrument was found to be fully invariant across gender in support of the Supplemental Research Hypothesis. For the LOT-R, although changes in fit statistics between the configural and metric models were outside of the acceptable range, the scalar model was acceptable. Moreover, as metric invariance was supported by a non-significant Satorra-Bentler scaled chi square difference test, the LOT-R factor structure was found to be sufficiently invariant across gender to support the Supplemental Research Hypothesis. For the Grit-S, the fit statistics for the configural model were good and the changes in fit statistics from the configural to the metric model were within range, but the degradation in fit statistics from the metric model to the scalar model were outside of the threshold range. While partial scalar invariance was tested by freeing two of the five Perseverance of Effort items, these adjustments did not achieve partial scalar invariance. Accordingly, the Supplemental Research Hypothesis was not supported for the Grit-S, suggesting that mean differences in the Grit-S data across gender cannot be clearly attributed to differences in the construct of grit.

APPENDIX O
FACTOR ANALYTIC MODELS: CHI-SQUARE TEST STATISTICS

Table O.1
Chi-Square Test Statistics for the Factor Analytic Models for each Scale

Models	χ^2	<i>df</i>	<i>p</i>-value	SCF
CHS				
One-factor CFA	49.159	9	.000	1.4195
Two-factor CFA	18.275	8	.019	1.3868
Two-factor ESEM	4.453	4	.348	0.7497
Two-factor CFA, modified				
LOT-R				
One-factor CFA	133.605	9	.000	1.3914
Two-factor CFA	8.691	8	.369	1.5212
Bifactor ("faking good")	3.138	6	.791	1.5430
Two-factor ESEM	2.735	4	.603	1.4843
ASRI				
One-factor CFA	1813.026	377	.000	1.3738
Two-factor CFA	1782.262	376	.000	1.3712
Two-factor ESEM	1058.382	349	.000	1.3591
ASRI-R				
Three-factor CFA	165.291	87	.000	1.4863
Three-factor ESEM	98.851	63	.003	1.3901
Grit-S				
One-factor CFA	262.191	20	.000	1.4147
Two-factor CFA	43.368	19	.001	1.4509
Two-factor ESEM	23.909	13	.032	1.4606
MES-S				
Four-factor CFA	207.331	38	.000	1.3105
Two-factor EFA	164.697	34	.000	1.2738
Three-factor EFA	69.340	25	.000	1.2499
Three-factor ESEM	69.339	25	.000	1.2499
Unidimensional Measures				
ARS one-factor CFA	96.352	9	.000	1.5149
LRS one-factor CFA	10.960	9	.279	1.4619
SWLS one-factor CFA	3.986	5	.552	1.5308
WEMWBS one-factor CFA	212.499	77	.000	1.5103
SDQII-S				
11-factor CFA + CU	2996.878	1168	.000	1.3760
11 factor ESEM + CU	1201.094	768	.000	1.2809
10-factor ESEM/CFA (Sch) + CU	1692.989	835	.000	1.2946
ROPELOC				
14-factor CFA	1369.000	728	.000	1.3428
14-factor ESEM	680.951	364	.000	0.9759

Note. χ^2 = chi square test statistic; *df* = degrees of freedom; SCF = scaling correlation factor; CHS = Children's Hope Scale; CFA = confirmatory factor analysis; ESEM = exploratory structural equation model; LOT-R = Life Orientation Test, Revised; Bifactor ("faking good") = a bifactor model with a general factor for all LOT-R items and a specific factor for the positively-worded items; ASRI = Adolescent Self-Regulatory Inventory; ASRI-R = Adolescent Self-Regulatory Inventory, Revised; Grit-S = Short Grit Scale; MES-S = Motivation and Engagement Scale-Short; ARS = Academic Resilience Scale; LRS = Life Resilience Scale; SWLS = Satisfaction with Life Scale; WEMWBS = Warwick-Edinburgh Mental Well-Being Scale; SDQII-S = Self-Description Questionnaire II-Short; ROPELOC = Review of Personal Effectiveness with Locus of Control.

APPENDIX P

LONGITUDINAL INVARIANCE MODELS: CHI-SQUARE TEST STATISTICS

Table P.1
Chi-Square Test Statistics for the Longitudinal Invariance Models for each Scale

Models	χ^2	df	p-value	SCF	Dif Test p-value
CHS Configural (1)	44.009	40	.306	1.0819	
CHS Metric (2)	53.522	45	.180	1.0692	.077
CHS Scalar (3a)	72.132	49	.017	1.0544	.000
CHS Partial Scalar (3b)	58.358	48	.145	1.0611	.172
LOT-R Configural (1)	44.721	42	.358	1.1993	
LOT-R Metric (2)	45.921	46	.476	1.2095	.836
LOT-R Scalar (3)	48.479	47	.413	1.1968	.044
ASRI-R Configural (1)	374.226	327	.037	1.1188	
ASRI-R Metric (2)	411.519	363	.040	1.1566	.370
ASRI-R Scalar (3)	425.038	374	.035	1.1517	.251
Grit-S Configural (1)	117.317	78	.003	1.1537	
Grit-S Metric (2)	134.92	90	.002	1.1854	.126
Grit-S Scalar (3)	146.333	96	.001	1.1725	.064
MES-S Configural (1)	203.992	151	.003	1.0567	
MES-S Metric (2)	213.998	175	.024	1.1145	.902
MES-S Scalar (3)	233.723	183	.007	1.0980	.002
ARS Configural (1)	113.687	47	.000	1.2626	
ARS Metric (2)	124.706	52	.000	1.2717	.050
ARS Scalar (3)	130.391	57	.000	1.2487	.522
LRS Configural (1)	50.927	47	.322	1.2686	
LRS Metric (2)	58.926	52	.237	1.2506	.135
LRS Scalar (3)	66.766	57	.177	1.2276	.137
SWLS Configural (1)	29.601	29	.434	1.1919	
SWLS Metric (2)	33.516	33	.442	1.2168	.415
SWLS Scalar (3a)	46.894	37	.128	1.1857	.489
SWLS Partial Scalar (3b)	37.083	36	.419	1.1965	.297
WEMWBS Configural (1)	490.301	335	.000	1.2267	
WEMWBS Metric (2)	513.848	348	.000	1.2196	.028
WEMWBS Scalar (3)	540.742	361	.000	1.2117	.008
SDQII-S Configural A (1)	6428.24	4097	.000	1.0102	
SDQII-S Metric A (2)	6768.634	4470	.000	1.0303	.341
SDQII-S Scalar A (3)	---	Did	Not	Converge	---
SDQII-S Configural B (1)	2628.671	1670	.000	1.0807	
SDQII-S Metric B (2)	2993.036	2043	.000	1.0951	.436
SDQII-S Scalar B (3)	3057.402	2083	.000	1.0897	.006
ROPELOC Configural A (1)	4359.683	2982	.000	1.0573	
ROPELOC Metric A (2)	4384.177	3010	.000	1.0587	.544
ROPELOC Scalar A (3)	4449.25	3037	.000	1.0580	.000
ROPELOC Configural B (1)	2123.369	1456	.000	1.1405	
ROPELOC Metric B (2)	2150.099	1484	.000	1.1412	.509
ROPELOC Scalar B (3)	2200.252	1512	.000	1.1366	.003

Note. χ^2 = chi square test statistic; df = degrees of freedom; SCF = scaling correlation factor; and the Dif Test p value represents the p-value of the Satorra-Bentler scaled chi square difference test. For the Metric model, the Dif Test is between the Configural model and the Metric model, and for the Scalar model or Partial Scalar model, the Dif Test is between the Metric model and the Scalar or Partial Scalar model. CHS = Children's Hope Scale; LOT-R = Life Orientation Test, Revised; ASRI-R = Adolescent Self-Regulatory Inventory, Revised; Grit-S = Short Grit Scale; MES-S = Motivation and Engagement Scale-Short; ARS = Academic Resilience Scale; LRS = Life Resilience Scale; SWLS = Satisfaction with Life Scale; WEMWBS = Warwick-Edinburgh Mental Well-being Scale; SDQII-S = Self-Description Questionnaire II-Short; ROPELOC = Review of Personal Effectiveness with Locus of Control. For SDQII-S and ROPELOC, A refers to the standard longitudinal invariance models with wide data used for this analysis and B refers to the alternative multigroup models with time as the grouping variable. Number of observations = 350 except for the B models, where the number of observations for T₁ (pre-test) = 319 and for T₂ (immediate post-test) = 292.

APPENDIX Q

FACTOR LOADINGS AND CORRELATIONS FOR SDQII-S ESEM/CFA

Table Q.1
Standardised Factor Loadings for the SDQII-S ESEM/CFA

SDQII-S Items	Factor Loadings										
	Phys Abil	Phys App	Op Sex Rel	Sm Sex Rel	Par Rel	Hon/Trust	Emot Stab	Math	Verbal	Global SE	School
Pab1	.88*	-.05*	-.02	.01	.05*	-.01	-.06*	-.02	.05*	-.01	
Pab2	.95*	.00	.00	-.01	.01	-.01	-.01	.00	.03	-.01	
Pab3 ^a	.72*	-.03	.02	.10*	-.04	.07*	.09*	-.03	-.11*	-.02	
Pab4	.69*	.12*	.01	-.11*	-.06	-.01	.01	.07*	.00	.05	
Pap1	-.01	.89*	-.06*	.06	.01	.02	.00	.05*	.04*	-.03	
Pap2	-.05*	.90*	.00	.03	-.01	.00	.03	-.01	.00	.05*	
Pap3	.02	.82*	.05	.00	-.05*	.06*	-.02	-.01	-.02	-.04	
Pap4	.13*	.62*	.07	-.06	.08*	-.10*	.01	.00	-.02	.09*	
OS1 ^a	.00	.10*	-.54*	.15*	-.05	.07*	.06	-.04	.03	-.11*	
OS2	.00	-.01	.90*	-.08*	.01	-.04	.02	.03	.02	.06*	
OS3 ^a	.06	-.06	.56*	.17*	-.03	.11*	.10*	-.06	.01	-.07	
OS4	-.02	.00	.86*	-.03	.03	-.01	-.04	.02	-.02	.02	
SS1 ^a	.01	-.04	.02	.75*	.00	-.02	-.01	-.01	.01	.01	
SS2	.03	.06	.09	.55*	.07	-.08*	-.06	.04	.01	.13*	
SS3 ^a	-.03	-.03	-.03	.69*	.00	-.01	.01	.01	.03	.04	
SS4	.06	.09*	.18*	.51*	.10*	-.08*	-.06	.04	.05	.06	
SS5	.02	.07	-.01	.58*	-.01	.07*	.09*	-.01	.03	.04	
PR1	-.03	.02	.01	.02	.85*	.02	.00	-.01	.02	-.03	
PR2	.03	-.04	.02	-.04	.80*	-.03	.01	.05*	.04	.04	
PR3	-.01	.00	.00	-.09*	.72*	.02	.06*	-.03	-.03	.15*	
PR4 ^a	.00	.03	-.05	.17*	.69*	.08*	-.05	-.03	-.01	-.11*	
Ho1	.02	.04	.07	-.02	.04	.54*	-.13*	-.10*	-.06	.22*	
Ho2 ^a	-.04	.05	.04	.01	-.05	.76*	.09*	.00	-.03	.04	
Ho3 ^a	.05	-.07	-.06	.00	-.02	.54*	.03	.11*	.07	.04	
Ho4 ^a	.00	-.15*	-.03	.10	.02	.46*	.03	.01	.02	.05	
Ho5 ^a	.05	-.04	.03	-.07*	.06	.69*	.01	.06*	.03	-.16*	
Ho6	-.01	.12*	.03	-.13*	.06	.71*	-.10*	-.02	.03	.02	
ES1 ^a	.01	-.02	.02	-.07	-.06	-.07*	.76*	-.04	-.02	.08	
ES2 ^a	.05	.11*	.03	.04	-.08	.03	.60*	-.02	.10*	-.05	
ES3 ^a	-.03	-.04	.00	.08*	.11*	.08*	.48*	.05	.00	.07	
ES4 ^a	.04	.06	.01	.11*	.05	.04	.41*	-.04	-.06	.03	
ES5 ^a	-.02	-.03	.04	-.12*	.06*	-.05	.89*	.05*	.01	.00	
Mh1	-.02	.03	.01	-.01	.01	.01	-.05*	.89*	-.02	-.05	
Mh2	.01	-.02	.01	-.01	.01	-.01	.01	.89*	.01	.06*	
Mh3	.01	.04	.01	.00	-.01	.00	.01	.89*	.01	.00	
Mh4 ^a	.01	-.04	-.06	.06	-.03	.06	.05	.57*	-.05	.03	
V1 ^a	.01	-.06	-.04	.16*	.01	.13*	.08*	-.01	.65*	-.08	
V2	.00	.03	.01	.00	-.02	-.05*	.00	-.06*	.79*	.08*	
V3	-.03	.04	.00	-.04	.02	.00	-.03	-.06*	.81*	-.12*	
V4	.00	-.02	.02	-.02	.02	-.02	.01	.07*	.86*	.03	
V5	.02	-.01	.01	-.05	-.02	-.01	-.03	.00	.83*	.12*	
GS1	.11*	.09*	-.04	.03	.20*	-.02	.07	-.01	.00	.50*	
GS2	.00	.02	.03	-.03	-.02	.09*	-.05	.00	.04	.76*	
GS3	.00	.01	-.05	.01	-.01	.02	.05	-.04	-.01	.63*	
GS4	.00	-.01	.06	.07*	-.04	-.01	.01	.15*	.04	.60*	
GS5	.02	.00	.01	.09*	-.03	.04	-.05	.00	.07	.55*	
GS6 ^a	-.03	.09*	-.06	.20*	.14*	.14*	.27*	.03	.02	.28*	
Sch1 ^a											.58*
Sch2											.72*
Sch3											.75*
Sch4											.78*

Note. SDQII-S = Self-Description Questionnaire II-Short. ESEM = exploratory structural equation model; CFA = confirmatory factor analysis; Phy Abil or Pab = Physical Ability Self-Concept; Phys App or Pap = Physical Appearance Self-Concept; Op Sex Rel or OS = Opposite Sex Relationships Self-Concept; Sm Sex Rel or SS = Same Sex Relationships Self-Concept; Par Rel or PR = Parent Relationships Self-Concept; Hon/Trust or Ho = Honesty/Trustworthiness Self-Concept; Emot Stab or ES = Emotional Stability Self-Concept; Math or Mh = Math Self-Concept; Verbal or V = Verbal Self-Concept; Global SE or GS = Global Self-Esteem/Self-Concept; School or Sch = School Self-Concept. Target loadings highlighted grey.

*indicates significant p -value < .05.

^a This item is negatively worded and was reverse-scored prior to analysis.

Table Q.2
Factor Correlations for the SDQII-S ESEM/CFA

Factor Correlations											
	Phys Abil	Phys App	Op Sex Rel	Sm Sex Rel	Par Rel	Honesty	Emot Stab	Math	Verbal	Global SE	School
Phys Abil	1.00*										
Phys App	.36*	1.00*									
Op-Sex Rel	.39*	.47*	1.00*								
Sm-Sex Rel	.16*	.18*	.28*	1.00*							
Parent Rel	.14*	.20*	.01	.31*	1.00*						
Honesty	.14*	.06	.09	.25*	.34*	1.00*					
Emot Stab	.11*	.14*	.12*	.33*	.24*	.26*	1.00*				
Math	.05	.10*	.00	.07	.04	.11*	.09	1.00*			
Verbal	-.02	.14*	.21*	.25*	.17*	.23*	.14*	-.03	1.00*		
Global SE	.29*	.36*	.17*	.28*	.44*	.23*	.22*	.31*	.29*	1.00*	
School	.12*	.23*	.12*	.30*	.26*	.29*	.17*	.57*	.53*	.75*	1.00*

Note. SDQII-S = Self-Description Questionnaire II-Short. ESEM = exploratory structural equation model; CFA = confirmatory factor analysis; Phys Abil = Physical Ability Self-Concept; Phys App = Physical Appearance Self-Concept; Op Sex Rel = Opposite Sex Relationships Self-Concept; Sm Sex Rel = Same Sex Relationships Self-Concept; Par Rel = Parent Relationships Self-Concept; Hon/Trust = Honesty/Trustworthiness Self-Concept; Emot Stab = Emotional Stability Self-Concept; Math = Math Self-Concept; Verbal = Verbal Self-Concept; Global SE = Global Self-Esteem/Self-Concept; School = School Self-Concept.

*indicates significant p -value < .05.

APPENDIX R

ROPELOC FACTOR LOADINGS AND CORRELATIONS FOR CFA

Table R.1
Standardised Factor Loadings and Correlations for ROPELOC CFA

ROPELOC Items	Factor Loadings													
	SC	SF	SM	OT	SE	CT	LA	TE	QS	CH	AI	OE	IL	EL
SC1	.70*													
SC2	.76*													
SC3	.79*													
SF1		.65*												
SF2		.76*												
SF3		.78*												
SM1			.65*											
SM2			.75*											
SM3			.79*											
OT1				.58*										
OT2				.65*										
OT3				.75*										
SE1					.70*									
SE2					.86*									
SE3					.78*									
CT1						.64*								
CT2						.79*								
CT3						.80*								
LA1							.83*							
LA2							.88*							
LA3							.81*							
TE1								.65*						
TE2								.76*						
TE3								.79*						
QS1									.57*					
QS2									.77*					
QS3									.73*					
CH1										.68*				
CH2										.83*				
CH3										.82*				
AI1											.64*			
AI2											.56*			
AI3											.65*			
OE1												.66*		
OE2												.83*		
OE3												.78*		
IL1													.60*	
IL2													.68*	
IL3													.74*	
EL1 ^a														.65*
EL2 ^a														.67*
EL3 ^a														.71*
Factor Correlations														
SC	1.00*													
SF	.70*	1.00*												
SM	.58*	.81*	1.00*											
OT	.73*	.68*	.62*	1.00*										
SE	.60*	.44*	.42*	.49*	1.00*									
CT	.53*	.42*	.44*	.58*	.58*	1.00*								
LA	.58*	.38*	.27*	.52*	.64*	.56*	1.00*							
TE	.63*	.57*	.53*	.54*	.37*	.39*	.38*	1.00*						
QS	.77*	.53*	.43*	.68*	.38*	.50*	.34*	.64*	1.00*					
CH	.65*	.77*	.78*	.71*	.52*	.50*	.38*	.59*	.56*	1.00*				
AI	.67*	.48*	.44*	.74*	.61*	.70*	.62*	.47*	.59*	.57*	1.00*			
OE	.83*	.68*	.56*	.70*	.54*	.47*	.53*	.64*	.71*	.68*	.59*	1.00*		
IL	.71*	.48*	.40*	.77*	.39*	.48*	.39*	.44*	.76*	.50*	.60*	.62*	1.00*	
EL	.17*	.02	.04	.09	.06	.12*	.12*	.09	.13*	.08	.07	.14*	.24*	1.00*

Note. ROPELOC = Review of Personal Effectiveness and Locus of Control; CFA = confirmatory factor analysis; SC = Self-confidence; SF = Self-Efficacy; SM = Stress Management; OT = Open Thinking; SE = Social Effectiveness; CT = Cooperative Teamwork; LA = Leadership Ability; TE = Time Efficiency; QS = Quality Seeking; CH = Coping with Change; AI = Active Involvement; OE = Overall Effectiveness; IL = Internal Locus of Control; EL = External Locus of Control.

*indicates significant p -value < .05.

^a This item is negatively worded and was reverse-scored prior to analysis.

APPENDIX S DESCRIPTIVE STATISTICS

Table S.1
Descriptive Statistics: Adventure Programs (taken together) (T1-T3)

Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE
Children's Hope Scale															
Agency	T1	153	3.80	0.61	-0.36	-0.22	0.05	Same-Sex Relat'ships SC ^a	T1	152	3.72	0.81	-0.63	0.62	0.07
	T2	128	3.97	0.59	-0.77	1.16	0.05		T2	128	3.71	0.72	0.00	-0.72	0.06
	T3	120	3.90	0.64	-0.62	0.33	0.06		T3	120	3.80	0.78	-0.48	0.00	0.07
Pathways (revised)	T1	153	3.82	0.58	-0.05	-0.56	0.05	Parent Relat'ships SC ^a	T1	153	3.89	0.87	-0.69	0.05	0.07
	T2	128	3.87	0.56	-0.17	0.23	0.05		T2	128	3.81	0.91	-0.59	-0.10	0.08
	T3	120	3.86	0.59	-0.19	-0.30	0.05		T3	120	3.86	0.86	-0.92	1.45	0.08
Life Orientation Test, Revised															
Optimism	T1	152	3.55	0.59	0.10	-0.21	0.05	Honesty-Trust SC ^a	T1	152	3.33	0.77	0.03	-0.44	0.06
	T2	128	3.62	0.60	0.36	-0.22	0.05		T2	128	3.38	0.78	-0.20	-0.37	0.07
	T3	120	3.52	0.54	0.24	0.07	0.05		T3	120	3.55	0.71	0.05	-0.55	0.06
Pessimism ^a	T1	152	2.63	0.68	-0.15	-0.36	0.06	Emotional Stability SC ^a	T1	153	2.72	0.75	0.03	-0.21	0.06
	T2	128	2.73	0.78	-0.02	0.19	0.07		T2	128	2.81	0.75	0.06	0.04	0.07
	T3	119	2.86	0.75	0.04	-0.52	0.07		T3	120	2.83	0.74	0.21	0.10	0.07
Adolescent Self-Regulatory Inventory, Revised															
Focus	T1	153	3.17	0.80	-0.43	0.08	0.06	Math SC ^a	T1	153	3.10	1.07	0.11	-0.75	0.09
	T2	128	3.27	0.71	-0.07	-0.15	0.06		T2	128	3.10	1.04	0.03	-0.69	0.09
	T3	120	3.31	0.70	-0.52	-0.07	0.06		T3	120	3.07	1.15	-0.01	-1.08	0.11
Goal Self-Regulation	T1	153	3.74	0.60	-0.24	-0.65	0.05	Verbal SC ^a	T1	152	3.22	0.96	-0.31	-0.17	0.08
	T2	128	3.82	0.60	-0.21	0.01	0.05		T2	128	3.43	0.82	0.20	-0.69	0.07
	T3	120	3.79	0.56	-0.04	-0.05	0.05		T3	120	3.47	0.91	-0.21	-0.40	0.08
Emotion Self-Reg	T1	153	3.45	0.58	-0.49	0.47	0.05	School SC ^a	T1	152	3.66	0.73	-0.51	0.30	0.06
	T2	128	3.48	0.64	0.09	0.11	0.06		T2	128	3.63	0.74	-0.48	0.38	0.07
	T3	120	3.52	0.59	-0.29	0.74	0.05		T3	120	3.69	0.80	-0.22	-0.10	0.07
Short Grit Scale															
Consistency of Interest ^a	T1	152	2.66	0.70	0.54	1.27	0.06	General Self-Esteem ^a	T1	153	3.74	0.57	-0.70	0.71	0.05
	T2	128	2.60	0.67	0.43	0.70	0.06		T2	128	3.82	0.65	-0.56	0.26	0.06
	T3	120	2.70	0.60	0.49	0.87	0.06		T3	120	3.83	0.64	-0.10	-0.81	0.06
Perseverance of Effort	T1	153	3.60	0.60	-0.05	-0.04	0.05	Review of Personal Effectiveness and Locus of Control							
	T2	128	3.60	0.63	0.09	-0.04	0.06	Self-Confidence	T1	152	3.79	0.70	-0.60	0.25	0.06
	T3	120	3.67	0.60	0.05	-0.30	0.06		T2	128	3.87	0.70	-0.26	-0.33	0.06
									T3	120	3.86	0.73	-0.63	1.21	0.07
								Self-Efficacy	T1	153	3.45	0.68	0.00	-0.42	0.05
									T2	128	3.66	0.73	-0.48	0.28	0.06
									T3	120	3.63	0.64	-0.43	0.48	0.06
Motivation and Engagement Scale - Short															
Booster Thoughts	T1	152	3.81	0.60	-0.12	-0.65	0.05	Stress Management	T1	152	3.41	0.75	-0.15	-0.20	0.06
	T2	128	3.79	0.62	0.08	-0.62	0.05		T2	128	3.53	0.75	-0.06	-0.30	0.07
	T3	120	3.87	0.60	-0.05	-0.81	0.05		T3	120	3.51	0.75	-0.39	0.36	0.07
Booster Behaviours	T1	153	3.29	0.78	-0.38	-0.01	0.06	Open Thinking	T1	153	3.91	0.61	-0.75	1.92	0.05
	T2	128	3.44	0.76	-0.12	0.11	0.07		T2	128	3.99	0.61	-0.33	-0.09	0.05
	T3	120	3.44	0.76	0.09	-0.33	0.07		T3	120	4.03	0.54	-0.17	-0.22	0.05
Hampering ^a	T1	153	2.79	0.63	0.25	0.86	0.05	Social Effectiveness	T1	153	3.59	0.69	-0.28	-0.14	0.06
	T2	128	2.89	0.61	-0.27	0.12	0.05		T2	128	3.67	0.74	-0.19	-0.16	0.06
	T3	120	2.92	0.62	0.04	0.05	0.06		T3	120	3.73	0.74	-0.23	-0.31	0.07
Life Resilience Scale															
Life Resilience	T1	153	3.46	0.60	-0.26	0.27	0.05	Cooperative Teamwork	T1	153	4.01	0.76	-0.95	1.70	0.06
	T2	128	3.61	0.61	-0.37	0.26	0.05		T2	128	4.08	0.60	-0.14	-0.68	0.05
	T3	120	3.67	0.60	-0.31	0.10	0.05		T3	120	4.05	0.62	-0.40	-0.13	0.06
Academic Resilience Scale															
Academic Resilience	T1	153	3.42	0.62	-0.22	0.74	0.05	Leadership Ability	T1	153	3.69	0.89	-0.62	0.19	0.07
	T2	158	3.49	0.64	-0.49	0.63	0.05		T2	128	3.77	0.90	-0.43	-0.50	0.08
	T3	157	3.53	0.63	-0.24	0.32	0.05		T3	120	3.76	0.91	-0.62	-0.04	0.08
Satisfaction with Life Scale															
Life Satisfaction	T1	153	3.32	0.74	-0.26	0.18	0.06	Time Efficiency	T1	152	3.21	0.73	-0.39	0.44	0.06
	T2	128	3.40	0.68	-0.25	0.16	0.06		T2	128	3.26	0.77	-0.11	0.24	0.07
	T3	120	3.48	0.74	-0.32	0.05	0.07		T3	120	3.34	0.72	-0.52	0.34	0.07
Warwick-Edinburgh Mental Well-Being Scale															
Wellbeing	T1	153	3.56	0.50	-0.31	-0.13	0.04	Quality Seeking	T1	152	4.08	0.60	-0.40	-0.01	0.05
	T2	128	3.72	0.58	-0.28	0.35	0.05		T2	128	4.07	0.57	-0.23	-0.07	0.05
	T3	120	3.64	0.56	-0.12	-0.19	0.05		T3	120	3.96	0.57	-0.18	0.16	0.05
Self-Description Questionnaire II-Short															
Physical Abilities SC ^a	T1	153	3.49	1.11	-0.61	-0.49	0.09	Coping with Change	T1	152	3.50	0.67	-0.32	-0.29	0.05
	T2	128	3.59	1.03	-0.64	-0.29	0.09		T2	128	3.59	0.74	-0.43	0.52	0.07
	T3	120	3.50	1.03	-0.48	-0.28	0.09		T3	120	3.60	0.72	-0.17	-0.24	0.07
Physical Appearance SC	T1	153	2.88	0.89	-0.11	-0.13	0.07	Active Involvement	T1	153	3.98	0.61	-0.63	0.13	0.05
	T2	128	3.00	0.87	-0.16	-0.04	0.08		T2	128	4.02	0.62	-0.26	-0.47	0.06
	T3	119	3.08	0.89	-0.01	-0.01	0.08		T3	120	3.94	0.62	-0.45	-0.02	0.06
Opp-Sex Relat'ships SC ^a	T1	152	3.31	0.96	-0.23	-0.49	0.08	Overall Effectiveness	T1	153	3.51	0.76	-0.10	0.09	0.06
	T2	128	3.46	0.95	-0.35	-0.28	0.08		T2	128	3.60	0.69	-0.16	0.74	0.06
	T3	120	3.58	0.84	-0.28	-0.17	0.08		T3	120	3.59	0.67	-0.29	0.99	0.06
Internal LOC															
Internal LOC	T1	153	4.16	0.56	-0.21	-0.66	0.04	External LOC ^a	T1	153	3.35	0.86	-0.31	-0.03	0.07
	T2	128	4.18	0.65	-0.46	-0.25	0.06		T2	128	3.25	0.75	-0.28	0.45	0.07
	T3	120	4.16	0.59	-0.30	-0.49	0.05		T3	120	3.52	0.71	-0.33	0.55	0.06

Note. THP = The Helmsman Project; SD = standard deviation; SE = standard error; T1 = pre-test; T2 = immediate post-test; T3 = three months post-test; Self-Reg = Self-Regulation; SC = Self-Concept; Relat'ships = Relationships; Trust = Trustworthiness; LOC = Locus of Control.

^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

Table S.2
Descriptive Statistics: Arctos Adventure Program (T1-T3)

Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE
Children's Hope Scale								Self-Description Questionnaire II-Short							
Agency	T1	57	3.73	0.60	0.07	-0.50	0.08	Same-Sex Relationships SC ^a	T1	56	3.87	0.80	-0.76	1.07	0.11
	T2	52	3.85	0.64	-0.97	1.14	0.09		T2	52	3.67	0.78	0.09	-0.95	0.11
	T3	48	3.85	0.63	-0.19	-0.39	0.09		T3	48	3.75	0.93	-0.49	-0.46	0.13
Pathways (revised)	T1	57	3.73	0.58	0.44	-0.34	0.08	Parent Relationships SC ^a	T1	57	3.80	0.82	-0.21	-0.76	0.11
	T2	52	3.73	0.58	0.17	-0.71	0.08		T2	52	3.79	0.90	-0.59	-0.47	0.12
	T3	48	3.75	0.65	-0.01	-0.71	0.09		T3	48	3.72	0.80	-0.96	1.56	0.11
Life Orientation Test, Revised								Honesty-Trust SC^a							
Optimism	T1	56	3.51	0.59	0.58	-0.40	0.08	Emotional Stability SC ^a	T1	56	3.33	0.77	0.23	-0.60	0.10
	T2	52	3.48	0.57	0.49	-0.38	0.08		T2	52	3.45	0.72	0.00	-0.64	0.10
	T3	48	3.42	0.50	-0.14	-0.95	0.07		T3	48	3.42	0.69	0.34	-0.58	0.10
Pessimism ^a	T1	56	2.74	0.66	-0.17	-0.02	0.09	Math SC ^a	T1	57	2.71	0.78	-0.08	-0.44	0.10
	T2	52	2.84	0.72	0.05	0.46	0.10		T2	52	2.90	0.75	-0.18	-0.38	0.10
	T3	48	3.03	0.63	-0.41	-0.58	0.09		T3	48	2.90	0.70	-0.10	0.17	0.10
Adolescent Self-Regulatory Inventory, Revised								Verbal SC^a							
Focus	T1	57	3.01	0.89	-0.31	-0.40	0.12	School SC ^a	T1	56	3.22	0.93	-0.37	0.11	0.12
	T2	52	3.18	0.70	0.43	0.02	0.10		T2	52	3.51	0.79	0.26	-0.96	0.11
	T3	48	3.26	0.66	0.01	0.00	0.10		T3	48	3.53	0.82	-0.22	0.00	0.12
Goal Self-Regulation	T1	57	3.62	0.60	-0.05	-0.57	0.08	General Self-Esteem ^a	T1	57	3.69	0.60	-0.80	0.93	0.08
	T2	52	3.72	0.65	0.30	-0.93	0.09		T2	52	3.70	0.66	-0.53	0.21	0.09
	T3	48	3.72	0.53	0.39	-0.57	0.08		T3	48	3.79	0.64	-0.24	-0.74	0.09
Emotion Self-Reg	T1	57	3.35	0.64	-0.51	0.79	0.09	Review of Personal Effectiveness and Locus of Control							
	T2	52	3.45	0.59	0.30	0.07	0.08	Self-Confidence	T1	56	3.75	0.63	-0.18	-0.62	0.08
	T3	48	3.32	0.55	-0.18	-0.28	0.08		T2	52	3.74	0.74	0.09	-0.99	0.10
Short Grit Scale									T3	48	3.82	0.68	-0.01	-0.54	0.10
Consistency of Interest ^a	T1	56	2.53	0.71	0.22	0.49	0.09	Self-Efficacy	T1	57	3.34	0.67	0.17	-0.17	0.09
	T2	52	2.60	0.57	0.11	-0.45	0.08		T2	52	3.52	0.69	-0.11	-0.60	0.10
	T3	48	2.72	0.54	0.13	-0.74	0.08		T3	48	3.44	0.62	-0.03	-0.07	0.09
Perseverance of Effort	T1	57	3.59	0.64	0.31	-0.46	0.08	Stress Management	T1	56	3.28	0.63	0.47	0.07	0.08
	T2	52	3.63	0.54	0.22	-0.70	0.07		T2	52	3.40	0.75	-0.24	-0.38	0.10
	T3	48	3.61	0.53	0.20	-0.91	0.08		T3	48	3.21	0.70	-0.06	-0.14	0.10
Motivation and Engagement Scale - Short								Open Thinking							
Booster Thoughts	T1	56*	3.77	0.60	0.07	-0.89	0.08	Social Effectiveness	T1	57	3.81	0.61	-0.04	-0.69	0.08
	T2	52	3.68	0.66	0.19	-0.84	0.09		T2	52	3.82	0.65	-0.24	-0.36	0.09
	T3	48	3.70	0.57	0.18	-0.81	0.08		T3	48	3.85	0.60	-0.09	-0.62	0.09
Booster Behaviours	T1	57	3.25	0.86	-0.19	-0.13	0.11	Cooperative Teamwork	T1	57	3.92	0.82	-0.97	1.34	0.11
	T2	52	3.33	0.72	0.07	-0.35	0.10		T2	52	4.05	0.64	0.04	-1.12	0.09
	T3	48	3.31	0.77	0.20	-0.13	0.11		T3	48	3.99	0.72	-0.34	-0.84	0.10
Hampering ^a	T1	57	2.88	0.61	0.65	1.79	0.08	Leadership Ability	T1	57	3.70	0.87	0.00	-1.25	0.12
	T2	52	2.85	0.58	-0.39	-0.37	0.08		T2	52	3.62	0.94	-0.34	-0.76	0.13
	T3	48	2.94	0.57	-0.27	0.31	0.08		T3	48	3.69	0.93	-0.36	-0.61	0.13
Life Resilience Scale								Time Efficiency							
Life Resilience	T1	57	3.35	0.60	-0.13	0.42	0.08	Quality Seeking	T1	56	4.04	0.67	-0.62	0.08	0.09
	T2	52	3.54	0.65	-0.25	0.34	0.09		T2	52	3.99	0.66	-0.15	-0.54	0.09
	T3	48	3.54	0.58	-0.31	1.03	0.08		T3	48	3.91	0.56	0.08	-0.77	0.08
Academic Resilience Scale								Coping with Change							
Academic Resilience	T1	57	3.34	0.61	0.25	0.68	0.08	Active Involvement	T1	56	3.33	0.58	0.21	-0.45	0.08
	T2	52	3.48	0.59	-0.53	0.74	0.08		T2	52	3.44	0.66	0.21	-0.55	0.09
	T3	48	3.60	0.57	-0.43	1.49	0.08		T3	48	3.42	0.68	0.13	-0.66	0.10
Satisfaction with Life Scale								Overall Effectiveness							
Life Satisfaction	T1	57	3.28	0.71	-0.17	0.80	0.09	Internal LOC	T1	57	3.44	0.79	0.17	-0.27	0.10
	T2	52	3.41	0.67	-0.14	0.62	0.09		T2	52	3.57	0.69	0.04	-0.26	0.10
	T3	48	3.40	0.71	-0.01	-0.16	0.10		T3	48	3.57	0.62	0.40	-0.60	0.09
Warwick-Edinburgh Mental Well-Being Scale								External LOC^a							
Wellbeing	T1	57	3.43	0.51	0.25	-0.16	0.07	Opp-Sex Relationships SC ^a	T1	56	3.28	1.02	-0.25	-0.68	0.14
	T2	52	3.58	0.65	-0.30	-0.20	0.09		T2	52	3.53	1.02	-0.56	-0.05	0.14
	T3	48	3.52	0.54	0.13	-0.60	0.08		T3	48	3.55	0.93	-0.48	-0.03	0.13
Self-Description Questionnaire II-Short															
Physical Abilities SC ^a	T1	57	3.46	1.07	-0.61	-0.23	0.14								
	T2	52	3.62	0.92	-0.60	0.04	0.13								
	T3	48	3.58	1.04	-0.56	-0.24	0.15								
Physical Appearance SC	T1	57	2.88	0.91	-0.10	-0.40	0.12								
	T2	52	2.99	0.83	-0.08	-0.04	0.11								
	T3	48	3.06	0.76	0.52	0.05	0.11								

Note. THP = The Helmsman Project; SD = standard deviation; SE = standard error; T1 = pre-test; T2 = immediate post-test; T3 = three months post-test; Self-Reg = Self-Regulation; SC = Self-Concept; Relationships = Relationships; Trust = Trustworthiness; LOC = Locus of Control.

^aAll or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

Table S.3
Descriptive Statistics: James Craig Adventure Program (T1-T3)

Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE
Children's Hope Scale								Self-Description Questionnaire II-Short							
Agency	T1	52	3.79	0.67	-0.53	-0.11	0.09	Same-Sex Relat'ships SC ^a	T1	52	3.74	0.73	-0.43	-0.20	0.10
	T2	40	3.96	0.55	-0.42	0.39	0.09		T2	40	3.80	0.67	-0.12	-0.66	0.11
	T3	34	3.93	0.60	-0.90	1.73	0.10		T3	34	3.99	0.57	0.00	-1.06	0.10
Pathways (revised)	T1	52	3.82	0.64	-0.25	-0.64	0.09	Parent Relat'ships SC ^a	T1	52	3.94	0.81	-0.78	0.57	0.11
	T2	40	4.06	0.50	0.23	-0.66	0.08		T2	40	3.91	0.87	-0.76	0.97	0.14
	T3	34	3.95	0.49	-0.35	0.09	0.08		T3	34	4.04	0.89	-1.09	1.73	0.15
Life Orientation Test, Revised								Honesty-Trust SC^a							
Optimism	T1	52	3.48	0.64	-0.12	-0.06	0.09	Emotional Stability SC ^a	T1	52	3.19	0.78	0.14	-0.05	0.11
	T2	40	3.86	0.61	0.14	-0.05	0.10		T2	40	3.18	0.83	-0.29	-0.20	0.13
	T3	34	3.66	0.54	0.38	0.75	0.09		T3	34	3.54	0.71	-0.20	-0.26	0.12
Pessimism ^a	T1	52	2.46	0.64	-0.04	-0.51	0.09	Math SC ^a	T1	52	2.83	0.78	0.17	-0.47	0.11
	T2	40	2.46	0.81	0.27	0.02	0.13		T2	40	2.86	0.75	0.49	0.66	0.12
	T3	34	2.65	0.80	0.60	-0.42	0.14		T3	34	2.91	0.82	0.71	0.03	0.14
Adolescent Self-Regulatory Inventory, Revised								Verbal SC^a							
Focus	T1	52	3.29	0.72	-0.15	-0.16	0.10	School SC ^a	T1	52	3.16	1.06	0.08	-0.58	0.15
	T2	40	3.39	0.71	0.00	-0.46	0.11		T2	40	3.12	1.10	0.09	-0.76	0.17
	T3	34	3.39	0.64	-0.73	-0.05	0.11		T3	34	2.93	1.15	0.13	-1.04	0.20
Goal Self-Regulation	T1	52	3.71	0.63	-0.35	-0.81	0.09	General Self-Esteem ^a	T1	52	3.27	1.01	-0.39	-0.22	0.14
	T2	40	3.90	0.65	-0.50	0.52	0.10		T2	40	3.53	0.83	0.13	-0.74	0.13
	T3	34	3.81	0.50	-0.10	-0.57	0.08		T3	34	3.55	0.83	-0.08	-0.86	0.14
Emotion Self-Reg	T1	52	3.56	0.49	-0.18	-0.70	0.07	School SC ^a	T1	52	3.62	0.83	-0.44	-0.07	0.12
	T2	40	3.76	0.62	0.23	-0.05	0.10		T2	40	3.60	0.74	-0.14	-0.27	0.12
	T3	34	3.74	0.53	0.76	0.10	0.09		T3	34	3.65	0.73	-0.19	0.05	0.13
Short Grit Scale								Review of Personal Effectiveness and Locus of Control							
Consistency of Interest ^a	T1	52	2.77	0.71	0.79	0.89	0.10	Self-Confidence	T1	52	3.67	0.80	-0.56	-0.06	0.11
	T2	40	2.64	0.85	0.48	0.06	0.14		T2	40	4.09	0.62	-0.07	-0.97	0.10
	T3	34	2.74	0.65	1.20	2.19	0.11		T3	34	3.91	0.55	0.22	-0.82	0.09
Perseverance of Effort	T1	52	3.54	0.63	-0.26	0.08	0.09	Self-Efficacy	T1	52	3.53	0.75	-0.14	-0.94	0.10
	T2	40	3.77	0.78	-0.26	-0.44	0.12		T2	40	3.84	0.72	-0.07	-1.12	0.11
	T3	34	3.65	0.55	0.19	-0.66	0.09		T3	34	3.80	0.59	-0.04	-1.12	0.10
Motivation and Engagement Scale - Short								Stress Management							
Booster Thoughts	T1	52	3.76	0.59	0.14	-0.65	0.08	Open Thinking	T1	52	3.53	0.93	-0.33	-0.85	0.13
	T2	40	3.91	0.63	-0.03	-0.75	0.10		T2	40	3.66	0.76	0.14	-0.78	0.12
	T3	34	4.02	0.54	0.03	-0.71	0.09		T3	34	3.80	0.73	-0.40	-0.29	0.13
Booster Behaviours	T1	52	3.22	0.74	-0.64	0.12	0.10	Social Effectiveness	T1	52	3.89	0.66	-1.57	4.59	0.09
	T2	40	3.65	0.73	-0.04	-0.33	0.12		T2	40	4.05	0.61	-0.27	-0.47	0.10
	T3	34	3.56	0.72	0.14	-0.50	0.12		T3	34	4.07	0.47	0.15	-0.03	0.08
Hampering ^a	T1	52	2.67	0.58	0.03	0.70	0.08	Cooperative Teamwork	T1	52	3.58	0.77	-0.46	-0.40	0.11
	T2	40	2.81	0.66	-0.13	0.51	0.10		T2	40	3.71	0.81	-0.64	0.37	0.13
	T3	34	2.86	0.63	0.78	0.86	0.11		T3	34	3.82	0.71	-0.68	0.82	0.12
Life Resilience Scale								Leadership Ability							
Life Resilience	T1	52	3.52	0.71	-0.35	-0.38	0.10	Time Efficiency	T1	52	3.58	1.01	-0.98	0.24	0.14
	T2	40	3.78	0.58	-0.52	-0.06	0.09		T2	40	3.91	0.93	-0.40	-0.59	0.15
	T3	34	3.73	0.62	-0.10	-0.88	0.11		T3	34	3.88	0.83	-0.37	-0.89	0.14
Academic Resilience Scale								Quality Seeking							
Academic Resilience	T1	52	3.47	0.72	-0.61	0.39	0.10	Coping with Change	T1	52	4.05	0.51	0.04	-0.61	0.07
	T2	40	3.66	0.72	-0.50	-0.07	0.11		T2	40	4.16	0.53	-0.12	0.05	0.08
	T3	34	3.55	0.63	-0.65	0.15	0.11		T3	34	3.39	0.65	0.26	-0.74	0.11
Satisfaction with Life Scale								Active Involvement							
Life Satisfaction	T1	52	3.22	0.73	-0.40	-0.09	0.10	Overall Effectiveness	T1	52	4.01	0.60	-0.85	-0.09	0.08
	T2	40	3.41	0.79	-0.36	-0.37	0.13		T2	40	4.08	0.69	-0.14	-1.04	0.11
	T3	34	3.56	0.69	-1.03	0.99	0.12		T3	34	4.06	0.59	-0.51	0.31	0.10
Warwick-Edinburgh Mental Well-Being Scale								Internal LOC							
Wellbeing	T1	52	3.58	0.54	-0.74	-0.10	0.07	External LOC ^a	T1	52	3.45	0.77	-0.49	0.58	0.11
	T2	40	3.87	0.56	0.33	-0.63	0.09		T2	40	3.63	0.81	-0.42	1.01	0.13
	T3	34	3.78	0.50	-0.16	-0.52	0.09		T3	34	3.51	0.65	-1.49	4.34	0.11
Self-Description Questionnaire II-Short								Internal LOC							
Physical Abilities SC ^a	T1	52	3.64	1.09	-0.79	-0.28	0.15	Internal LOC	T1	52	4.13	0.50	-0.62	0.23	0.07
	T2	40	3.73	1.14	-0.69	-0.61	0.18		T2	40	4.29	0.59	-0.31	-0.85	0.09
	T3	34	3.62	1.03	-0.69	0.22	0.18		T3	34	4.16	0.45	0.20	0.06	0.08
Physical Appearance SC	T1	52	2.83	0.91	0.07	-0.15	0.13	External LOC ^a	T1	52	3.21	0.89	-0.03	-0.32	0.12
	T2	40	3.09	0.99	-0.23	-0.42	0.16		T2	40	3.09	0.87	-0.26	0.13	0.14
	T3	34	3.22	0.93	-0.31	-0.31	0.16		T3	34	3.53	0.77	-0.81	1.63	0.13
Opp-Sex Relat'ships SC ^a	T1	52	3.35	0.93	-0.42	-0.18	0.13								
	T2	40	3.54	0.96	-0.42	-0.55	0.15								
	T3	34	3.75	0.74	-0.07	-0.91	0.13								

Note. THP = The Helmsman Project; SD = standard deviation; SE = standard error; T1 = pre-test; T2 = immediate post-test; T3 = three months post-test; Self-Reg = Self-Regulation; SC = Self-Concept; Relat'ships = Relationships; Trust = Trustworthiness; LOC = Locus of Control.

^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

Table S.4
Descriptive Statistics: Outward Bound Adventure Program (T1-T3)

Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE
Children's Hope Scale								Self-Description Questionnaire II-Short							
Agency	T1	44	3.91	0.55	-0.58	-0.32	0.08	Same-Sex Relationships SC ^a	T1	44	3.50	0.87	-0.56	0.33	0.13
	T2	36	4.16	0.54	-0.49	0.16	0.09		T2	36	3.67	0.68	0.02	-0.63	0.11
	T3	38	3.95	0.70	-0.86	0.12	0.11		T3	38	3.69	0.73	-0.15	-0.49	0.12
Pathways (revised)	T1	44	3.94	0.49	-0.31	-0.94	0.07	Parent Relationships SC ^a	T1	44	3.94	1.00	-0.98	0.11	0.15
	T2	36	3.85	0.53	-0.90	2.62	0.09		T2	36	3.72	0.96	-0.37	-0.73	0.16
	T3	38	3.92	0.58	-0.11	-0.27	0.09		T3	38	3.89	0.90	-0.88	1.09	0.15
Life Orientation Test, Revised								Honesty-Trust SC^a							
Optimism	T1	44	3.67	0.54	-0.04	-0.57	0.08	Emotional Stability SC ^a	T1	44	3.51	0.74	-0.34	-0.63	0.11
	T2	36	3.56	0.55	0.30	-0.32	0.09		T2	36	3.51	0.79	-0.18	-1.10	0.13
	T3	38	3.53	0.58	0.37	-0.37	0.09		T3	38	3.74	0.71	-0.09	-0.75	0.12
Pessimism ^a	T1	44	2.71	0.73	-0.31	-0.64	0.11	Math SC ^a	T1	44	2.60	0.66	-0.22	-0.13	0.10
	T2	36	2.88	0.76	-0.27	0.39	0.13		T2	36	2.63	0.73	-0.13	-0.60	0.12
	T3	37	2.82	0.81	0.14	-0.37	0.13		T3	38	2.68	0.73	-0.12	-0.85	0.12
Adolescent Self-Regulatory Inventory, Revised								Verbal SC^a							
Focus	T1	44	3.22	0.75	-0.59	0.35	0.11	School SC ^a	T1	44	3.30	1.12	-0.03	-1.08	0.17
	T2	36	3.27	0.71	-0.85	-0.04	0.12		T2	36	3.33	1.05	-0.31	-0.85	0.17
	T3	38	3.31	0.79	-0.78	-0.37	0.13		T3	38	3.26	1.19	-0.18	-1.21	0.19
Goal Self-Regulation	T1	44	3.94	0.54	-0.18	-1.03	0.08	General Self-Esteem ^a	T1	44	3.17	0.96	-0.13	-0.68	0.14
	T2	36	3.88	0.43	-1.03	2.51	0.07		T2	36	3.22	0.84	0.28	-0.60	0.14
	T3	38	3.85	0.64	-0.41	0.25	0.10		T3	38	3.33	1.07	-0.09	-0.90	0.17
Emotion Self-Reg	T1	44	3.45	0.58	-0.33	-0.84	0.09	T1	44	3.76	0.62	-1.03	2.47	0.09	
	T2	36	3.20	0.61	-0.35	-0.91	0.10		T2	36	3.67	0.86	-0.88	1.03	0.14
	T3	38	3.58	0.62	-1.05	1.56	0.10		T3	38	3.66	0.94	-0.36	-0.37	0.15
Short Grit Scale								Review of Personal Effectiveness and Locus of Control							
Consistency of Interest ^a	T1	44	2.70	0.66	0.73	2.21	0.10	Self-Confidence	T1	44	3.97	0.61	-0.81	0.90	0.09
	T2	36	2.56	0.58	0.25	0.49	0.10		T2	36	3.81	0.68	-0.82	0.93	0.11
	T3	38	2.64	0.65	0.09	-0.17	0.10		T3	38	3.85	0.91	-1.02	1.11	0.15
Perseverance of Effort	T1	44	3.68	0.51	-0.23	-0.21	0.08	Self-Efficacy	T1	44	3.52	0.59	-0.02	0.10	0.09
	T2	36	3.38	0.48	-0.31	-0.18	0.08		T2	36	3.65	0.77	-1.34	1.98	0.13
	T3	38	3.74	0.73	-0.23	-0.38	0.12		T3	38	3.71	0.66	-1.12	2.26	0.11
Motivation and Engagement Scale - Short								Stress Management							
Booster Thoughts	T1	44	3.94	0.60	-0.70	0.04	0.09	Open Thinking	T1	44	3.42	0.64	-1.03	1.09	0.10
	T2	36	3.81	0.52	0.19	-0.37	0.09		T2	36	3.56	0.73	-0.07	-0.30	0.12
	T3	38	3.95	0.64	-0.32	-0.93	0.10		T3	38	3.62	0.72	-1.06	2.78	0.12
Booster Behaviours	T1	44	3.42	0.73	-0.36	-0.64	0.11	Social Effectiveness	T1	44	4.05	0.55	-0.18	-0.22	0.08
	T2	36	3.35	0.80	-0.38	0.49	0.13		T2	36	4.18	0.48	0.17	-0.75	0.08
	T3	38	3.49	0.79	-0.02	-0.66	0.13		T3	38	4.21	0.47	0.16	-1.22	0.08
Hampering ^a	T1	44	2.83	0.69	-0.03	-0.36	0.10	Cooperative Teamwork	T1	44	3.68	0.66	-0.08	-0.46	0.10
	T2	36	3.06	0.59	-0.25	-0.39	0.10		T2	36	3.81	0.60	0.07	-0.82	0.10
	T3	38	2.93	0.68	-0.26	-0.81	0.11		T3	38	3.82	0.67	-0.33	-0.17	0.11
Life Resilience Scale								Leadership Ability							
Life Resilience	T1	44	3.54	0.45	-0.07	0.26	0.07	Time Efficiency	T1	44	3.82	0.73	-0.28	0.03	0.11
	T2	36	3.53	0.56	-0.41	0.04	0.09		T2	36	3.84	0.79	-0.58	-0.37	0.13
	T3	38	3.78	0.58	-0.56	-0.21	0.09		T3	38	3.72	0.96	-0.96	0.55	0.16
Academic Resilience Scale								Quality Seeking							
Academic Resilience	T1	44	3.46	0.50	0.02	0.85	0.08	Coping with Change	T1	44	4.17	0.60	-0.27	-0.77	0.09
	T2	36	3.63	0.46	-0.19	-0.02	0.08		T2	36	4.08	0.48	-0.21	-0.24	0.08
	T3	38	3.69	0.64	-0.30	-0.46	0.10		T3	38	3.35	0.79	-0.70	0.18	0.13
Satisfaction with Life Scale								Active Involvement							
Life Satisfaction	T1	44	3.48	0.76	-0.29	-0.43	0.11	Overall Effectiveness	T1	44	4.00	0.67	-0.47	0.36	0.11
	T2	36	3.38	0.56	-0.14	-1.20	0.09		T2	36	4.09	0.43	-0.28	0.72	0.07
	T3	38	3.50	0.83	-0.20	-0.35	0.13		T3	38	4.00	0.63	-0.56	-0.05	0.10
Warwick-Edinburgh Mental Well-Being Scale								Internal LOC							
Wellbeing	T1	44	3.71	0.40	0.01	-0.29	0.06	External LOC ^a	T1	44	3.68	0.68	0.25	-1.01	0.10
	T2	36	3.74	0.46	-0.76	0.94	0.08		T2	36	3.62	0.57	0.14	-0.24	0.09
	T3	38	3.68	0.62	-0.29	0.11	0.10		T3	38	3.69	0.76	-0.20	-0.44	0.12
Self-Description Questionnaire II-Short								Internal LOC							
Physical Abilities SC ^a	T1	44	3.37	1.18	-0.40	-1.03	0.18	Internal LOC	T1	44	4.24	0.56	0.05	-1.35	0.08
	T2	36	3.38	1.04	-0.63	-0.59	0.17		T2	36	4.25	0.68	-1.16	1.67	0.11
	T3	38	3.30	1.01	-0.19	-0.76	0.16		T3	38	4.18	0.62	-0.59	-0.15	0.10
Physical Appearance SC	T1	44	2.93	0.84	-0.37	0.08	0.13	External LOC ^a	T1	44	3.31	0.86	-0.45	-0.16	0.13
	T2	36	2.94	0.80	-0.29	0.01	0.13		T2	36	3.29	0.70	0.41	-0.38	0.12
	T3	37	2.96	1.00	-0.06	-0.24	0.16		T3	38	3.44	0.82	0.05	-0.68	0.13
Opp-Sex Relationships SC ^a	T1	44	3.30	0.92	0.07	-0.87	0.14								
	T2	36	3.26	0.81	0.06	-0.60	0.14								
	T3	38	3.46	0.79	0.10	-0.78	0.13								

Note. THP = The Helmsman Project; SD = standard deviation; SE = standard error; T1 = pre-test; T2 = immediate post-test; T3 = three months post-test; Self-Reg = Self-Regulation; SC = Self-Concept; Relationships = Relationships; Trust = Trustworthiness; LOC = Locus of Control.

^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

Table S.7
Descriptive Statistics: Control Group Adventure Programs (taken together) (T1-T4)

Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE						
Children's Hope Scale								Self-Description Questionnaire II-Short													
Agency	T1	114	3.85	0.62	-0.45	0.00	0.06	Same-Sex Relationships SC ^a	T1	114	3.87	0.78	-0.53	-0.43	0.07						
	T2	121	3.86	0.61	-0.48	0.69	0.06		T2	121	3.82	0.82	-0.44	-0.42	0.07						
	T3	123	3.93	0.60	-0.20	-0.57	0.05		T3	123	3.85	0.85	-0.61	0.33	0.08						
	T4	82	3.96	0.59	-0.21	-0.43	0.06		T4	82	3.84	0.79	-0.44	-0.45	0.09						
Pathways (revised)	T1	114	3.79	0.55	-0.09	-0.31	0.05	Parent Relationships SC ^a	T1	114	4.02	0.78	-0.83	0.30	0.07						
	T2	121	3.76	0.61	0.05	-0.04	0.06		T2	121	3.89	0.91	-0.82	0.15	0.08						
	T3	123	3.86	0.55	-0.13	-0.41	0.05		T3	123	3.92	0.88	-0.66	-0.03	0.08						
	T4	82	3.82	0.58	-0.16	-0.36	0.06		T4	82	3.89	0.91	-0.72	-0.28	0.10						
Life Orientation Test, Revised								Honesty-Trust SC^a													
Optimism	T1	114	3.68	0.68	-0.01	-0.37	0.06	T1	114	3.39	0.71	0.17	-0.76	0.07							
	T2	121	3.63	0.66	0.05	-0.60	0.06	T2	121	3.43	0.73	0.11	-0.44	0.07							
	T3	123	3.62	0.62	0.08	0.15	0.06	T3	123	3.48	0.69	0.07	-0.66	0.06							
	T4	82	3.69	0.66	-0.20	-0.48	0.07	T4	82	3.59	0.69	0.11	-0.78	0.08							
Pessimism ^a	T1	114	2.80	0.71	0.13	0.38	0.07	Emotional Stability SC ^a	T1	114	2.92	0.86	0.08	-0.31	0.08						
	T2	121	2.82	0.76	0.27	-0.21	0.07		T2	121	2.96	0.82	0.20	-0.25	0.07						
	T3	123	2.91	0.76	0.15	0.19	0.07		T3	123	3.03	0.85	0.30	-0.43	0.08						
	T4	82	2.87	0.84	0.38	-0.24	0.09		T4	82	3.02	0.83	0.25	-0.85	0.09						
Adolescent Self-Regulatory Inventory, Revised								Math SC^a													
Focus	T1	114	3.21	0.82	-0.28	-0.47	0.08	T1	114	3.10	1.14	-0.02	-0.98	0.11							
	T2	121	3.15	0.76	-0.07	-0.36	0.07	T2	121	3.03	1.14	0.13	-1.01	0.10							
	T3	123	3.26	0.77	-0.14	-0.09	0.07	T3	123	3.05	1.06	-0.09	-0.73	0.10							
	T4	82	3.31	0.71	-0.04	-0.92	0.08	T4	82	3.11	0.98	0.12	-0.67	0.11							
Goal Self-Regulation	T1	114	3.81	0.57	-0.12	-0.14	0.05	Verbal SC ^a	T1	114	3.24	0.91	-0.23	-0.14	0.09						
	T2	121	3.80	0.60	-0.19	-0.06	0.05		T2	121	3.24	0.98	-0.27	-0.46	0.09						
	T3	123	3.81	0.58	-0.04	-0.59	0.05		T3	123	3.30	0.98	-0.32	-0.33	0.09						
	T4	82	3.84	0.52	-0.02	-0.35	0.06		T4	82	3.35	0.89	-0.39	-0.29	0.10						
Emotion Self-Reg	T1	114	3.59	0.65	-0.29	0.14	0.06	School SC ^a	T1	114	3.70	0.75	-0.70	0.36	0.07						
	T2	121	3.55	0.62	-0.38	0.81	0.06		T2	121	3.60	0.74	-0.44	0.08	0.07						
	T3	123	3.59	0.61	-0.16	-0.50	0.05		T3	123	3.69	0.67	-0.28	0.15	0.06						
	T4	82	3.71	0.63	-0.49	0.71	0.07		T4	82	3.72	0.68	-0.49	0.89	0.07						
Short Grit Scale								General Self-Esteem^a													
Consistency of Interest ^a	T1	114	2.65	0.63	0.12	0.42	0.06	T1	114	3.80	0.65	-0.86	1.84	0.06							
	T2	121	2.67	0.60	0.32	0.14	0.05	T2	121	3.85	0.63	-0.47	0.37	0.06							
	T3	123	2.82	0.62	0.34	0.43	0.06	T3	123	3.84	0.61	-0.28	-0.05	0.06							
	T4	82	2.78	0.59	0.25	-0.48	0.06	T4	82	3.92	0.65	-0.46	-0.72	0.07							
Perseverance of Effort	T1	114	3.73	0.60	-0.18	-0.29	0.06	Review of Personal Effectiveness and Locus of Control													
	T2	121	3.68	0.66	0.12	-0.26	0.06	Self-Confidence													
	T3	123	3.73	0.62	0.20	-0.56	0.06	T1	114	3.88	0.72	-0.47	0.52	0.07							
	T4	82	3.77	0.59	0.26	-0.67	0.07	T2	121	3.84	0.73	-0.07	-0.79	0.07							
Motivation and Engagement Scale - Short								Self-Efficacy													
Booster Thoughts	T1	114	3.80	0.66	-0.47	0.44	0.06	T1	114	3.58	0.77	-0.07	-0.08	0.07							
	T2	121	3.73	0.69	-0.27	-0.23	0.06	T2	121	3.51	0.71	0.16	-0.22	0.06							
	T3	123	3.83	0.67	0.02	-0.60	0.06	T3	123	3.64	0.73	-0.12	-0.43	0.07							
	T4	82	3.82	0.69	-0.41	0.21	0.08	T4	82	3.71	0.61	0.09	-0.66	0.07							
Booster Behaviours	T1	114	3.37	0.73	-0.37	-0.45	0.07	Stress Management	T1	114	3.43	0.77	-0.11	-0.31	0.07						
	T2	121	3.30	0.77	-0.02	-0.48	0.07		T2	121	3.51	0.78	-0.13	0.27	0.07						
	T3	123	3.35	0.79	0.03	-0.46	0.07		T3	123	3.47	0.78	-0.40	0.29	0.07						
	T4	82	3.43	0.76	-0.03	-0.54	0.08		T4	82	3.61	0.66	-0.20	0.08	0.07						
Hampering ^a	T1	114	2.85	0.70	-0.02	-0.38	0.07	Open Thinking	T1	114	3.96	0.55	-0.34	0.08	0.05						
	T2	121	2.88	0.60	0.48	0.52	0.05		T2	121	3.95	0.61	-0.15	-0.29	0.06						
	T3	123	2.99	0.59	-0.08	-0.18	0.05		T3	123	3.97	0.59	-0.30	0.02	0.05						
	T4	82	3.01	0.64	0.24	-0.81	0.07		T4	82	3.98	0.53	-0.49	1.56	0.06						
Life Resilience Scale								Social Effectiveness													
Life Resilience	T1	114	3.49	0.73	-0.57	1.15	0.07	T1	114	3.61	0.84	-0.24	0.02	0.08							
	T2	121	3.60	0.60	0.13	0.02	0.05	T2	121	3.60	0.84	-0.26	-0.42	0.08							
	T3	123	3.66	0.64	-0.25	0.01	0.06	T3	123	3.72	0.82	-0.52	-0.08	0.07							
	T4	82	3.72	0.56	0.08	-0.36	0.06	T4	82	3.69	0.76	-0.20	-0.50	0.08							
Academic Resilience Scale								Cooperative Teamwork													
Academic Resilience	T1	114	3.46	0.69	-0.48	0.09	0.06	T1	114	4.18	0.68	-0.46	-0.77	0.06							
	T2	121	3.55	0.67	-0.17	-0.01	0.06	T2	121	4.08	0.71	-0.98	2.26	0.06							
	T3	123	3.67	0.64	-0.11	-0.25	0.06	T3	123	4.08	0.60	-0.29	-0.25	0.05							
	T4	82	3.57	0.59	0.21	-0.48	0.06	T4	82	4.18	0.66	-0.34	-0.57	0.07							
Satisfaction with Life Scale								Leadership Ability													
Life Satisfaction	T1	114	3.42	0.67	-0.30	0.30	0.06	T1	114	3.88	0.88	-0.70	0.44	0.08							
	T2	121	3.40	0.74	-0.04	-0.17	0.07	T2	121	3.88	0.90	-0.63	0.14	0.08							
	T3	123	3.47	0.72	-0.25	0.16	0.06	T3	123	3.93	0.87	-0.76	0.48	0.08							
	T4	82	3.50	0.76	-0.13	-0.76	0.08	T4	82	3.94	0.92	-0.88	0.43	0.10							
Warwick-Edinburgh Mental Well-Being Scale								Time Efficiency													
Wellbeing	T1	114	3.61	0.54	-0.49	0.88	0.05	T1	114	3.29	0.79	-0.17	-0.47	0.07							
	T2	121	3.64	0.57	-0.19	0.55	0.05	T2	121	3.25	0.80	0.04	-0.10	0.07							
	T3	123	3.62	0.58	-0.29	0.67	0.05	T3	123	3.37	0.76	-0.05	-0.43	0.07							
	T4	82	3.71	0.53	-0.12	-0.52	0.06	T4	82	3.43	0.72	0.07	-0.69	0.08							
Self-Description Questionnaire II-Short								Quality Seeking													
Physical Abilities SC ^a	T1	114	3.66	1.01	-0.34	-0.99	0.09	T1	114	4.14	0.64	-0.56	-0.18	0.06							
	T2	121	3.61	1.02	-0.48	-0.57	0.09	T2	121	4.00	0.63	-0.49	-0.09	0.06							
	T3	123	3.74	1.00	-0.49	-0.55	0.09	T3	123	3.99	0.60	-0.34	-0.15	0.05							
	T4	82	3.75	0.99	-0.60	-0.48	0.11	T4	82	4.05	0.59	-0.08	-0.66	0.07							
Physical Appearance SC	T1	113	3.00	0.98	-0.11	-0.40	0.09	Coping with Change	T1	114	3.62	0.67	-0.27	0.33	0.06						
	T2	121	3.10	1.02	-0.11	-0.59	0.09		T2	121	3.58	0.78	-0.35	-0.12	0.07						
	T3	123	3.23	0.97	-0.09	-0.35	0.09		T3	123	3.65	0.71	-0.75	1.76	0.06						
	T4	82	3.10	0.80	-0.27	0.65	0.09		T4	82	3.68	0.71	-0.31	0.01	0.08						
Opp-Sex Relationships SC ^a	T1	114	3.44	0.92	-0.45	-0.35	0.09	Active Involvement	T1	114	4.11	0.70	-0.69	-0.18	0.07						
	T2	121	3.48	0.93	-0.43	-0.27	0.08		T2	121	3.99	0.75	-0.70	0.05	0.07						
	T3	123	3.50	0.95	-0.41	-0.35	0.09		T3	123	4.05	0.68	-0.48	-0.06	0.06						
	T4	82	3.54	0.76	-0.63	0.25	0.08		T4	82	4.04	0.74	-0.47	-0.62	0.08						
								Overall Effectiveness													
								T1								114	3.59	0.68	-0.04	0.22	0.06
								T2								121	3.65	0.68	0.42	-0.72	0.06
								T3								123	3.64	0.71	-0.24	0.46	0.06
								T4								82	3.72	0.64	-0.20	-0.47	0.07
								Internal LOC													
								T1								114	4.24	0.53	-0.51	-0.42	0.05
								T2								121	4.20	0.62	-0.75	0.63	0.06
								T3								123	4.22	0.58	-0.25	-0.83	0.05
								T4								82	4.24	0.60	-0.45	-0.49	0.07
								External LOC^a													
								T1								114	3.50	0.75	-0.4		

Table S.8
Descriptive Statistics: Control Group Arctos Adventure Program (T1-T4)

Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	
Children's Hope Scale								Self-Description Questionnaire II-Short								
Agency	T1	41	3.83	0.73	-0.37	-0.57	0.11	Same-Sex Relationships SC ^a	T1	41	3.99	0.79	-0.68	-0.55	0.12	
	T2	45	3.70	0.63	-0.59	0.81	0.09		T2	45	3.68	0.88	-0.32	-0.72	0.13	
	T3	45	3.88	0.62	0.06	-0.52	0.09		T3	45	3.82	0.81	-0.78	1.34	0.12	
	T4	28	4.05	0.43	0.24	-0.63	0.08		T4	28	4.06	0.61	-0.49	-0.64	0.12	
Pathways (revised)	T1	41	3.73	0.60	-0.15	-0.66	0.09	Parent Relationships SC ^a	T1	41	4.14	0.74	-0.99	0.47	0.12	
	T2	45	3.64	0.57	0.06	0.09	0.08		T2	45	3.94	0.93	-0.81	-0.02	0.14	
	T3	45	3.82	0.55	-0.06	-0.15	0.08		T3	45	4.11	0.84	-1.27	2.34	0.13	
	T4	28	3.86	0.49	0.07	-0.56	0.09		T4	28	4.09	0.93	-1.20	0.23	0.18	
Life Orientation Test, Revised								Honesty-Trust SC ^a	T1	41	3.51	0.80	0.08	-1.01	0.13	
Optimism	T1	41	3.58	0.70	-0.23	0.04	0.11		T2	45	3.52	0.81	-0.08	-0.87	0.12	
	T2	45	3.49	0.68	0.15	-0.54	0.10		T3	45	3.62	0.71	-0.36	-0.75	0.11	
	T3	45	3.57	0.62	-0.03	0.73	0.09		T4	28	3.79	0.70	-0.22	-0.63	0.13	
	T4	28	3.80	0.58	0.09	-0.81	0.11		Emotional Stability SC ^a	T1	41	3.06	0.92	-0.14	-0.41	0.14
Pessimism ^a	T1	41	2.74	0.75	0.79	1.07	0.12		T2	45	3.00	0.81	0.16	-0.49	0.12	
	T2	45	2.83	0.70	0.44	0.59	0.10		T3	45	3.12	0.87	-0.06	-0.79	0.13	
	T3	45	2.99	0.61	0.67	0.89	0.09		T4	28	3.27	0.78	-0.17	-0.47	0.15	
	T4	28	3.07	0.73	0.38	-0.33	0.14		Math SC ^a	T1	41	3.06	1.00	0.03	-0.67	0.16
Adolescent Self-Regulatory Inventory, Revised									T2	45	2.84	1.12	0.17	-0.91	0.17	
Focus	T1	41	3.11	0.92	-0.10	-0.42	0.14		T3	45	2.98	1.03	-0.18	-0.44	0.15	
	T2	45	3.09	0.77	0.22	-0.24	0.11		T4	28	3.12	0.98	-0.05	-0.87	0.19	
	T3	45	3.18	0.79	0.19	-0.40	0.12		Verbal SC ^a	T1	41	3.17	0.99	-0.03	-0.76	0.16
	T4	28	3.35	0.72	-0.21	-1.01	0.14		T2	45	3.24	1.06	-0.25	-0.84	0.16	
Goal Self-Regulation	T1	41	3.76	0.71	-0.08	-0.72	0.11		T3	45	3.29	0.94	-0.29	-0.25	0.14	
	T2	45	3.68	0.69	-0.07	-0.46	0.10		T4	28	3.50	0.96	-0.12	-1.28	0.18	
	T3	45	3.77	0.60	-0.04	-0.81	0.09		School SC ^a	T1	41	3.62	0.87	-0.65	-0.07	0.14
	T4	28	3.87	0.52	-0.27	0.48	0.10		T2	45	3.48	0.75	-0.32	0.21	0.11	
Emotion Self-Reg	T1	41	3.66	0.75	-0.60	0.41	0.12		T3	45	3.66	0.63	0.09	-0.12	0.09	
	T2	45	3.48	0.65	-0.50	0.85	0.10		T4	28	3.73	0.68	-0.11	-0.58	0.13	
	T3	45	3.52	0.55	0.21	-0.20	0.08		General Self-Esteem ^a	T1	41	3.76	0.73	-0.71	0.74	0.11
	T4	28	3.68	0.70	-0.30	0.40	0.13		T2	45	3.65	0.70	-0.30	0.36	0.10	
Short Grit Scale									T3	45	3.80	0.64	-0.55	0.85	0.10	
Consistency of Interest ^a	T1	41	2.71	0.63	-0.14	-0.16	0.10		T4	28	4.07	0.50	-0.65	0.26	0.09	
	T2	45	2.74	0.61	0.55	0.30	0.09		Review of Personal Effectiveness and Locus of Control							
	T3	45	2.98	0.61	0.01	0.18	0.09		Self-Confidence	T1	41	3.74	0.85	-0.55	0.26	0.13
	T4	28	2.82	0.59	0.06	-0.85	0.11		T2	45	3.58	0.75	0.07	-0.85	0.11	
Perseverance of Effort	T1	41	3.68	0.71	-0.27	-0.44	0.11		T3	45	3.86	0.65	-0.15	-0.71	0.10	
	T2	45	3.46	0.71	0.09	-0.59	0.11		T4	28	3.96	0.52	0.34	-0.22	0.10	
	T3	45	3.71	0.64	0.16	-0.69	0.10		Self-Efficacy	T1	41	3.59	0.76	0.07	-0.79	0.12
	T4	28	3.71	0.56	0.22	-1.21	0.11		T2	45	3.37	0.72	0.54	-0.33	0.11	
Motivation and Engagement Scale - Short									T3	45	3.55	0.79	0.03	-0.50	0.12	
Booster Thoughts	T1	41	3.79	0.76	-0.43	-0.16	0.12		T4	28	3.65	0.66	0.33	-0.40	0.12	
	T2	45	3.63	0.69	-0.11	-0.35	0.10		Stress Management	T1	41	3.43	0.89	-0.35	-0.68	0.14
	T3	45	3.81	0.66	0.11	-0.70	0.10		T2	45	3.47	0.78	0.02	0.13	0.12	
	T4	28	3.96	0.60	-0.39	0.16	0.11		T3	45	3.33	0.79	-0.09	0.04	0.12	
Booster Behaviours	T1	41	3.29	0.80	0.13	-0.82	0.12		T4	28	3.67	0.65	-0.29	0.10	0.12	
	T2	45	3.13	0.75	0.56	0.00	0.11		Open Thinking	T1	41	3.91	0.59	-0.33	0.16	0.09
	T3	45	3.30	0.79	0.27	-0.51	0.12		T2	45	3.83	0.59	-0.01	-0.15	0.09	
	T4	28	3.46	0.75	-0.10	-0.30	0.14		T3	45	3.90	0.62	-0.27	0.42	0.09	
Hampering ^a	T1	41	2.96	0.70	0.00	-0.46	0.11		T4	28	4.05	0.37	0.97	1.11	0.07	
	T2	45	2.91	0.61	1.00	1.52	0.09		Social Effectiveness	T1	41	3.66	0.81	0.01	-0.18	0.13
	T3	45	3.03	0.55	0.12	-0.46	0.08		T2	45	3.50	0.84	-0.25	-0.32	0.13	
	T4	28	3.19	0.53	-0.29	-1.30	0.10		T3	45	3.67	0.75	-0.52	0.63	0.11	
Life Resilience Scale									T4	28	3.86	0.67	0.18	-0.94	0.13	
Life Resilience	T1	41	3.50	0.80	-0.41	0.34	0.13		Cooperative Teamwork	T1	41	4.34	0.66	-0.80	-0.41	0.10
	T2	45	3.49	0.64	0.21	0.03	0.10		T2	45	4.16	0.62	-0.38	-0.68	0.09	
	T3	45	3.58	0.68	-0.14	-0.30	0.10		T3	45	4.14	0.51	-0.27	-0.08	0.08	
	T4	28	3.83	0.55	0.46	-0.73	0.10		T4	28	4.31	0.57	-0.07	-1.45	0.11	
Academic Resilience Scale									Leadership Ability	T1	41	3.80	0.97	-0.77	0.07	0.15
Academic Resilience	T1	41	3.50	0.77	-0.73	0.54	0.12		T2	45	3.84	0.89	-0.84	0.57	0.13	
	T2	45	3.44	0.72	-0.02	0.08	0.11		T3	45	3.90	0.94	-1.01	1.04	0.14	
	T3	45	3.65	0.60	0.03	-0.45	0.09		T4	28	4.02	0.98	-1.37	1.78	0.19	
	T4	28	3.60	0.64	0.17	-0.44	0.12		Time Efficiency	T1	41	3.32	0.88	-0.05	-0.67	0.14
Satisfaction with Life Scale									T2	45	3.13	0.83	-0.04	0.22	0.12	
Life Satisfaction	T1	41	3.42	0.75	-0.53	-0.12	0.12		T3	45	3.37	0.69	0.14	0.10	0.10	
	T2	45	3.36	0.72	-0.31	0.04	0.11		T4	28	3.49	0.71	0.05	-0.93	0.13	
	T3	45	3.42	0.74	-0.64	0.71	0.11		Quality Seeking	T1	41	4.11	0.69	-0.75	0.25	0.11
	T4	28	3.68	0.71	-0.09	-0.74	0.13		T2	45	3.87	0.68	-0.63	0.14	0.10	
Warwick-Edinburgh Mental Well-Being Scale									T3	45	3.91	0.63	-0.47	0.30	0.09	
Wellbeing	T1	41	3.56	0.59	-0.85	0.85	0.09		T4	28	4.07	0.63	-0.14	-0.70	0.12	
	T2	45	3.53	0.62	-0.40	0.56	0.09		Coping with Change	T1	41	3.57	0.71	0.36	-0.60	0.11
	T3	45	3.55	0.62	-0.77	1.69	0.09		T2	45	3.61	0.73	-0.28	-0.32	0.11	
	T4	28	3.78	0.44	-0.21	0.56	0.08		T3	45	3.52	0.74	-0.65	0.91	0.11	
Self-Description Questionnaire II-Short									T4	28	3.77	0.70	-0.20	-0.30	0.13	
Physical Abilities SC ^a	T1	41	3.84	1.01	-0.67	-0.81	0.16		Active Involvement	T1	41	4.04	0.74	-0.55	-0.58	0.12
	T2	45	3.67	1.01	-0.60	-0.56	0.15		T2	45	3.88	0.75	-0.57	-0.10	0.11	
	T3	45	3.86	0.94	-0.57	-0.64	0.14		T3	45	3.91	0.75	-0.56	-0.01	0.11	
	T4	28	4.02	0.98	-1.35	1.53	0.19		T4	28	4.13	0.74	-0.92	0.20	0.14	
Physical Appearance SC	T1	41	2.82	0.86	-0.13	0.01	0.13		Overall Effectiveness	T1	41	3.55	0.79	-0.16	-0.04	0.12
	T2	45	2.92	0.93	0.19	-0.15	0.14		T2	45	3.45	0.65	0.77	-0.03	0.10	
	T3	45	3.17	0.98	-0.13	-0.32	0.15		T3	45	3.55	0.76	-0.39	0.40	0.11	
	T4	28	3.14	0.66	0.10	-0.37	0.12		T4	28	3.89	0.57	-0.08			

Table S.9
Descriptive Statistics: Control Group James Craig Adventure Program (T1-T4)

Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	
Children's Hope Scale								Self-Description Questionnaire II-Short								
Agency	T1	40	3.94	0.55	-0.32	-0.72	0.09	Same-Sex Relationships SC ^a	T1	40	3.87	0.71	-0.55	-0.06	0.11	
	T2	39	3.90	0.55	0.28	-0.84	0.09		T2	39	3.94	0.76	-0.38	-0.54	0.12	
	T3	40	3.95	0.60	-0.33	-0.85	0.09		T3	40	3.80	0.95	-0.71	-0.15	0.15	
	T4	27	4.05	0.56	0.10	-1.01	0.11		T4	27	3.78	0.85	-0.61	-0.32	0.16	
Pathways (revised)	T1	40	3.84	0.54	0.24	-0.41	0.08	Parent Relationships SC ^a	T1	40	3.92	0.86	-0.47	-0.83	0.14	
	T2	39	3.74	0.60	0.01	-0.08	0.10		T2	39	3.67	1.03	-0.49	-0.60	0.16	
	T3	40	3.89	0.54	-0.17	-0.73	0.08		T3	40	3.71	0.99	-0.18	-1.16	0.16	
	T4	27	3.82	0.46	0.01	-0.83	0.09		T4	27	3.83	0.97	-0.71	-0.27	0.19	
Life Orientation Test, Revised								Honesty-Trust SC ^a	T1	40	3.21	0.63	0.17	-0.51	0.10	
Optimism	T1	40	3.75	0.70	0.29	-1.07	0.11		T2	39	3.30	0.59	0.50	0.80	0.09	
	T2	39	3.67	0.69	0.26	-0.82	0.11		T3	40	3.33	0.59	0.24	-0.85	0.09	
	T3	40	3.73	0.63	0.34	-0.35	0.10		T4	27	3.30	0.56	0.01	-1.00	0.11	
	T4	27	3.71	0.62	-0.10	-0.95	0.12		Emotional Stability SC ^a	T1	40	2.90	0.84	-0.21	-0.14	0.13
Pessimism ^a	T1	40	2.65	0.65	-0.54	0.29	0.10		T2	39	2.98	0.84	0.25	-0.45	0.14	
	T2	39	2.78	0.76	0.49	-0.12	0.12		T3	40	3.02	0.84	0.38	-0.03	0.13	
	T3	40	2.79	0.76	0.32	-0.40	0.12		T4	27	2.96	0.87	0.42	-0.90	0.17	
	T4	27	2.88	0.93	0.56	-0.52	0.18		Math SC ^a	T1	40	3.03	1.20	-0.08	-1.17	0.19
Adolescent Self-Regulatory Inventory, Revised								T2	39	3.08	1.14	0.18	-1.23	0.18		
Focus	T1	40	3.17	0.76	-0.39	-0.86	0.12		T3	40	3.02	1.06	0.14	-1.10	0.17	
	T2	39	3.08	0.71	0.05	-0.83	0.11		T4	27	3.12	0.98	0.00	-0.71	0.19	
	T3	40	3.15	0.73	-0.53	0.62	0.12		Verbal SC ^a	T1	40	3.35	0.98	-0.32	-0.26	0.16
	T4	27	3.15	0.66	0.19	-0.82	0.13		T2	39	3.39	1.01	-0.43	-0.18	0.16	
Goal Self-Regulation	T1	40	3.83	0.53	-0.09	-0.58	0.08		T3	40	3.50	1.02	-0.36	-0.48	0.16	
	T2	39	3.80	0.56	-0.10	-0.65	0.09		T4	27	3.40	0.77	-0.86	1.08	0.15	
	T3	40	3.82	0.59	0.14	-1.00	0.09		School SC ^a	T1	40	3.82	0.66	-0.36	-0.08	0.10
	T4	27	3.80	0.52	-0.05	-1.23	0.10		T2	39	3.65	0.72	-0.28	-0.58	0.12	
Emotion Self-Reg	T1	40	3.56	0.60	-0.36	-0.60	0.10		T3	40	3.74	0.70	0.00	-0.59	0.11	
	T2	39	3.64	0.60	0.44	-0.33	0.10		T4	27	3.86	0.51	0.64	-0.20	0.10	
	T3	40	3.56	0.68	-0.17	-0.94	0.11		General Self-Esteem ^a	T1	40	3.88	0.66	-1.09	2.95	0.10
	T4	27	3.76	0.62	-0.49	0.03	0.12		T2	39	3.90	0.56	-0.21	-1.21	0.09	
Short Grit Scale								T3	40	3.84	0.63	-0.01	-1.30	0.10		
Consistency of Interest ^a	T1	40	2.56	0.70	0.37	0.91	0.11		T4	27	3.97	0.64	-0.62	-0.87	0.12	
	T2	39	2.62	0.60	0.17	-0.28	0.10		Review of Personal Effectiveness and Locus of Control							
	T3	40	2.70	0.68	0.56	0.87	0.11		Self-Confidence	T1	40	4.05	0.63	-0.17	-1.06	0.10
	T4	27	2.63	0.63	0.46	-0.62	0.12		T2	39	3.91	0.72	0.26	-1.34	0.12	
Perseverance of Effort	T1	40	3.86	0.54	-0.11	-0.80	0.09		T3	40	3.94	0.72	-0.12	-1.18	0.11	
	T2	39	3.83	0.62	0.26	-0.93	0.10		T4	27	3.96	0.72	-0.43	-0.66	0.14	
	T3	40	3.71	0.67	0.24	-0.81	0.11		Self-Efficacy	T1	40	3.73	0.86	-0.53	0.56	0.14
	T4	27	3.83	0.53	-0.31	-0.16	0.10		T2	39	3.57	0.72	0.10	-0.41	0.12	
Motivation and Engagement Scale - Short								T3	40	3.70	0.71	0.04	-0.62	0.11		
Booster Thoughts	T1	40	3.89	0.59	0.32	-0.82	0.09		T4	27	3.77	0.57	-0.04	-0.65	0.11	
	T2	39	3.77	0.67	0.17	-0.85	0.11		Stress Management	T1	40	3.46	0.72	0.13	-0.12	0.11
	T3	40	3.96	0.69	-0.29	-0.66	0.11		T2	39	3.56	0.83	-0.30	0.70	0.13	
	T4	27	3.83	0.62	0.26	-1.04	0.12		T3	40	3.52	0.75	-0.80	1.48	0.12	
Booster Behaviours	T1	40	3.33	0.73	-0.47	-0.64	0.12		T4	27	3.65	0.67	0.43	-0.71	0.13	
	T2	39	3.32	0.83	-0.25	-0.61	0.13		Open Thinking	T1	40	4.03	0.62	-0.28	-0.93	0.10
	T3	40	3.27	0.83	0.00	-0.20	0.13		T2	39	3.97	0.67	-0.25	-0.62	0.11	
	T4	27	3.33	0.72	0.24	-0.72	0.14		T3	40	3.90	0.61	-0.09	-0.84	0.10	
Hampering ^a	T1	40	2.70	0.68	-0.16	0.07	0.11		T4	27	4.00	0.49	-0.21	-0.35	0.09	
	T2	39	2.75	0.56	-0.01	0.24	0.09		Social Effectiveness	T1	40	3.67	0.89	-0.31	-0.83	0.14
	T3	40	2.89	0.62	-0.13	0.06	0.10		T2	39	3.72	0.86	-0.37	-0.43	0.14	
	T4	27	2.81	0.67	0.86	0.42	0.13		T3	40	3.92	0.78	-0.49	-0.72	0.12	
Life Resilience Scale								T4	27	3.79	0.77	-0.44	-0.43	0.15		
Life Resilience	T1	40	3.48	0.71	-1.06	2.29	0.11		Cooperative Teamwork	T1	40	4.19	0.70	-0.36	-1.23	0.11
	T2	39	3.67	0.59	0.20	-0.05	0.10		T2	39	3.93	0.85	-1.37	2.51	0.14	
	T3	40	3.72	0.64	-0.30	0.40	0.10		T3	40	3.95	0.71	-0.21	-0.84	0.11	
	T4	27	3.67	0.52	0.07	-0.67	0.10		T4	27	4.16	0.69	-0.23	-1.02	0.13	
Academic Resilience Scale								Leadership Ability	T1	40	3.93	0.93	-0.71	0.33	0.15	
Academic Resilience	T1	40	3.41	0.69	-0.34	-0.87	0.11		T2	39	3.75	0.98	-0.37	-0.24	0.16	
	T2	39	3.55	0.66	0.03	-0.51	0.11		T3	40	3.86	0.85	-0.36	-0.34	0.13	
	T3	40	3.62	0.73	0.04	-0.64	0.12		T4	27	3.93	0.91	-0.76	-0.16	0.17	
	T4	27	3.56	0.53	0.53	-0.57	0.10		Time Efficiency	T1	40	3.18	0.81	-0.18	-0.63	0.13
Satisfaction with Life Scale								T2	39	3.32	0.78	0.39	-0.33	0.12		
Life Satisfaction	T1	40	3.49	0.67	0.11	-0.04	0.11		T3	40	3.25	0.83	0.03	-0.78	0.13	
	T2	39	3.50	0.74	0.19	-0.08	0.12		T4	27	3.23	0.69	-0.17	-1.30	0.13	
	T3	40	3.52	0.71	0.31	-0.81	0.11		Quality Seeking	T1	40	4.14	0.65	-0.36	-0.94	0.10
	T4	27	3.60	0.76	0.02	-1.09	0.15		T2	39	4.02	0.59	-0.13	-1.01	0.09	
Warwick-Edinburgh Mental Well-Being Scale								T3	40	4.03	0.56	-0.30	-1.13	0.09		
Wellbeing	T1	40	3.68	0.54	-0.25	0.41	0.09		T4	27	4.07	0.54	-0.40	-0.80	0.10	
	T2	39	3.72	0.55	0.31	-0.53	0.09		Coping with Change	T1	40	3.69	0.71	-0.97	1.57	0.11
	T3	40	3.63	0.62	0.20	-0.80	0.10		T2	39	3.50	0.79	-0.12	-0.48	0.13	
	T4	27	3.80	0.56	-0.05	-1.27	0.11		T3	40	3.73	0.62	0.24	-0.26	0.10	
Self-Description Questionnaire II-Short								T4	27	3.73	0.60	-0.15	-0.61	0.12		
Physical Abilities SC ^a	T1	40	3.50	1.07	-0.11	-1.19	0.17		Active Involvement	T1	40	4.15	0.80	-0.81	-0.42	0.13
	T2	39	3.34	1.14	-0.21	-0.97	0.18		T2	39	3.97	0.82	-0.74	-0.21	0.13	
	T3	40	3.54	1.10	-0.16	-1.06	0.17		T3	40	4.08	0.69	-0.04	-1.23	0.11	
	T4	27	3.44	1.05	-0.05	-1.32	0.20		T4	27	4.16	0.74	-0.55	-0.60	0.14	
Physical Appearance SC	T1	39	3.15	1.04	-0.19	-0.59	0.17		Overall Effectiveness	T1	40	3.64	0.68	0.41	-0.47	0.11
	T2	39	3.19	1.10	-0.22	-0.66	0.18		T2	39	3.73	0.73	0.41	-1.05	0.12	
	T3	40	3.23	1.05	0.10	-0.77	0.17		T3	40	3.59	0.77	-0.11	-0.16	0.12	
	T4	27	3.16	0.82	-0.04	0.67	0.16		T4	27	3.56	0.64	-0.35	-0.74	0.12	
Opp-Sex Relationships SC ^a																

Table S.10
Descriptive Statistics: Control Group Outward Bound Adventure Program (T1-T4)

Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE
Children's Hope Scale								Self-Description Questionnaire II-Short							
Agency	T1	33	3.78	0.55	-0.68	0.90	0.10	Same-Sex Relationships SC ^a	T1	33	3.73	0.85	-0.29	-0.78	0.15
	T2	37	4.02	0.62	-0.86	1.19	0.10		T2	37	3.88	0.80	-0.52	-0.35	0.13
	T3	38	3.96	0.60	-0.36	-0.47	0.10		T3	38	3.94	0.82	-0.15	-0.93	0.13
	T4	27	3.79	0.72	-0.07	-1.14	0.14		T4	27	3.67	0.87	0.10	-1.05	0.17
Pathways (revised)	T1	33	3.82	0.53	-0.27	-0.24	0.09	Parent Relationships SC ^a	T1	33	3.98	0.72	-1.14	2.04	0.13
	T2	37	3.93	0.65	-0.11	-0.30	0.11		T2	37	4.08	0.71	-0.97	1.02	0.12
	T3	38	3.89	0.57	-0.16	-0.64	0.09		T3	38	3.93	0.79	-0.43	-0.52	0.13
	T4	27	3.76	0.75	-0.11	-1.10	0.15		T4	27	3.74	0.81	-0.13	-0.78	0.16
Life Orientation Test, Revised								Honesty-Trust SC^a							
Optimism	T1	33	3.71	0.61	-0.08	-1.07	0.11	Emotional Stability SC ^a	T1	33	3.46	0.66	0.00	-1.27	0.12
	T2	37	3.76	0.60	-0.26	-0.60	0.10		T2	37	3.45	0.77	-0.05	-0.66	0.13
	T3	38	3.56	0.61	-0.12	-0.69	0.10		T3	38	3.49	0.73	0.25	-0.50	0.12
	T4	27	3.56	0.77	-0.17	-0.84	0.15		T4	27	3.68	0.74	0.11	-1.20	0.14
Pessimism ^a	T1	33	3.06	0.69	-0.27	-0.74	0.12	Math SC ^a	T1	33	2.78	0.81	0.67	-0.29	0.14
	T2	37	2.86	0.83	-0.06	-1.03	0.14		T2	37	2.89	0.82	0.18	-0.08	0.14
	T3	38	2.93	0.91	-0.09	-0.30	0.15		T3	38	2.92	0.85	0.65	-0.30	0.14
	T4	27	2.65	0.82	0.28	-0.43	0.16		T4	27	2.83	0.81	0.57	-0.79	0.16
Adolescent Self-Regulatory Inventory, Revised								General Self-Esteem^a							
Focus	T1	33	3.38	0.75	-0.29	-1.00	0.13	Verbal SC ^a	T1	33	3.23	1.24	-0.04	-1.31	0.22
	T2	37	3.30	0.78	-0.55	0.01	0.13		T2	37	3.21	1.15	0.03	-1.13	0.19
	T3	38	3.47	0.78	-0.30	-0.43	0.13		T3	38	3.16	1.11	-0.24	-0.85	0.18
	T4	27	3.44	0.73	-0.13	-1.04	0.14		T4	27	3.08	1.01	0.37	-0.75	0.20
Goal Self-Regulation	T1	33	3.84	0.43	0.22	-0.07	0.07	School SC ^a	T1	33	3.66	0.69	-0.80	-0.06	0.12
	T2	37	3.96	0.50	0.12	0.29	0.08		T2	37	3.70	0.75	-0.73	0.41	0.12
	T3	38	3.85	0.54	-0.25	0.04	0.09		T3	38	3.68	0.70	-0.90	0.68	0.11
	T4	27	3.86	0.56	0.20	-0.69	0.11		T4	27	3.56	0.80	-0.66	0.33	0.15
Emotion Self-Reg	T1	33	3.56	0.59	0.40	-0.44	0.10	Self-Confidence	T1	33	3.83	0.60	0.12	-0.39	0.10
	T2	37	3.56	0.62	-0.94	0.90	0.10		T2	37	4.08	0.61	-0.33	-0.40	0.10
	T3	38	3.70	0.59	-0.51	-0.16	0.10		T3	38	3.96	0.68	-0.53	0.01	0.11
	T4	27	3.70	0.59	-0.69	1.10	0.11		T4	27	3.84	0.86	-0.18	-1.30	0.17
Short Grit Scale								Review of Personal Effectiveness and Locus of Control							
Consistency of Interest ^a	T1	33	2.67	0.54	0.15	-0.96	0.09	Self-Efficacy	T1	33	3.38	0.64	0.30	-0.58	0.11
	T2	37	2.66	0.57	0.10	-0.28	0.09		T2	37	3.61	0.66	-0.24	0.60	0.11
	T3	38	2.75	0.53	0.55	-0.07	0.09		T3	38	3.69	0.69	-0.44	-0.49	0.11
	T4	27	2.88	0.53	0.43	-0.04	0.10		T4	27	3.72	0.63	-0.04	-1.27	0.12
Perseverance of Effort	T1	33	3.62	0.51	0.24	-0.92	0.09	Stress Management	T1	33	3.37	0.69	0.13	-0.58	0.12
	T2	37	3.80	0.56	0.74	-0.22	0.09		T2	37	3.52	0.74	-0.09	-0.68	0.12
	T3	38	3.78	0.54	0.31	-0.47	0.09		T3	38	3.59	0.79	-0.39	-0.41	0.13
	T4	27	3.77	0.69	0.54	-0.92	0.13		T4	27	3.51	0.69	-0.66	-0.04	0.13
Motivation and Engagement Scale - Short								Open Thinking							
Booster Thoughts	T1	33	3.72	0.62	-1.24	1.08	0.11	Social Effectiveness	T1	33	3.92	0.40	-0.91	1.46	0.07
	T2	37	3.81	0.71	-0.83	0.39	0.12		T2	37	4.08	0.55	-0.09	-0.38	0.09
	T3	38	3.71	0.63	0.21	-0.32	0.10		T3	38	4.13	0.49	-0.28	-0.29	0.08
	T4	27	3.65	0.82	-0.41	-0.29	0.16		T4	27	3.89	0.69	-0.51	0.35	0.13
Booster Behaviours	T1	33	3.51	0.62	-1.07	0.93	0.11	Cooperative Teamwork	T1	33	3.49	0.83	-0.42	0.98	0.15
	T2	37	3.48	0.70	-0.39	-0.41	0.12		T2	37	3.60	0.82	-0.14	-0.80	0.13
	T3	38	3.51	0.75	-0.15	-1.02	0.12		T3	38	3.56	0.90	-0.43	-0.58	0.15
	T4	27	3.49	0.83	-0.22	-0.82	0.16		T4	27	3.41	0.78	-0.03	-0.84	0.15
Hampering ^a	T1	33	2.89	0.72	0.06	-1.18	0.13	Leadership Ability	T1	33	3.92	0.72	-0.15	-0.54	0.14
	T2	37	2.98	0.63	0.17	-0.96	0.10		T2	37	4.05	0.82	-0.53	-0.68	0.14
	T3	38	3.04	0.61	-0.14	-0.66	0.10		T3	38	4.05	0.81	-0.65	-0.44	0.13
	T4	27	3.02	0.69	0.22	-1.26	0.13		T4	27	3.85	0.89	-0.20	-1.03	0.17
Life Resilience Scale								Time Efficiency							
Life Resilience	T1	33	3.49	0.67	-0.15	0.19	0.12	Quality Seeking	T1	33	3.38	0.65	-0.30	-0.94	0.11
	T2	37	3.65	0.55	0.07	-0.43	0.09		T2	37	3.32	0.80	-0.12	-0.84	0.13
	T3	38	3.69	0.61	-0.28	-0.32	0.10		T3	38	3.48	0.75	-0.22	-0.74	0.12
	T4	27	3.67	0.63	-0.14	-0.65	0.12		T4	27	3.56	0.76	0.19	-0.76	0.15
Academic Resilience Scale								Coping with Change							
Academic Resilience	T1	33	3.46	0.58	-0.04	-0.44	0.10	Active Involvement	T1	33	4.17	0.57	-0.29	-0.89	0.10
	T2	37	3.66	0.60	-0.51	0.16	0.10		T2	37	4.44	0.60	-0.36	-0.96	0.10
	T3	38	3.74	0.59	-0.45	0.25	0.10		T3	38	4.04	0.60	-0.11	-0.58	0.10
	T4	27	3.57	0.60	-0.03	-1.01	0.11		T4	27	4.01	0.62	0.23	-0.84	0.12
Satisfaction with Life Scale								Overall Effectiveness							
Life Satisfaction	T1	33	3.33	0.58	-0.59	0.60	0.10	Internal LOC	T1	33	3.80	0.63	-0.02	-0.81	0.10
	T2	37	3.35	0.78	0.01	-0.87	0.13		T2	37	3.80	0.58	0.43	-0.32	0.09
	T3	38	3.46	0.71	-0.28	-0.27	0.12		T3	38	3.80	0.58	0.43	-0.32	0.09
	T4	27	3.22	0.75	-0.24	-1.33	0.14		T4	27	3.69	0.67	0.01	-1.06	0.13
Warwick-Edinburgh Mental Well-Being Scale								External LOC^a							
Wellbeing	T1	33	3.59	0.47	0.07	0.01	0.08	Internal LOC	T1	33	4.27	0.48	-0.65	0.04	0.08
	T2	37	3.68	0.50	-0.09	0.08	0.08		T2	37	4.42	0.46	-0.72	0.49	0.08
	T3	38	3.68	0.51	-0.05	-0.44	0.08		T3	38	4.24	0.48	-0.07	-0.88	0.08
	T4	27	3.54	0.57	0.04	-0.76	0.11		T4	27	4.11	0.71	-0.38	-0.92	0.14
Self-Description Questionnaire II-Short								External LOC^a							
Physical Abilities SC ^a	T1	33	3.63	0.91	-0.16	-0.88	0.16	Internal LOC	T1	33	3.60	0.54	-0.23	1.27	0.09
	T2	37	3.82	0.86	-0.27	-0.91	0.14		T2	37	3.57	0.63	0.17	-0.29	0.10
	T3	38	3.81	0.95	-0.71	0.20	0.15		T3	38	3.64	0.71	-0.94	1.00	0.11
	T4	27	3.77	0.90	-0.40	-0.79	0.17		T4	27	3.48	0.71	0.28	-0.47	0.14
Physical Appearance SC	T1	33	3.04	1.04	-0.19	-0.79	0.18	External LOC ^a	T1	33	3.60	0.54	-0.23	1.27	0.09
	T2	37	3.21	1.04	-0.38	-0.89	0.17		T2	37	3.57	0.63	0.17	-0.29	0.10
	T3	38	3.31	0.89	-0.29	-0.11	0.14		T3	38	3.64	0.71	-0.94	1.00	0.11
	T4	27	3.01	0.93	-0.43	-0.06	0.18		T4	27	3.48	0.71	0.28	-0.47	0.14
Opp-Sex Relationships SC ^a	T1	33	3.30	0.95	0.03	-0.60	0.16	External LOC ^a	T1	33	3.60	0.54	-0.23	1.27	0.09
	T2	37	3.47	0.92	-0.10	-0.95	0.15		T2	37	3.57	0.63	0.17	-0.29	0.10
	T3	38	3.48	0.90	-0.21	-0.60	0.15		T3	38	3.64	0.71	-0.94	1.00	0.11
	T4	27	3.40	0.67	0.35	0.03	0.13		T4	27	3.48	0.71	0.28	-0.47	0.14

Note. THP = The Helmsman Project; SD = standard deviation; SE = standard error; T1, T2, T3 = extended baseline pre-test; T4 = immediate post-test; Self-Reg = Self-Regulation; SC = Self-Concept; Relationships = Relationships; Trust = Trustworthiness; LOC = Locus of Control.

^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

Table S.11
Descriptive Statistics: Control Group Coaching Only Program (T1-T4)

Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE	Scale	Wave	Numbers	Mean	SD	Skewness	Kurtosis	SE
Children's Hope Scale								Self-Description Questionnaire II-Short							
Agency	T1	26	4.18	0.53	-0.18	-0.78	0.10	Same-Sex Relationships SC ^a	T1	26	3.89	0.71	-0.53	-0.45	0.14
	T2	23	3.88	0.72	0.05	-1.32	0.15		T2	23	3.80	0.87	-1.36	2.17	0.18
	T3	26	3.91	0.47	0.04	-0.16	0.09		T3	26	4.00	0.59	-0.32	-0.32	0.12
	T4	18	4.09	0.47	0.37	-0.46	0.11		T4	18	3.87	0.72	-0.02	-1.13	0.17
Pathways (revised)	T1	26	4.06	0.54	0.04	-0.63	0.10	Parent Relationships SC ^a	T1	26	4.37	0.80	-1.69	2.16	0.16
	T2	23	3.90	0.68	-0.29	-0.63	0.14		T2	23	4.28	1.00	-1.62	1.72	0.21
	T3	26	3.97	0.54	-0.04	-0.15	0.11		T3	26	4.36	0.66	-0.83	-0.41	0.13
	T4	18	4.11	0.60	-0.07	-1.23	0.14		T4	18	4.21	0.82	-0.51	-1.19	0.19
Life Orientation Test, Revised								Honesty-Trust SC^a							
Optimism	T1	26	3.97	0.65	-0.31	-0.74	0.13	T1	26	3.53	0.79	-0.31	-0.40	0.15	
	T2	23	3.70	0.80	-0.17	-0.62	0.17	T2	23	3.41	0.91	-0.84	-0.08	0.19	
	T3	26	3.82	0.61	-0.29	-0.34	0.12	T3	26	3.66	0.80	-0.50	-0.46	0.16	
	T4	18	3.65	0.93	-0.36	0.00	0.22	T4	18	3.55	0.80	0.39	-1.47	0.19	
Pessimism ^a	T1	26	2.76	0.79	0.69	-0.52	0.15	Emotional Stability SC ^a	T1	26	2.88	0.73	-0.56	-0.01	0.14
	T2	23	2.77	0.98	-0.14	-1.08	0.20		T2	22	2.81	0.90	-0.46	-0.84	0.19
	T3	26	2.96	0.89	0.23	0.02	0.17		T3	26	2.75	0.74	0.07	0.29	0.15
	T4	18	2.94	1.07	-0.04	-0.54	0.25		T4	18	2.98	0.78	0.61	0.41	0.18
Adolescent Self-Regulatory Inventory, Revised								Math SC^a							
Focus	T1	26	3.68	0.79	-0.70	-0.52	0.15	T1	26	3.27	1.11	-0.15	-0.74	0.22	
	T2	23	3.63	0.87	-0.62	-0.18	0.18	T2	23	2.87	1.21	0.10	-1.25	0.25	
	T3	26	3.63	0.73	-0.52	-0.85	0.14	T3	26	3.09	1.03	0.06	-0.84	0.20	
	T4	18	3.59	0.91	-1.27	1.66	0.21	T4	18	2.86	1.03	0.09	-0.80	0.24	
Goal Self-Regulation	T1	26	4.18	0.52	-0.16	-1.20	0.10	Verbal SC ^a	T1	26	3.44	0.98	-0.51	-0.11	0.19
	T2	23	4.07	0.67	-0.23	-0.81	0.14		T2	23	3.56	0.97	-0.34	-0.94	0.20
	T3	26	3.94	0.47	-0.10	-0.31	0.09		T3	26	3.60	0.96	-0.50	0.07	0.19
	T4	18	4.06	0.54	-0.37	0.10	0.13		T4	18	3.71	0.94	-0.10	-0.99	0.22
Emotion Self-Reg	T1	26	3.69	0.68	-0.17	-0.97	0.13	School SC ^a	T1	26	3.88	0.81	-0.85	0.27	0.16
	T2	23	3.79	0.75	-0.52	-0.46	0.16		T2	23	3.71	0.78	0.03	-1.08	0.16
	T3	26	4.01	0.56	-0.41	0.07	0.11		T3	26	3.80	0.56	-0.27	-0.58	0.11
	T4	18	3.88	0.55	-0.37	0.07	0.13		T4	18	3.67	0.73	-0.34	-0.42	0.17
Short Grit Scale								General Self-Esteem^a							
Consistency of Interest ^a	T1	26	2.61	0.73	0.04	-0.86	0.14	T1	26	4.08	0.56	-0.16	-0.63	0.11	
	T2	23	2.66	0.82	-0.41	-1.15	0.17	T2	23	3.99	0.68	-0.14	-1.30	0.14	
	T3	26	2.58	0.69	-0.73	-0.32	0.14	T3	26	4.03	0.48	-0.25	-0.13	0.09	
	T4	18	2.93	0.89	-0.38	-1.01	0.21	T4	18	4.12	0.62	-0.41	-1.09	0.15	
Perseverance of Effort	T1	26	4.07	0.52	-0.32	-0.36	0.10	Review of Personal Effectiveness and Locus of Control							
	T2	23	3.74	0.61	-0.71	-0.45	0.13	Self-Confidence							
	T3	26	3.94	0.58	0.01	-0.59	0.11	T1	26	4.26	0.66	-0.87	0.50	0.13	
	T4	18	3.99	0.71	-0.15	-1.29	0.17	T2	23	4.01	0.84	-0.36	-1.16	0.17	
Motivation and Engagement Scale - Short								Self-Efficacy							
Booster Thoughts	T1	26	4.10	0.70	-0.34	-0.99	0.14	T1	26	3.77	0.82	-0.41	-0.53	0.16	
	T2	22	3.82	0.83	-0.32	-0.63	0.18	T2	23	3.80	0.87	-0.18	-1.16	0.18	
	T3	26	3.78	0.70	-0.02	-1.18	0.14	T3	26	3.77	0.65	0.32	-0.75	0.13	
	T4	18	3.93	0.98	-1.35	1.93	0.23	T4	18	3.98	0.63	0.03	-1.23	0.15	
Booster Behaviours	T1	26	3.92	0.78	-0.72	-0.10	0.15	Stress Management	T1	26	3.56	0.88	-0.64	-0.34	0.17
	T2	22	3.71	0.76	0.00	-0.93	0.16		T2	23	3.62	0.82	-0.15	-0.95	0.17
	T3	26	3.69	0.62	-0.18	-1.14	0.12		T3	26	3.87	0.54	0.45	-0.48	0.11
	T4	18	3.57	1.05	-0.78	-0.24	0.25		T4	18	3.93	0.64	0.48	-1.14	0.15
Hampering ^a	T1	26	3.02	0.59	-0.03	-0.14	0.12	Open Thinking	T1	26	4.33	0.51	-0.52	0.16	0.10
	T2	23	2.83	0.93	0.12	-0.57	0.19		T2	23	4.25	0.57	-0.50	-0.85	0.12
	T3	26	2.98	0.46	0.01	-0.60	0.09		T3	26	4.12	0.53	-0.27	-0.31	0.10
	T4	18	2.93	0.81	0.64	0.24	0.19		T4	18	4.19	0.75	-1.22	1.38	0.18
Life Resilience Scale								Social Effectiveness							
Life Resilience	T1	26	3.71	0.65	-0.31	-0.90	0.13	T1	26	3.86	0.78	-0.06	-0.98	0.15	
	T2	23	3.84	0.74	-0.21	-0.84	0.16	T2	23	3.93	0.72	-0.39	0.26	0.15	
	T3	26	3.90	0.57	0.13	-0.71	0.11	T3	26	3.88	0.69	-0.29	-0.55	0.13	
	T4	18	4.02	0.62	-0.05	-0.56	0.15	T4	18	4.07	0.75	-0.83	0.81	0.18	
Academic Resilience Scale								Cooperative Teamwork							
Academic Resilience	T1	26	3.66	0.70	0.03	-1.00	0.14	T1	26	4.13	0.70	-0.37	-0.39	0.14	
	T2	23	3.67	0.89	-0.59	-0.77	0.19	T2	23	4.06	0.77	-0.27	-1.26	0.16	
	T3	26	3.78	0.59	-0.27	-0.35	0.12	T3	26	4.04	0.62	0.20	-1.03	0.12	
	T4	18	3.69	0.66	0.06	-0.29	0.15	T4	18	4.24	0.59	-0.50	-0.80	0.14	
Satisfaction with Life Scale								Leadership Ability							
Life Satisfaction	T1	26	3.82	0.65	-0.73	0.12	0.13	T1	26	4.03	0.86	-0.45	-1.03	0.17	
	T2	22	3.57	0.92	-0.28	-1.17	0.20	T2	23	3.93	0.85	-0.47	-0.60	0.18	
	T3	26	3.69	0.67	-0.29	-0.70	0.13	T3	26	3.92	0.80	-0.75	0.43	0.16	
	T4	18	3.79	0.73	0.08	-1.31	0.17	T4	18	4.02	0.90	-0.23	-1.68	0.21	
Warwick-Edinburgh Mental Well-Being Scale								Time Efficiency							
Wellbeing	T1	26	4.02	0.56	0.12	-0.80	0.11	T1	26	3.46	0.97	-0.53	-0.72	0.19	
	T2	23	3.76	0.73	-0.03	-1.08	0.15	T2	23	3.46	0.88	-0.61	0.76	0.18	
	T3	26	3.89	0.54	0.28	-0.62	0.11	T3	26	3.64	0.82	-0.37	-0.90	0.16	
	T4	18	3.99	0.62	0.39	-1.12	0.15	T4	18	3.70	1.10	-0.76	-0.21	0.26	
Self-Description Questionnaire II-Short								Quality Seeking							
Physical Abilities SC ^a	T1	26	3.50	1.25	-0.73	-0.82	0.24	T1	26	4.40	0.50	-0.38	-1.08	0.10	
	T2	23	3.95	0.96	-1.42	1.58	0.20	T2	23	4.23	0.63	-0.58	-0.57	0.13	
	T3	26	3.51	1.10	-0.47	-0.82	0.22	T3	26	4.10	0.65	-0.51	-0.89	0.13	
	T4	18	3.82	1.14	-0.81	-0.28	0.27	T4	18	4.15	0.73	-0.64	-0.51	0.17	
Physical Appearance SC	T1	26	3.18	1.01	-0.37	-0.54	0.20	Coping with Change	T1	26	3.79	0.56	0.17	0.33	0.11
	T2	23	3.25	0.91	0.06	0.32	0.19		T2	23	4.12	0.62	-0.12	-0.99	0.13
	T3	26	3.30	0.90	-0.42	0.22	0.18		T3	26	3.94	0.57	0.20	-0.55	0.11
	T4	18	3.42	0.82	0.40	-0.75	0.19		T4	18	4.15	0.59	-0.28	-0.51	0.14
Opp-Sex Relationships SC ^a	T1	26	3.46	0.96	-0.14	-0.72	0.19	Active Involvement	T1	26	4.00	0.71	-0.41	-0.65	0.14
	T2	22	3.56	0.78	0.12	-0.83	0.17		T2	23	4.17	0.69	-0.39	-1.30	0.14
	T3	26	3.53	0.76	0.06	-0.68	0.15		T3	26	4.04	0.66	-0.06	-1.24	0.13
	T4	18	3.53	0.72	0.15	-0.96	0.17		T4	18	4.06	0.74	0.04	-1.50	0.18
Internal LOC								Overall Effectiveness							
External LOC^a								Internal LOC							
								T1							
								T2							
								T3							
								T4							

Note. THP = The Helmsman Project; SD = standard deviation; SE = standard error; T1, T2, T3 = extended baseline pre-test; T4 = immediate post-test; Self-Reg = Self-Regulation; SC = Self-Concept; Relationships = Relationships; Trust = Trustworthiness; LOC = Locus of Control.

^a All or some of the items in this scale are negatively worded and were reverse-scored prior to analysis.

APPENDIX T

BASELINE (T1) GROUP DIFFERENCES

Table T.1
Group Differences at Baseline (T1)

Scale	Baseline (T1) Differences: ES (SE)					
	Adv vs. C	Arctos vs. C	JC vs. C	OB vs. C	CO vs. C	Adv vs. CO
<i>High Relevance</i>						
Agency	.033 (.074)	.024 (.104)	.012 (.094)	.065 (.121)	-.086 (.079)	.120 (.086)
Pathways Thinking	.144 (.071)*	.130 (.112)	.105 (.116)	.196 (.094)*	-.032 (.045)	.176 (.064)
Goal Self-Regulation	-.002 (.061)	-.110 (.099)	.025 (.074)	.078 (.077)	.075 (.076)	-.077 (.077)
General Self-Esteem/SC	.058 (.067)	.123 (.106)	.016 (.074)	.035 (.090)	-.008 (.069)	.066 (.056)
Self-Confidence	.008 (.035)	.057 (.056)	-.102 (.072)	.069 (.043)	-.047 (.078)	.055 (.083)
Self-Efficacy	.044 (.051)	-.003 (.076)	.108 (.069)	.027 (.061)	-.076 (.068)	.120 (.063)
Open Thinking	-.019 (.033)	-.079 (.052)	.010 (.037)	.014 (.044)	-.032 (.092)	.014 (.086)
Cooperative Teamwork	-.087 (.087)	-.115 (.135)	-.018 (.114)	-.129 (.092)	-.052 (.157)	-.035 (.145)
Active Involvement	-.044 (.062)	-.023 (.101)	-.041 (.070)	-.069 (.065)	-.068 (.120)	.024 (.111)
<i>Mean for High Relevance</i>	<i>.015 (.022)</i>	<i>.000 (.031)</i>	<i>.013 (.029)</i>	<i>.032 (.025)</i>	<i>-.036 (.037)</i>	<i>.051 (.034)</i>
<i>Moderate Relevance</i>						
Optimism	-.100 (.094)	-.071 (.133)	-.187 (.109)	-.042 (.116)	-.284 (.102)**	.184 (.084)*
Emotional Self-Regulation	-.080 (.055)	-.167 (.096)	.081 (.071)	-.155 (.065)*	-.009 (.143)	-.071 (.148)
Booster Behaviors	-.089 (.100)	-.055 (.142)	-.079 (.124)	-.134 (.158)	-.039 (.181)	-.050 (.177)
Booster Thoughts	.037 (.067)	.044 (.085)	.017 (.102)	.050 (.084)	.053 (.145)	-.017 (.145)
Wellbeing	.055 (.063)	.026 (.074)	-.043 (.070)	.181 (.090)*	-.133 (.189)	.187 (.188)
Same-Sex Relationships SC	-.085 (.136)	.206 (.196)	-.063 (.136)	-.396 (.226)	-.022 (.136)	-.062 (.149)
Emotional Stability SC	-.155 (.099)	-.010 (.147)	-.096 (.111)	-.360 (.129)**	.000 (.217)	-.155 (.209)
Stress Management	.129 (.063)*	.083 (.087)	.211 (.088)*	.093 (.072)	.026 (.118)	.103 (.116)
Social Effectiveness	.056 (.097)	.082 (.143)	.030 (.103)	.057 (.098)	-.035 (.123)	.091 (.101)
Leadership Ability	-.088 (.098)	-.015 (.160)	-.178 (.139f)	-.070 (.123)	.013 (.117)	-.101 (.112)
Coping with Change	.026 (.059)	-.067 (.087)	.093 (.079)	.053 (.056)	-.012 (.068)	.038 (.064)
Internal LOC	-.027 (.053)	-.008 (.057)	-.080 (.075)	.008 (.085)	-.023 (.145)	-.004 (.023)
<i>Mean for Moderate Relevance</i>	<i>-.027 (.017)</i>	<i>.004 (.023)</i>	<i>-.025 (.020)</i>	<i>-.060 (.018)***</i>	<i>-.039 (.037)</i>	<i>.012 (.036)</i>
<i>Low Relevance</i>						
Pessimism	-.146 (.087)	-.049 (.120)	-.292 (.096)**	-.097 (.126)	.117 (.089)	-.263 (.084)**
Focus	-.040 (.094)	-.211 (.146)	.183 (.096)	-.092 (.121)	.061 (.107)	-.101 (.091)
Consistency of Interest	.032 (.097)	-.133 (.139)	.188 (.114)	.042 (.121)	.177 (.137)	-.145 (.124)
Perseverance of Effort	-.054 (.063)	-.001 (.080)	-.049 (.064)	-.110 (.098)	-.079 (.081)	.025 (.070)
Hampering	-.043 (.087)	.038 (.109)	-.082 (.129)	-.084 (.148)	-.027 (.083)	-.016 (.088)
Life Resilience	.158 (.076)*	.160 (.106)	.172 (.091)	.141 (.096)	.257 (.112)	-.099 (.104)
Academic Resilience	.120 (.076)	.192 (.130)	.121 (.106)	.047 (.077)	.328 (.148)*	-.208 (.145)
Life Satisfaction	-.025 (.089)	.053 (.129)	-.208 (.118)	.081 (.129)	-.095 (.200)	.070 (.201)
Physical Abilities SC	-.063 (.119)	-.056 (.168)	.120 (.103)	-.251 (.233)	-.021 (.182)	-.041 (.176)
Physical Appearance SC	-.018 (.142)	.070 (.168)	-.096 (.178)	-.028 (.232)	.095 (.181)	-.113 (.190)
Opposite-Sex Relationships SC	-.017 (.126)	.027 (.216)	.005 (.156)	-.085 (.142)	.055 (.236)	-.072 (.229)
Parent Relationships SC	-.046 (.099)	.001 (.164)	-.086 (.117)	-.053 (.185)	.107 (.164)	-.153 (.182)
Honesty-Trustworthiness SC	-.063 (.119)	-.104 (.156)	-.094 (.147)	.010 (.187)	.126 (.238)	-.189 (.234)
Math SC	.097 (.123)	-.021 (.149)	.087 (.184)	.225 (.157)	.208 (.152)	-.111 (.140)
Verbal SC	-.052 (.091)	-.114 (.135)	.146 (.136)	-.189 (.121)	.178 (.109)	-.231 (.097)*
School SC	.037 (.075)	-.005 (.095)	.079 (.105)	.039 (.096)	.126 (.116)	-.089 (.111)
Time Efficiency	.021 (.065)	-.037 (.080)	.079 (.105)	.021 (.069)	.037 (.085)	-.016 (.080)
Quality Seeking	-.023 (.050)	-.037 (.071)	-.035 (.056)	.004 (.085)	-.067 (.153)	.044 (.154)
Overall Effectiveness	-.003 (.048)	.000 (.064)	-.095 (.095)	.087 (.043)	-.137 (.058)	.134 (.064)*
External LOC	-.075 (.116)	.107 (.153)	-.209 (.156)	-.123 (.175)	.050 (.216)	-.124 (.211)
<i>Mean for Low Relevance</i>	<i>-.010 (.021)</i>	<i>-.006 (.029)</i>	<i>-.003 (.027)</i>	<i>-.021 (.029)</i>	<i>.075 (.036)</i>	<i>-.085 (.035)*</i>

Note. T1 = pre-test; ES = standardised effect size; SE = standard error; Adv = Adventure Programs (as a whole); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; C = waitlist control group. SC = Self-Concept. High, Moderate, and Low Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. The Mean for each of the High, Moderate, and Low Relevance scales represents the mean effect for that group of scales. Significant differences are in bold: * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

APPENDIX U

WAITLIST CONTROL GROUP EXTENDED BASELINE COMPARISONS

Table U.1
Waitlist Control Group Within-Subjects Extended Baseline Differences (T1, T2, T3)

Scale	Waitlist Control Group Extended Baseline Comparisons: ES (SE)														
	Adventure Programs (together)			Arctos Adventure Program			James Craig Adventure Program			Outward Bound Adventure Program			Coaching Only Program		
	T1/T2	T1/T3	T2/T3	T1/T2	T1/T3	T2/T3	T1/T2	T1/T3	T2/T3	T1/T2	T1/T3	T2/T3	T1/T2	T1/T3	T2/T3
<i>High Relevance</i>															
Agency	.056 (.085)	.068 (.078)	.012 (.069)	-.167 (.123)	-.017 (.133)	.150 (.103)	.061 (.139)	-.011 (.119)	-.072 (.119)	.273 (.119)*	.232 (.110)*	-.041 (.127)	-.100 (.262)	-.108 (.175)	-.008 (.244)
Pathways Thinking	-.016 (.085)	.059 (.074)	.075 (.071)	-.139 (.103)	-.004 (.129)	.135 (.075)	-.057 (.185)	.07 (.136)	.127 (.168)	.146 (.125)	.11 (.084)	-.037 (.109)	-.434 (.185)*	-.075 (.175)	.359 (.197)
Goal Self-Regulation	-.029 (.088)	-.039 (.094)	-.010 (.061)	-.040 (.112)	-.029 (.116)	.011 (.092)	-.093 (.169)	-.081 (.153)	.012 (.135)	.045 (.125)	-.006 (.124)	-.051 (.077)	-.081 (.164)	-.046 (.132)	.035 (.078)
General Self-Esteem/SC	.070 (.066)	-.012 (.083)	-.082 (.062)	.005 (.111)	.051 (.112)	.047 (.090)	-.014 (.115)	-.103 (.158)	-.089 (.134)	.220 (.103)*	.016 (.100)	-.204 (.078)*	-.110 (.136)	.008 (.153)	.119 (.148)
Self-Confidence	-.062 (.083)	-.029 (.076)	.034 (.062)	-.222 (.117)	-.097 (.140)	.125 (.108)	-.067 (.140)	.009 (.102)	.076 (.120)	.102 (.104)	.002 (.100)	-.100 (.088)	-.194 (.179)	-.009 (.161)	.185 (.117)
Self-Efficacy	-.035 (.088)	-.017 (.092)	.018 (.053)	-.131 (.119)	-.152 (.123)	-.021 (.104)	-.108 (.165)	-.062 (.15)	.046 (.086)	.133 (.117)	.162 (.136)	.029 (.082)	.145 (.198)	.273 (.187)	.128 (.108)
Open Thinking	-.043 (.084)	-.078 (.078)	-.035 (.070)	-.053 (.135)	-.134 (.136)	-.080 (.111)	-.186 (.147)	-.283 (.127)*	-.097 (.159)	.110 (.109)	.182 (.093)*	.072 (.069)	-.005 (.143)	-.126 (.148)	-.122 (.096)
Cooperative Teamwork	-.093 (.069)	-.101 (.079)	-.008 (.068)	-.104 (.124)	-.245 (.132)	-.14 (.098)	-.222 (.120)	-.178 (.146)	.044 (.134)	.048 (.100)	.119 (.093)	.071 (.098)	.011 (.213)	-.075 (.153)	-.086 (.157)
Active Involvement	-.138 (.081)	-.113 (.071)	.025 (.076)	-.132 (.115)	-.203 (.121)	-.071 (.11)	-.271 (.168)	-.194 (.142)	.077 (.166)	-.011 (.089)	.059 (.079)	.070 (.090)	-.012 (.176)	-.077 (.125)	-.065 (.120)
<i>Moderate Relevance</i>															
Optimism	-.011 (.074)	-.109 (.083)	-.098 (.053)	-.024 (.116)	-.038 (.127)	-.013 (.094)	.028 (.106)	-.008 (.096)	-.036 (.100)	-.037 (.122)	-.283 (.134)*	-.245 (.093)*	-.230 (.119)	.056 (.135)	.285 (.113)**
Emotion Self-Regulation	.010 (.085)	.021 (.095)	.011 (.064)	-.109 (.111)	-.142 (.149)	-.033 (.108)	.081 (.126)	.088 (.128)	.007 (.131)	.058 (.126)	.118 (.121)	.060 (.085)	.259 (.191)	.406 (.173)**	.147 (.147)
Booster Behaviours	-.076 (.088)	-.050 (.098)	.026 (.056)	-.167 (.099)	-.048 (.125)	.119 (.096)	.052 (.182)	.033 (.168)	-.019 (.105)	-.112 (.151)	-.134 (.167)	-.022 (.082)	-.236 (.217)	-.160 (.170)	.076 (.142)
Booster Thoughts	-.058 (.080)	-.035 (.087)	.024 (.070)	-.103 (.112)	-.016 (.123)	.087 (.087)	-.154 (.164)	.005 (.150)	.160 (.116)	.083 (.111)	-.093 (.105)	-.176 (.133)	-.195 (.171)	-.230 (.170)	-.035 (.113)
Wellbeing	.098 (.092)	-.010 (.087)	-.109 (.067)	.160 (.153)	.028 (.125)	-.132 (.105)	.110 (.135)	-.053 (.139)	-.163 (.136)	.026 (.114)	-.006 (.140)	-.031 (.099)	-.345 (.223)	-.014 (.144)	.332 (.198)
Same-Sex Relationships SC	-.075 (.068)	-.028 (.070)	.047 (.065)	-.231 (.119)	-.096 (.108)	.135 (.101)	.079 (.124)	.033 (.147)	-.045 (.125)	-.073 (.092)	-.020 (.093)	.053 (.100)	-.062 (.269)	.221 (.144)	.283 (.201)
Emotional Stability SC	.076 (.084)	.057 (.103)	-.019 (.092)	-.024 (.122)	.079 (.136)	.103 (.102)	.151 (.155)	.162 (.161)	.012 (.146)	.100 (.150)	-.072 (.191)	-.172 (.152)	-.285 (.203)	-.281 (.142)*	.004 (.126)
Stress Management	.068 (.099)	.041 (.103)	-.027 (.069)	.065 (.131)	-.152 (.148)	-.217 (.107)*	.100 (.181)	.104 (.138)	.004 (.142)	.039 (.130)	.170 (.153)	.131 (.094)	.383 (.240)	.568 (.209)**	.185 (.145)
Social Effectiveness	.029 (.112)	.158 (.115)	.129 (.066)	-.074 (.121)	.096 (.141)	.171 (.110)	.134 (.235)	.327 (.228)	.193 (.104)	.026 (.187)	.050 (.173)	.024 (.117)	.218 (.187)	.223 (.169)	.006 (.123)
Leadership Ability	.009 (.063)	.039 (.067)	.030 (.060)	.049 (.102)	.057 (.113)	.008 (.095)	-.114 (.091)	-.057 (.097)	.057 (.111)	.092 (.110)	.116 (.114)	.024 (.085)	-.270 (.161)	-.103 (.144)	.168 (.108)
Coping with Change	.018 (.127)	.042 (.106)	.024 (.081)	.118 (.132)	-.024 (.153)	-.142 (.131)	-.061 (.244)	.045 (.168)	.106 (.156)	-.003 (.170)	.105 (.131)	.109 (.108)	.579 (.192)**	.375 (.169)*	-.203 (.106)
Internal Locus of Control	-.063 (.078)	-.092 (.070)	-.029 (.067)	-.195 (.124)	-.157 (.138)	.037 (.123)	-.172 (.143)	-.131 (.110)	.040 (.131)	.179 (.098)	.014 (.094)	-.165 (.081)*	-.247 (.151)	-.311 (.155)*	-.064 (.110)

Note. ES = standardized effect sizes; SE = standard error; T1, T2, T3 = three extended baseline data waves in that order, collected approximately three months apart; SC = Self-Concept. High and Moderate Relevance indicate the level of relevance of the scales to the THP program design and aims as rated by three raters. Significant effects are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

continues

Table U.1 (continued)
 Waitlist Control Group Within-Subjects Extended Baseline Differences (T1, T2, T3)

Scale	Waitlist Control Group Extended Baseline Comparisons: ES (SE)														
	Adventure Programs (together)			Arctos Adventure Program			James Craig Adventure Program			Outward Bound Adventure Program			Coaching Only Program		
	T1/T2	T1/T3	T2/T3	T1/T2	T1/T3	T2/T3	T1/T2	T1/T3	T2/T3	T1/T2	T1/T3	T2/T3	T1/T2	T1/T3	T2/T3
<i>Low Relevance</i>															
Pessimism ^a	.021 (.070)	.083 (.090)	.062 (.071)	.051 (.115)	.180 (.117)	.129 (.101)	.139 (.106)	.267 (.182)	.128 (.142)	-.127 (.140)	-.198 (.167)	-.071 (.106)	-.037 (.173)	.180 (.180)	.217 (.180)
Focus	-.068 (.085)	-.001 (.098)	.066 (.063)	-.047 (.126)	-.038 (.139)	.009 (.110)	-.088 (.124)	-.008 (.128)	.080 (.122)	-.069 (.122)	.041 (.134)	.110 (.074)	.131 (.236)	.161 (.200)	.029 (.112)
Consistency of Interest	.006 (.080)	.172 (.066)**	.166 (.069)**	-.002 (.121)	.301 (.112)**	.303 (.106)**	.029 (.162)	.161 (.092)	.132 (.151)	-.009 (.133)	.054 (.140)	.063 (.092)	.022 (.195)	.001 (.196)	-.021 (.173)
Perseverance of Effort	-.035 (.074)	-.015 (.076)	.020 (.062)	-.203 (.104)	.003 (.106)	.205 (.112)	-.017 (.132)	-.167 (.116)	-.151 (.121)	.114 (.113)	.121 (.123)	.007 (.077)	-.428 (.153)**	-.053 (.141)	.374 (.111)***
Hampering ^a	.021 (.078)	.142 (.090)	.120 (.062)	-.084 (.137)	.038 (.145)	.123 (.106)	.113 (.133)	.319 (.158)*	.206 (.125)	.036 (.132)	.068 (.140)	.032 (.080)	-.247 (.196)	-.143 (.158)	.104 (.127)
Life Resilience	.213 (.109)*	.267 (.121)*	.054 (.068)	.086 (.139)	.103 (.168)	.017 (.118)	.370 (.232)	.507 (.236)*	.137 (.133)	.184 (.142)	.191 (.149)	.007 (.084)	.312 (.235)	.543 (.215)*	.231 (.148)
Academic Resilience	.159 (.095)	.220 (.095)*	.061 (.069)	-.033 (.121)	.133 (.134)	.166 (.106)	.289 (.173)	.274 (.140)	-.015 (.113)	.221 (.167)	.254 (.163)	.033 (.112)	-.009 (.196)	.250 (.137)	.258 (.168)
Life Satisfaction	.042 (.073)	.076 (.074)	.034 (.071)	.081 (.108)	.019 (.108)	-.062 (.099)	-.002 (.122)	.059 (.139)	.062 (.170)	.048 (.132)	.150 (.120)	.102 (.086)	-.163 (.210)	-.063 (.101)	.101 (.201)
Physical Abilities SC	-.081 (.064)	-.040 (.064)	.040 (.053)	-.083 (.087)	-.029 (.091)	.055 (.085)	-.164 (.137)	-.107 (.125)	.057 (.119)	.006 (.091)	.016 (.089)	.010 (.053)	.067 (.152)	-.025 (.152)	-.092 (.104)
Physical Appearance SC	.114 (.084)	.276 (.087)***	.162 (.071)**	.188 (.155)	.387 (.136)**	.199 (.105)	.073 (.121)	.183 (.135)	.109 (.112)	.081 (.144)	.258 (.159)	.177 (.118)	.147 (.241)	.232 (.197)	.085 (.172)
Opposite-Sex Relationships SC	.050 (.062)	.102 (.073)	.052 (.059)	.140 (.114)	.274 (.102)**	.134 (.104)	.039 (.095)	.070 (.154)	.031 (.120)	-.030 (.115)	-.039 (.114)	-.009 (.068)	.190 (.179)	.086 (.143)	-.103 (.130)
Parent Relationships SC	-.075 (.067)	-.064 (.064)	.012 (.058)	-.063 (.117)	.010 (.101)	.073 (.102)	-.307 (.103)**	-.165 (.091)	.142 (.120)	.144 (.108)	-.036 (.113)	-.179 (.065)**	.101 (.190)	.195 (.148)	.094 (.170)
Honesty-Trustworthiness SC	.059 (.063)	.069 (.065)	.009 (.058)	.043 (.106)	.122 (.097)	.080 (.082)	.226 (.104)*	.189 (.115)	-.037 (.113)	-.091 (.095)	-.106 (.120)	-.014 (.093)	.011 (.168)	.203 (.133)	.192 (.178)
Math SC	-.042 (.064)	-.059 (.064)	-.017 (.045)	-.115 (.105)	-.095 (.092)	.020 (.077)	.042 (.104)	-.009 (.104)	-.051 (.084)	-.054 (.091)	-.075 (.104)	-.021 (.074)	-.004 (.156)	-.038 (.122)	-.034 (.080)
Verbal SC	.027 (.061)	.049 (.070)	.022 (.063)	.120 (.095)	.101 (.106)	-.020 (.100)	.085 (.101)	.153 (.099)	.068 (.109)	-.125 (.109)	-.108 (.122)	.017 (.103)	.201 (.141)	.185 (.114)	-.016 (.097)
School SC	-.036 (.069)	-.045 (.085)	-.009 (.059)	-.062 (.106)	-.018 (.107)	.043 (.090)	-.056 (.108)	-.066 (.141)	-.011 (.116)	.008 (.117)	-.052 (.114)	-.060 (.074)	.012 (.129)	.096 (.134)	.085 (.085)
Time Efficiency	-.077 (.079)	-.030 (.089)	.047 (.064)	-.153 (.111)	-.047 (.102)	.106 (.092)	.113 (.149)	.016 (.165)	-.098 (.126)	-.191 (.118)	-.057 (.119)	.134 (.094)	.013 (.242)	.265 (.165)	.251 (.185)
Quality Seeking	-.157 (.084)	-.187 (.080)*	-.030 (.068)	-.289 (.127)*	-.265 (.138)	.024 (.130)	-.149 (.138)	-.125 (.122)	.023 (.117)	-.033 (.105)	-.171 (.104)	-.138 (.093)	-.260 (.168)	-.205 (.156)	.055 (.095)
Overall Effectiveness	-.028 (.079)	-.046 (.079)	-.018 (.064)	-.162 (.144)	-.105 (.151)	.056 (.116)	.040 (.130)	-.053 (.119)	-.093 (.120)	.037 (.088)	.019 (.092)	-.018 (.078)	-.122 (.161)	.101 (.152)	.223 (.112)*
External Locus of Control ^a	.040 (.080)	.172 (.082)*	.132 (.069)	-.035 (.127)	.081 (.149)	.116 (.101)	.184 (.170)	.389 (.162)*	.205 (.135)	-.030 (.117)	.046 (.105)	.076 (.112)	-.472 (.293)	.114 (.149)	.586 (.274)*

Note. ES = standardized effect sizes; SE = standard error; T1, T2, T3 = three extended baseline data waves in that order, collected approximately three months apart; SC = Self-Concept. Low Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. Significant effects are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

APPENDIX V

INTERVENTION GROUP WITHIN-SUBJECTS PRE-POST ANALYSIS

Table V.1
Results of Within-Subjects Pre-Post Analysis for Intervention Participants

Scale	Intervention Group Within-Subjects Pre-Post Program Effects: ES (SE)					
	Adventure	Arctos	James Craig	Outward Bound	Coaching Only	Adv vs. CO
<i>High Relevance</i>						
Agency	.194 (.133)	.138 (.119)	.229 (.350)	.214 (.148)	-.101 (.144)	.295 (.187)
Pathways Thinking	.170 (.097)	.014 (.115)	.482 (.189)*	.015 (.151)	-.004 (.166)	.174 (.177)
Goal Self-Regulation	.317 (.094)***	.168 (.121)	.691 (.223)**	.092 (.111)	-.012 (.207)	.330 (.227)
General Self-Esteem/SC	.278 (.098)**	-.025 (.097)	.640 (.201)***	.219 (.138)	.236 (.147)	.042 (.156)
Self-Confidence	.320 (.082)***	.033 (.102)	.807 (.141)***	.120 (.135)	.246 (.113)*	.074 (.121)
Self-Efficacy	.391 (.080)***	.204 (.109)	.660 (.159)***	.307 (.115)**	.310 (.151)*	.081 (.156)
Open Thinking	.349 (.100)***	.057 (.110)	.663 (.230)**	.326 (.135)*	-.110 (.169)	.458 (.191)*
Cooperative Teamwork	.158 (.104)	.166 (.135)	.102 (.208)	.207 (.177)	-.076 (.154)	.234 (.185)
Active Involvement	.271 (.092)**	.016 (.091)	.492 (.216)*	.306 (.125)*	-.038 (.139)	.309 (.164)
<i>Moderate Relevance</i>						
Optimism	.113 (.098)	.018 (.093)	.269 (.225)	.052 (.155)	.347 (.110)**	-.234 (.146)
Emotion Self-Regulation	.270 (.095)**	.127 (.114)	.648 (.221)**	.035 (.124)	-.021 (.176)	.292 (.197)
Booster Behaviours	.144 (.091)	-.043 (.104)	.360 (.200)	.114 (.137)	.043 (.195)	.101 (.213)
Booster Thoughts	.088 (.069)	-.056 (.102)	.220 (.120)	.100 (.127)	-.174 (.197)	.262 (.205)
Wellbeing	.432 (.129)***	.244 (.121)*	.867 (.303)**	.187 (.191)	.355 (.216)	.078 (.239)
Same-Sex Relationships SC	-.038 (.089)	-.163 (.100)	.035 (.189)	.016 (.158)	.141 (.238)	-.179 (.262)
Emotional Stability SC	.126 (.084)	.126 (.120)	.244 (.193)	.008 (.125)	.332 (.153)*	-.206 (.182)
Stress Management	.388 (.090)***	.217 (.105)*	.652 (.222)**	.295 (.117)*	.459 (.212)*	-.071 (.228)
Social Effectiveness	.257 (.111)*	.023 (.098)	.530 (.262)*	.220 (.168)	.222 (.120)	.035 (.157)
Leadership Ability	.205 (.093)*	-.056 (.087)	.613 (.206)**	.059 (.163)	.176 (.114)	.029 (.144)
Coping with Change	.269 (.116)*	.144 (.104)	.378 (.290)	.286 (.148)	.218 (.183)	.051 (.214)
Internal Locus of Control	.235 (.092)*	-.068 (.115)	.617 (.186)***	.155 (.158)	-.250 (.154)	.485 (.179)**
<i>Low Relevance</i>						
Pessimism ^a	.156 (.088)	.107 (.118)	.161 (.160)	.200 (.145)	-.182 (.218)	.339 (.220)
Focus	.330 (.085)***	.171 (.110)	.672 (.200)***	.147 (.100)	.006 (.163)	.324 (.176)
Consistency of Interest	-.020 (.084)	.082 (.113)	.010 (.187)	-.151 (.133)	-.515 (.145)***	.496 (.155)***
Perseverance of Effort	.051 (.097)	.028 (.100)	.394 (.221)	-.269 (.147)	-.211 (.172)	.263 (.195)
Hampering ^a	.113 (.078)	-.045 (.095)	.139 (.142)	.245 (.133)	-.244 (.203)	.357 (.211)
Life Resilience	.337 (.084)***	.191 (.096)	.724 (.197)***	.096 (.121)	.136 (.139)	.201 (.160)
Academic Resilience	.366 (.085)***	.169 (.097)	.488 (.172)**	.442 (.103)***	-.158 (.172)	.524 (.178)**
Life Satisfaction	.181 (.124)	.108 (.116)	.493 (.298)	-.057 (.180)	.371 (.134)**	-.190 (.176)
Physical Abilities SC	.163 (.068)*	.205 (.077)**	.236 (.166)	.049 (.085)	.129 (.143)	.034 (.154)
Physical Appearance SC	.157 (.099)	.086 (.099)	.347 (.226)	.039 (.167)	.432 (.222)	-.275 (.237)
Opposite-Sex Relationships SC	.146 (.079)	.184 (.076)*	.329 (.167)*	-.076 (.151)	.243 (.127)	-.097 (.151)
Parent Relationships SC	.021 (.077)	-.017 (.088)	.177 (.170)	-.097 (.123)	.227 (.144)	-.206 (.163)
Honesty-Trustworthiness SC	.089 (.104)	.218 (.089)*	.029 (.270)	.020 (.126)	.137 (.208)	-.048 (.233)
Math SC	-.067 (.095)	.038 (.097)	-.276 (.236)	.035 (.136)	.031 (.152)	-.099 (.180)
Verbal SC	.152 (.083)	.219 (.110)*	.230 (.195)	.006 (.112)	.194 (.106)	-.042 (.128)
School SC	.164 (.093)	.064 (.103)	.340 (.198)	.088 (.109)	.122 (.118)	.042 (.134)
Time Efficiency	.164 (.119)	.119 (.107)	.352 (.300)	.022 (.137)	.218 (.164)	-.053 (.203)
Quality Seeking	.147 (.085)	-.013 (.118)	.413 (.168)*	.041 (.130)	-.106 (.189)	.253 (.204)
Overall Effectiveness	.364 (.109)***	.101 (.111)	.880 (.261)***	.112 (.133)	.475 (.122)***	-.111 (.152)
External Locus of Control ^a	-.076 (.074)	-.167 (.105)	.004 (.120)	-.065 (.136)	-.287 (.137)*	.211 (.146)

Note. ES = standardized effect sizes; SE = standard error; Adventure/Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; James Craig = James Craig Adventure Program; Outward Bound = Outward Bound Adventure Program; and Coaching Only/CO = Coaching Only Program; SC = Self-Concept. High, Moderate, and Low Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. Significant effects are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for these scales were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

APPENDIX W
MULTIVARIATE OMNIBUS TESTS OF HIGH RELEVANCE SCALES

Table W.1

Results of Multivariate Omnibus Tests of Intervention Effects for High Relevance Scales using Wald Test

Model	Wald statistic	df	p-value
Short-Term Analysis (T1-T2)	311.33	36	.0000
Long-Term Analysis (T1-T3)	460.478	36	.0000
Follow-Up Analysis (T2-T3)	221.107	36	.0000

Note. *df* = degrees of freedom; T1 = pre-test; T2 = immediate post-test; and T3 = three months post-test.

APPENDIX X

GENDER-TREATMENT INTERACTIONS: SIGNIFICANT RESULTS

Table X.1
Gender-Treatment Interactions for Short-Term and Long-Term Analyses: Significant Results

Scale	Parameter	Gender-Treatment Interaction Effects (Gender x Program Group): ES (SE)									
		Adv vs. C		Arctos vs. C		James Craig vs. C		Outward Bound vs. C		Coaching Only vs. C	
		T2	T3	T2	T3	T2	T3	T2	T3	T2	T3
<i>High Relevance</i>											
Agency	b1			.343 (.136)*	.242 (.128)						
	b2			.241 (.121)*	.138 (.104)						
	b3			-.807 (.166)***	-.915 (.161)***						
	b1 + b3			-.464 (.093)**	-.673 (.312)*						
Pathways Thinking	b1			.198 (.148)	.206 (.082)*						
	b2			.040 (.113)	.165 (.098)						
	b3			-.511 (.188)**	-.804 (.126)***						
	b1 + b3			-.313 (.121)**	-.597 (.295)*						
Goal Self-Regulation	b1			.249 (.127)	.199 (.125)	.860 (.182)***			.214 (.202)	.387 (.150)**	
	b2			.195 (.133)	.186 (.118)	.195 (.133)			.195 (.133)	.186 (.118)	
	b3			-.534 (.176)**	-.730 (.168)***	-.723 (.244)**			-.622 (.252)*	-.523 (.204)**	
	b1 + b3			-.285 (.133)*	-.531 (.320)	.138 (.158)			-.409 (.135)**	-.137 (.143)	
General Self-Esteem	b1			.583 (.314)	.278 (.156)			.373 (.170)*	.347 (.129)**	.204 (.149)	
	b2			.290 (.126)*	.222 (.120)			.290 (.126)*	.290 (.126)*	.222 (.120)	
	b3			-.513 (.231)*	-.756 (.190)***			-.879 (.341)**	-.780 (.236)***	-.473 (.217)*	
	b1 + b3			-.394 (.126)**	-.478 (.324)			-.507 (.294)	-.433 (.187)*	-.270 (.149)	
Self Confidence	b1			.222 (.178)	.075 (.113)	.893 (.173)***			.329 (.161)*		
	b2			.196 (.137)	.020 (.112)	.196 (.137)			.196 (.137)		
	b3			-.623 (.233)**	-.370 (.159)*	-.532 (.239)*			-.467 (.217)*		
	b1 + b3			-.400 (.141)**	-.295 (.127)*	.361 (.142)*			-.139 (.140)		
Self Efficacy	b1					.858 (.144)***			.545 (.255)*		
	b2					.254 (.128)*			.254 (.128)*		
	b3					-.667 (.207)***			-.831 (.298)**		
	b1 + b3					.191 (.137)			-.285 (.138)*		

Note. ES = standardized effect size; SE = standard error. Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. T2 = significant gender-treatment interaction effects for the Short-Term Analysis (pre-test to immediate post-test) and T3 = significant gender-treatment interaction effects for the Long-Term Analysis (pre-test to three months post-test). High Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. Only parameter coefficients for significant interaction effects are shown: b1 = main effect coefficient for program group; b2 = main effect coefficient for gender (reflecting effect for females when compared to control group); b3 = gender-treatment interaction effect coefficient; b1 + b3 = effect for males when compared to control group. Significant parameters are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

continues

Table X.1 (continued)
Gender-Treatment Interactions for Short-Term and Long-Term Analyses: Significant Results

Scale	Parameter	Gender-Treatment Interaction Effects (Gender x Program Group): ES (SE)									
		Adv vs. C		Arctos vs. C		James Craig vs. C		Outward Bound vs. C		Coaching Only vs. C	
		T2	T3	T2	T3	T2	T3	T2	T3	T2	T3
<i>High Relevance</i>											
Open Thinking	b1			.142 (.153)	.124 (.102)	.792 (.148)***					
	b2			.169 (.143)	.177 (.134)	.169 (.143)					
	b3			-.533 (.198)**	-.851 (.171)***	-.541 (.249)*					
	b1 + b3			-.391 (.147)**	-.727 (.325)*	.251 (.185)					
Cooperative Teamwork	b1			.350 (.112)**	.194 (.103)						
	b2			.205 (.140)	.108 (.109)						
	b3			-.741 (.185)***	-.674 (.152)***						
	b1 + b3			-.391 (.147)**	-.480 (.310)						
Active Involvement	b1			.231 (.139)	.076 (.081)		.672 (.194)***				
	b2			.149 (.126)	.163 (.116)		.163 (.115)				
	b3			-.909 (.188)***	-.781 (.128)***		-.535 (.248)*				
	b1 + b3			-.679 (.124)***	-.705 (.293)*		.137 (.159)				
<i>Moderate Relevance</i>											
Optimism	b1				-.001 (.108)					.279 (.140)*	
	b2				.049 (.091)					-.066 (.116)	
	b3				-.361 (.132)**					-.410 (.150)**	
	b1 + b3				-.361 (.083)***					-.131 (.083)	
Emotion Self-Regulation	b1				-.040 (.162)						
	b2				.144 (.144)						
	b3				-.447 (.215)*						
	b1 + b3				-.487 (.105)***						
Booster Behaviours	b1				.191 (.119)					.567 (.256)*	.381 (.235)
	b2				.124 (.131)					.187 (.162)	.124 (.131)
	b3				-.372 (.165)*					-1.056 (.309)**	-.624 (.271)*
	b1 + b3				-.181 (.128)					-.489 (.164)**	-.243 (.145)

Note. ES = standardized effect size; SE = standard error. Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. T2 = significant gender-treatment interaction effects for the Short-Term Analysis (pre-test to immediate post-test) and T3 = significant gender-treatment interaction effects for the Long-Term Analysis (pre-test to three months post-test). High Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. Only parameter coefficients for significant interaction effects are shown: b1 = main effect coefficient for program group; b2 = main effect coefficient for gender (reflecting effect for females when compared to control group); b3 = gender-treatment interaction effect coefficient; b1 + b3 = effect for males when compared to control group. Significant parameters are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

continues

Table X.1 (continued)

Gender-Treatment Interactions for Short-Term and Long-Term Analyses: Significant Results

Scale	Parameter	Gender-Treatment Interaction Effects (Gender x Program Group): ES (SE)									
		Adv vs. C		Arctos vs. C		James Craig vs. C		Outward Bound vs. C		Coaching Only vs. C	
		T2	T3	T2	T3	T2	T3	T2	T3	T2	T3
<i>Moderate Relevance</i>											
Booster Thoughts	b1			.217 (.125)							
	b2			.194 (.129)							
	b3			-.872 (.161)***							
	b1 + b3			-.655 (.106)***							
Wellbeing	b1			.147 (.146)							.310 (.123)*
	b2			.189 (.120)							.189 (.120)
	b3			-.425 (.187)*							-.465 (.231)*
	b1 + b3			-.278 (.107)**							-155 (.194)
Same Sex Rel'ships SC	b1			-.043 (.129)			.502 (.095)***		-.135 (.199)		
	b2			-.062 (.136)			.220 (.103)*		-.062 (.136)		
	b3			-.532 (.194)**			-.474 (.242)*		.636 (.244)**		
	b1 + b3			-.575 (.147)***			.028 (.216)		.501 (.140)***		
Emotional Stability	b1								-.294 (.158)		-.457 (.167)**
	b2								-.162 (.091)		-.217 (.125)
	b3								.581 (.229)*		.863 (.237)***
	b1 + b3								.287 (.161)		.406 (.157)**
Stress Management	b1										.525 (.178)**
	b2										.166 (.154)
	b3										-1.055 (.218)***
	b1 + b3										-.530 (.132)***
Social Effectiveness	b1			.172 (.115)							
	b2			.098 (.132)							
	b3			-.921 (.155)***							
	b1 + b3			-.749 (.101)***							
Internal Locus of Control	b1			.140 (.177)		.142 (.122)					
	b2			.119 (.167)		.069 (.122)					
	b3			-.894 (.214)***		-.852 (.164)***					
	b1 + b3			-.754 (.105)***		-.710 (.107)***					

Note. ES = standardised effect size; SE = standard error. Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. T2 = significant gender-treatment interaction effects for the Short-Term Analysis (pre-test to immediate post-test) and T3 = significant gender-treatment interaction effects for the Long-Term Analysis (pre-test to three months post-test). Moderate Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. Only parameter coefficients for significant interaction effects are shown: b1 = main effect coefficient for program group; b2 = main effect coefficient for gender (reflecting effect for females when compared to control group); b3 = gender-treatment interaction effect coefficient; b1 + b3 = effect for males when compared to control group. Significant parameters are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

continues

Table X.1 (continued)
Gender-Treatment Interactions for Short-Term and Long-Term Analyses: Significant Results

Scale	Parameter	Gender-Treatment Interaction Effects (Gender x Program Group): ES (SE)									
		Adv vs. C		Arctos vs. C		James Craig vs. C		Outward Bound vs. C		Coaching Only vs. C	
		T2	T3	T2	T3	T2	T3	T2	T3	T2	T3
<i>Low Relevance</i>											
Pessimism ^a	b1			.030 (.111)							
	b2			-.107 (.111)							
	b3			.312 (.150)*							
	b1 + b3			.343 (.096)***							
Focus	b1					.949 (.111)***				.326 (.168)	.407 (.097)***
	b2					.024 (.106)				.024 (.106)	.264 (.127)*
	b3					-.728 (.180)***				-1.046 (.442)*	-.958 (.160)***
	b1 + b3					.220 (.136)				-.720 (.417)	-.551 (.224)*
Consistency of Interest	b1			-.054 (.127)	-.062 (.113)	.497 (.212)*					
	b2			-.179 (.118)	-.008 (.118)	-.179 (.118)					
	b3			.541 (.162)***	.890 (.161)**	-.556 (.267)*					
	b1 + b3			.487 (.108)***	.828 (.291)**	-.059 (.167)					
Perseverance of Effort	b1				.113 (.109)					-.191 (.081)*	
	b2				.078 (.123)					.000 (.117)	
	b3				-.410 (.148)**					-.382 (.122)**	
	b1 + b3				-.297 (.093)***					-.573 (.116)***	
Academic Resilience	b1									.083 (.120)	.132 (.160)
	b2									-.055 (.137)	.193 (.135)
	b3									-.844 (.394)*	-.578 (.262)*
	b1 + b3									-.761 (.379)*	-.445 (.201)*
Life Satisfaction	b1			.149 (.158)							.150 (.140)
	b2			.083 (.112)							.021 (.113)
	b3			-.415 (.197)*							-.433 (.220)*
	b1 + b3			-.266 (.123)*							-.283 (.174)
Physical Appearance SC	b1					.504 (.256)*			.113 (.142)		.468 (.153)**
	b2					.322 (.141)*			.322 (.141)*		.322 (.141)*
	b3					-.587 (.276)*			-.387 (.159)*		-.710 (.178)***
	b1 + b3					-.083 (.154)			-.275 (.101)**		-.241 (.089)**

Note. ES = standardised effect size; SE = standard error. Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. T2 = significant gender-treatment interaction effects for the Short-Term Analysis (pre-test to immediate post-test) and T3 = significant gender-treatment interaction effects for the Long-Term Analysis (pre-test to three months post-test). Low Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. Only parameter coefficients for significant interaction effects are shown: b1 = main effect coefficient for program group; b2 = main effect coefficient for gender (reflecting effect for females when compared to control group); b3 = gender-treatment interaction effect coefficient; b1 + b3 = effect for males when compared to control group; SC = Self-Concept. Significant parameters are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a The items for this scale were reverse-scored so that positive effects indicate a decrease in the relevant outcome variable.

continues

Table X.1 (continued)

Gender-Treatment Interactions for Short-Term and Long-Term Analyses: Significant Results

Scale	Parameter	Gender-Treatment Interaction Effects (Gender x Program Group): ES (SE)									
		Adv vs. C		Arctos vs. C		James Craig vs. C		Outward Bound vs. C		Coaching Only vs. C	
		T2	T3	T2	T3	T2	T3	T2	T3	T2	T3
<i>Low Relevance</i>											
Opposite-Sex Relationships SC	b1										
	b2										
	b3										
	b1 + b3										
Parent Relationships SC	b1										
	b2										
	b3										
	b1 + b3										
Hon/Trustworthiness SC	b1										
	b2										
	b3										
	b1 + b3										
Math SC	b1										
	b2										
	b3										
	b1 + b3										
School SC	b1										
	b2										
	b3										
	b1 + b3										
Time Efficiency	b1										
	b2										
	b3										
	b1 + b3										
Quality Seeking	b1										
	b2										
	b3										
	b1 + b3										
Overall Effectiveness	b1										
	b2										
	b3										
	b1 + b3										

Note. ES = standardised effect size; SE = standard error. Adv = Adventure Programs (taken together); Arctos = Arctos Adventure Program; JC = James Craig Adventure Program; OB = Outward Bound Adventure Program; CO = Coaching Only Program; and C = waitlist control group. T2 = significant gender-treatment interaction effects for the Short-Term Analysis (pre-test to immediate post-test) and T3 = significant gender-treatment interaction effects for the Long-Term Analysis (pre-test to three months post-test). Low Relevance indicates the level of relevance of the scales to the THP program design and aims as rated by three raters. Only parameter coefficients for significant interaction effects are shown: b1 = main effect coefficient for program group; b2 = main effect coefficient for gender (reflecting effect for females when compared to control group); b3 = gender-treatment interaction effect coefficient; b1 + b3 = effect for males when compared to control group; SC = Self-Concept. Significant parameters are bold for ease of reference. * $p < .05$; ** $p < .01$; *** $p < .001$.

APPENDIX Y SUBJECT-OBJECT INTERVIEW PROMPTS

Category A

1. PROUD, SUCCESSFUL

If you were to think back over your time in the Program, and you had to think about times you felt proud or successful, for example, because you had achieved something that was difficult for you, are there one or two things that come to mind? Take a minute to think about it, if you like, and just jot down on the card whatever you need to remind you of what they were.

2. MOVED, TOUCHED

If you think back over your time in the Program, and you had to think about times you felt quite touched by something you saw, or thought or heard, perhaps something that moved you, are there one or two moments that come to mind? Take a minute to think about it, if you like, and just jot down on the card whatever you need to remind you of what they were.

3. CHANGE

As you look back at your past, if you had to think of some ways in which you think you've changed since participating in the Program, are there some ways that come to mind? For example, you might feel changed in your attitude towards school or your schoolwork, or you might feel different in how you are at home, or you might notice that there are things you do now that you couldn't or didn't do before. You might notice these changes in a number of different ways, including in relation to how you feel or how you act. Take a minute to think about these changes and then you can make a few notes on the card as a reminder.

Category B

4. ANXIOUS, SCARED, WORRIED

If you were to think of sometimes during the Program when you found yourself being really scared about something, or worried, or nervous or anxious about something, are there one or two things that come to mind? Again, take a minute to think about it and then make some notes on the card to remind you of those times.

5. UNCOMFORTABLE, AWKWARD, UNEASY

If you were to think of some times during the Program when you felt uncomfortable or awkward or uneasy about something, are there one or two moments that come to mind? Again, take a minute to think about it and then make some notes on the card to remind you of those times.

6. ANGRY

If you think back over your time in the Program, and you had to think about times during the Program you felt really angry about something, or times you got really mad or felt a sense of outrage or frustration, are there one or two things that come to mind? Again, take a minute to think about it and then make some notes on the card to remind you of those times.

7. TORN

As you think about your time in the Program, if you were to think of times you felt really in conflict about something, where someone or some part of you felt one way or was urging you on in one direction, and someone else or some other part of you was feeling another way; times when you really felt kind of torn about something, are there one or two things that come to mind? Again, take a minute to think about it and then make some notes on the card to remind you of those times.

8. SAD

As you think about your time in the Program, if you had to think about any times during the Program that you felt really sad about something, perhaps something that even made you cry, or left you feeling on the verge of tears, is

there anything that comes to mind? Again, take a minute to think about it and then make some notes on the card to remind you of those times.

Category C

9. IMPORTANT TO ME

If I were to ask you “What is it from your experience of the Program that is most important to you?” or “What from your experience of the Program matters most to you?” Is there anything that comes to mind? Take a moment to make some notes.

APPENDIX Z
PARTICIPANT INTERVIEW: INTRODUCTION TO THE INTERVIEW

1. Prepare cards
2. Greet and explain:
 - a. This is going to be an hour long interview about whatever experiences of the Program you want to discuss. The aim is for me to understand how you understood those experiences.
 - b. You don't need to talk about anything that you don't want to talk about and your participation is completely voluntary. You can stop the interview or change topics at any time.
 - c. There are two parts to the interview. In the first part, I'll give you some cards with some topics on them on which you can make some notes. These cards are for you. Nobody will see them, and you can take them with you when you finish.
 - d. The second part is when I'll ask you to choose some of the experiences you've made notes about on the cards to talk about. We won't get to all of them, so you can choose which ones you'd most like to discuss. This part of the interview will be recorded.
3. Questions:
 - a. What's going to happen? – I'll ask you to tell me about an experience you had on the Program and I might ask you questions about that experience until I feel like I understand what you are saying as best as I can.
 - b. Confidentiality – Your name and anyone and any place you mention will not appear in the transcribed version of the interview or anything that is written about the interview. In addition to anyone who transcribes the interview, only I will listen to the recording.
4. Get consent form signed.
5. Part I of the interview: I'm going to read a little about each card asking you to think about a time in the Program when you felt a particular way. You can

then write down some notes to remind you of the particular time. You don't need to write the whole thing out – just a few notes.

6. Interview: [Start recording]
 - a. I'm going to start the recording now.
 - b. So I'm going to ask you to start us off. Is there a particular card that stood out for you, that you might like to tell me about?

