Enhancing Adolescent Self-Concept, Life Effectiveness and Locus of Control: The Immediate and Lasting Effects of Outdoor Adventure Education

Mathew J. Pfeiffer

BA (Psych), BComm, PGDipPsych, DipPsychPrac

A thesis submitted in fulfilment of the requirements for degree of Doctor of Philosophy

Institute for Positive Psychology and Education

Faculty of Health Sciences

Australian Catholic University

Principal Supervisor: Professor Herbert W. Marsh

Co-Supervisor: Dr Theresa Dicke

Assistant Supervisors: Dr Jiesi Guo and Dr Keong Yap

January, 2020

STATEMENT OF AUTHENTICATION

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text.

I hereby declare that 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).

…….

Mathew J. Pfeiffer

January, 2020

ACKNOWLEDGEMENTS

There are too many people I would like to mention, so I must limit my comments to those who contributed particularly to this research.

To my principle supervisor, Herb. Our initial encounter that led to this project was nothing short of serendipitous. You made this opportunity possible, but more importantly, your encouragement and support ensured that I could be successful. Your expertise and knowledge brought a critical dimension to the project and your enthusiasm during our meetings ensured I was always excited about what was to come. You will always have my gratitude.

To my co-supervisor, Dr. Theresa Dicke, your attention to detail and assurance that I understood the 'why' and 'how' was critical to my understanding of psychometric analyses and research skills. You regularly invited my ideas and assisted me in putting them to practice. Your sensitivity to the personal challenges of undertaking a project like this also brought reassurance at just the right moments.

To my assistant supervisors, Dr. Jiesi Guo and Dr. Keong Yap. Jiesi, thank you for your unwavering commitment to my project. Your rapid replies to my emails, texts and phone calls truly knew no limit. Your expertise in coding and statistical analyses coupled with your jovial and fun nature made learning even the most complex ideas a pleasure. Keong, before transferring to the PhD you taught and supervised me in the Master program where your passion and energy always left me inspired. Your warm and thoughtful manner meant there was never a professional matter that I couldn't bring to your door and your clinical training has left me with a great deal to aspire to. Thank you.

To my dear wife, Susan. You made many sacrifices so I could complete this project. Without your patience, encouragement and support, it would not have been

feasible. In the time it took to complete this thesis, we married, built a home together, and were blessed with our greatest joy, our son, Toby. I love you both so much. Thank you.

My gratitude goes to all my family and in particular my parents, who provided me with every opportunity and who always unconditionally supported my passions. There were many times throughout my tertiary education where I felt like 'throwing in the towel', but you were always there with an encouraging hand to keep me on track.

Warm appreciation is due to my mentors who were instrumental in shaping my interest and career in psychology. My gratitude goes out to Dr William Clark, Dr Tom Cerni and Dr Alison Cornish.

To my colleagues at the IPPE, thank you for your friendship and for ensuring that there was always a laugh and helping hand only a chair-spin away. In particular, I would like to mention Ms Jillian Murphy for assisting with data collection and Ms Wendy Gwen, who generously contributed her ideas and wisdom while completing her PhD research in parallel to mine.

Thank you to the principal, staff and students at The Scots College, without whom this research would not have been possible. My gratitude goes out to the Glengarry staff who committed their time to ensure the research ran smoothly. They even welcomed me enthusiastically at the Coledale Surf Club for data collection after a 16km student hike. And of course, thank you to all the 2016 and 2017 Year 9 students who brought their insights, curiosity and interest to the project.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND SYMBOLS

Abbreviation / Symbol	Description
SEM	structural equation modelling
CFA	confirmatory factor analysis
ESEM	exploratory structural equation modelling
ω	omega reliability coefficient
df	degrees of freedom
CI	confidence interval
χ^2	chi squared
ES	standard deviation unit effect size
$\triangle$	delta (change)
CFI	comparative fit index
TLI	Tucker–Lewis Index
RMSEA	root mean squared error of approximation
MTMM	multitrait-multimethod
MTHM	multitrait-heteromethod
HTHM	heterotrait-heteromethod
HTMM	heterotrait-multimethod
M, Kurt, Skew	mean, kurtosis, skewness
SD, SE	standard deviation, standard error
r	product moment correlation
T0, T1, T2, T3	Time 0, Time 1, Time 2, Time 3
OAE	outdoor adventure education
TGP	The Glengarry Program
WLC	waitlist-control
PGE	post-group euphoria
SDT	Self Determination Theory
SDQ	self-description questionnaire
ROPELOC	review of personal effectiveness with locus of control
BPNSFS	basic psychological needs satisfaction and frustration scale
GHQ-12	general health questionnaire
MES	motivation engagement scale
Wellbeing	short Warwick-Edinburgh mental well-being scale

ABSTRACT

Outdoor adventure education (OAE) is embedded in the educational policy of countries around the world, with the objective of enhancing self-development in teenagers. Despite the apparent effectiveness of OAE, the research literature is criticised as being somewhat fragmented and as lacking methodological rigour. This has implications for undermining the position of OAE in the Australian curriculum. Accordingly, this investigation provides a critical review of outdoor education literature, conducts psychometric validation of the research instrumentation and reports on a longitudinal, controlled trial investigation of an OAE program.

Research indicates that OAE generally has a positive impact on the most widely studied psychological outcomes (e.g. self-concept, locus of control; ES ~ .32). Despite this, limitations of the existing quantitative literature include insufficient methodological rigour and the lack of a unified theory identifying the underlying mechanisms of change. Thus, the promising results are often described to be anecdotal.

To address these shortcomings, this research surveyed 346 male students enrolled in a single-sex Australian high school who were randomly assigned to one of two conditions: either attending a five-month OAE program, or continuing with their regular schooling i.e. the waitlist-control condition. The quantitative survey tool used contained 32 scales from nine established psychological instruments. Self Determination Theory (SDT) provided the framework through which changes in students' self-concept, life effectiveness, locus of control and other key psychological resources are inferred.

Study 1 investigated the psychometric properties of the survey instrument. Structural equation modelling supported the combined instruments' *a priori* factor structure (TLI = .93; CFI = .94; RMSEA = .02). Strong measurement invariance was

observed and multitrait-multimethod analysis found good support for convergent and discriminant validity.

Study 2 conducted a longitudinal investigation of the OAE outcomes. Short-term gains were observed on 19 of 32 outcomes (standardised ES = .18 to .71), with the greatest benefits found in social-emotional functioning, wellbeing, parent relations self-concept and locus of control. Six-month follow-up revealed maintenance of effects on 10 outcomes and the emergence of new gains on 7 outcomes (standardised ES = .17 to .39). The greatest benefits were found in social-emotional functioning, wellbeing, resilience, and academic engagement (persistence, planning and task management).

Study 2 further investigated the impact of students' baseline aptitude on treatment effectiveness, finding no systematic effects. Thus, the benefits from the outdoor experience appeared to be comparably available to all students regardless of whether they exhibited strength, or weakness against a particular outcome at the beginning of the program. Investigation into the mediating role of SDT's basic psychological needs revealed that from the 19 short-term effects, psychological needs satisfaction primarily mediated 10 outcomes and partially mediated 8 outcomes.

The findings from this research demonstrate the potential for OAE to benefit students' long-term psychosocial development that compliments traditional schooling. Furthermore, the principals of SDT appear to be largely applicable in OAE settings and it is suggested that basic psychological needs may be manipulated to foster greater program engagement and transference of health outcomes. These promising results, along with the research strengths and limitations, are discussed within the context of existing OAE literature. The implications for policy and practice, as well as directions for future research are further explored. It is hoped that this methodologically rigorous investigation will provide a strong basis for future studies to build on, while informing a more unified

framework for OAE program design and implementation so that the most effective educational experiences can be provided.

CHAPTER 1: INTRODUCTION

Introduction and Research Case

Adolescence is a complex yet exciting stage of life, in which the exploration of one's rapidly evolving physical, social and intrapersonal conditions are fundamental tasks. Although this critical developmental phase is characterised by tendencies for risk taking, impulsivity, and emotional dysregulation (Spear, 2000), and consequently is associated with increasing risk of mental illness (Kessler et al., 2005; Paus, Keshavan, & Giedd, 2008), adolescence is also a period of opportunities for positive development. Such opportunities include the formation of positive self-concept, the establishment of strong friendships, and the development of life skills and inner resources necessary to meeting life's demands.

Schools have been identified as pivotal settings for fostering healthy development in students (e.g. Kaplan & Flum, 2012; O'Mara, Green, & Marsh, 2006). For decades, Outdoor Adventure Education (OAE) has been embedded in educational policy of countries around the world, with the aim of enhancing personal resources and self-development among students (Marsh, Richards, & Barnes, 1986a; Richards, 1977; Zink, 2010). National advocates for OAE in schools believe "**all** Australian children and Australian society benefit immensely from outdoor experiences" (Outdoor Education Australia, 2018, Education, para. 1; bolding in the original text). This experiential form of education generally involves groups of students to spend time in the outdoors, where they engage in adventure activities that incorporate mental and physical challenges. The rationale is that skills learnt in unfamiliar environments can be assimilated and applied to daily life. Gutman and Schoon (2013) note that educational policy makers are placing increasing emphasis on non-cognitive soft skills, attitudes and strengths that may be even more important than cognitive hard skills in determining academic and vocational

outcomes. Such soft skills are frequently associated with OAE (Richmond, Sibthorp, Gookin, Annarella, & Ferri, 2018), and findings from meta-analyses examining OAE outcomes suggest favourable results (Cason & Gillis, 1994; Hattie, Marsh, Neill, & Richards, 1997; Neill, 2002; Wilson & Lipsey, 2000). However, methodological weaknesses (e.g. Dillon, 2013; Marsh, Richards, & Barnes, 1986c; Rickinson et al., 2004; Scrutton & Beames, 2015) and the lack of identification of specific mechanisms of change within OAE literature, undermine these largely promising findings (Sibthorp, Furman, Paisley, Gookin, & Schumann, 2011; Sibthorp, Paisley, & Gookin, 2007). Sibthorp, Paisley, and Gookin (2007) refers to the perceived over-dependence on anecdotal evidence as "Black Box programming, where it seems that simple participation is assumed to lead to participant development without any ability to describe the specific mechanisms through which change may occur" (p. 1). The lack of empirical grounding within OAE literature further led Neill (2002) to conclude "It is no longer sufficient for outdoor education (OE) advocates to simply believe in the benefits of OE, or to accept anecdotal evidence at face value" (p. 8). Consequently, the place of OAE programs in high-schools is under threat (O'Brien et al., 2011), despite the growing body of evidence linking OAE experiences with some of the most highly regarded psychosocial outcomes in developmental and educational psychology.

This investigation addresses these stated limitations by adopting a strong research design that integrates an empirical investigation into the mechanisms through which OAE outcomes come about. This is achieved by implementing best practice design standards recommended in major OAE reviews including the use of a randomly assigned control group, follow-up to post-treatment measurement, and a well-defined method for administering the OAE program (e.g., Cason & Gillis, 1994; Dillon, 2013; Marsh et al., 1986c; Scrutton & Beames, 2015; Wilson & Lipsey, 2000). Additionally, this research

examines the influence of participant individual differences on OAE program success and incorporates a strong theoretical model to quantitatively investigate the causal mechanisms through which the OAE fosters healthy psychological change. In doing so, the present investigation will fill a need in the current research and thus be well-placed within the broader literature in this area.

The OAE intervention program in this research, named The Glengarry Program (TGP), is a five-month residential experience for all Year 9 students attending The Scots College, an urban boys-school in New South Wales (NSW), Australia. The program constitutes a compulsory aspect of students' schooling and incorporates a formal academic curriculum (see Chapter 3: The Intervention Program). TGP is located at a bushland setting in Kangaroo Valley (a river valley along Kangaroo River, approximately 150 km south-west from Sydney). It is anticipated that this research will derive much-needed empirical evidence that is timely in relation to Australian policy debates. One such debate regards the National School Reform Agreement between the federal government, states and territories that commenced January 2019 and stipulates a series of five-year reform actions to improve student outcomes (Australian Government Department of Education and Training, 2019).

Purpose and Project Aims

The overarching purpose of this investigation is to quantitatively examine the effects of OAE on healthy adolescent development. More specifically, short- and long-term changes in self-concept (Marsh, 1990a), life effectiveness (Neill, 2008) and locus of control (LoC; Strickland, 1989) were examined in male students attending the single-sex high school. These constructs, and other secondary psychological resource outcomes (see Chapter 4: General Method) were selected for analysis due to their predominant position within OAE research and their intentional alignment with the idiosyncratic aims of the

TGP, as described in Chapter 3. The influence of students' baseline aptitude on treatment gains and the mediating role of basic psychological needs satisfaction were further analysed among the male adolescents attending the five-month OAE program (see Chapter 2: Literature Review).

In summary, this research aims to address three notable gaps within OAE scientific research: (a) a lack of well-defined outcome measures that reflect the specific goals of the OAE program and which are supported by robust psychometric measurement; (b) a lack of longitudinal, controlled trial studies with sufficient sample size and statistical power; and (c) insufficient exploration of independent factors that explain the positive outcomes from OAE. In achieving these aims, it is expected that this research will lead to a more unified framework for OAE program design and implementation, while providing a strong platform for future studies to build on. The two overarching studies in this thesis will now be presented.

The Studies

This project's aims and concepts are anchored in current knowledge and debate, and are pursued through the application of new and presently-evolving methods to address the need for the greater rigour pertaining to OAE theory and research.

Chapter 2 commences with a review of the developmental literature, and of adolescent health information of contemporary relevance. Self Determination Theory (SDT), a leading empirical theory of human development, is then introduced and conceptualised both practically and theoretically in relation to the OAE context. It concludes with a review of prior OAE literature and the consolidation of research findings regarding the factors that influence OAEs' success. This section provides a much-needed synthesis of OAE literature.

The OAE program that provides the intervention in this research, TGP, constitutes a promising treatment model that is further detailed in Chapter 3: The Intervention Program. TGP involves an extended residential stay during which a holistic approach to adolescent boys' education and personal development is adopted. TGP is school-integrated, maintaining well-defined goals that underscore the philosophy and values of The Scots College. Chapter 3 provides a detailed account of this OAE program to convey the research context, including program history, aims and philosophy, content and design; each of these points is salient to the identification of primary outcomes in this research.

A long-standing criticism of OAE research concerns the lack of appropriately selected survey scales and of well-defined measurements of the outcomes being explored (Dillon, 2013; Marsh et al., 1986c; Rickinson et al., 2004; Scrutton & Beames, 2015). To address the need for greater psychometric rigour, Study 1 (Chapter 5) adopted a methodical approach to psychometric validation of the research instrumentation. This was deemed critical, to establish an appropriate measurement regime with the research sample. The internal consistency of outcome scales was tested through statistical analysis before structural equation modelling (SEM) procedures were employed to derive each model's representations of best fit. Each measure was analysed in relation to invariance over time, and tested in regard to convergent and discriminant validity. The use of SEM procedures allowed for latently-derived factor scores to be computed (see Chapter 4: General Method), thus increasing the power of the intervention analysis in Chapter 6 (Study 2).

The overarching aims of this investigation, and two predominant recommendations within OAE literature, are addressed in greater detail in Chapter 6 (Study 2). They include:

(a) a randomised controlled trial (RCT) longitudinal evaluation of the short- and long-term effects of OAE on self-concept, life effectiveness and LoC (see Chapter 2: Literature Review); and

(b) an empirical investigation into the role of specific mechanisms of change through which OAE outcomes transpire.

Applying a RCT design, Study 2 thoroughly investigates the potential for OAE to provide short- and long-term benefits in multiple domains of psychosocial development among male youths. Furthermore, students' baseline aptitude (i.e. strengths and weaknesses on constructs measured), and SDT's basic psychological needs are examined as to their influence on the OAE program's outcomes.

Chapter 7 discusses the findings from Study 1 and Study 2 in relation to the theory and practice of OAE programming. The theoretical and methodological contributions of the investigation and implications for national policy are considered, and recommendations for future research and practice provided.

In summary, this research provides much-needed hard evidence by empirically evaluating the effects of OAE on adolescent males' positive psychological development, whilst exploring the mechanisms influencing these effects. Only a small percentage of OAE studies undertake follow-up evaluation, or use controlled trials. This investigation thus offers a strong, theoretically grounded and empirically derived platform to generate new insights into the best-practice design and implementation of OAE programming. Additionally, this research provides a sound empirical basis for future OAE studies to build on. In doing so, the most effective youth-orientated interventions might be developed, to promote healthy living and positive psychological development for adolescents.

Anticipated Research Impact

As Australia embarks on a new educational era in seeking to implement a new, national curriculum, the place of OAE in schools has come under question (O'Brien et al., 2011). Many schools have OAE programs (Wang, Ang, Teo-Koh, & Kahlid, 2004) but, as is the case with many applied and field educational interventions, little OAE research has come in the form of high quality, extensive controlled trials (Cason & Gillis, 1994). Furthermore, only a small percentage of OAE programs have been subjected to empirical evaluation (Bowen & Neill, 2013), despite meta-analytic reviews reporting an average of small to medium effect sizes across a range of psychosocial indicators. The O'Brien et al. (2011) review of OAE warns "lack of evaluation of learning outcomes as well as other personal and social development outcomes and health and well-being outcomes leaves the sector in a vulnerable position" (p. 369). This warning may account for the declining position of OAE in Australia's national curriculum, despite the apparently very positive outcomes that it can promote.

By addressing the limitations that have inhibited the progress of OAE, the findings from this research will contribute to the broader body of literature by empirically evaluating the psychosocial and emotional benefits associated with OAE experiences. Additionally, new insights will be generated into the underlying mechanisms through which these benefits relate to youths. It is anticipated that the results of this research will inform national policy change regarding the accessibility and design of OAE within schools' curriculum.

Summary

Adolescence is often depicted as a stressful time. However, the rapid biopsychosocial changes that happen throughout this period present a valuable opportunity to engage youths in enriching experiences that can foster personal effectiveness and inform

critical developmental processes. These include the formation of self-concept and the establishment of strong relationships.

Schools have been identified as effective platforms through which youth-interventions may be integrated. However, OAE represents a treatment protocol that has in effect been marginalised in educational settings partly through lack of strong empirical scientific research, despite strong anecdotal indicators of these programs' effectiveness. The present research aims to close this gap by studying an extended, school-embedded OAE program for male adolescents, and quantitatively evaluating changes in key developmental outcomes associated with it. Additionally, this investigation explores the specific mechanisms through which OAE may achieve these outcomes. The following chapter provides a critical review of the literature pertinent to these aims.

CHAPTER 2: LITERATURE REVIEW

Adolescent Development

Mental health and risk factors. Adolescence is often depicted as a turbulent stage of life, due to the rapid changes in social, biological, emotional and sexual functioning. It is therefore defined as a time of transition, wherein identity formation is a central task (Erikson, 1968). Adolescence is characterised by under-developed self-regulation devices, a state that contributes to heightened emotionality, sensitivity to peer influence, impulsivity and novelty-seeking (Spear, 2000). Rankin, Lane, Gibbons, and Gerrard (2004) demonstrate how self-consciousness during adolescence moderates a greater frequency of self-reflection and self-scrutiny than that observed during pre-adolescent years. With adolescence demonstrating heightened attention to both internal and external aspects of the self (Lerner, Kier, & Brown, 2002; Neff & McGehee, 2010; Santrock, 2003), this critical period of development is characterised by peak onset of mental illness (Paus et al., 2008). On average, one in five youths are diagnosed with a psychological disorder that will persist into adulthood (Kessler et al., 2005) and suicide presents as the leading cause of death amongst Australian children 5 to 17 years of age (Australian Bureau of Statistics, 2016). The 2018 Youth Survey (Mission Australia, 2018) found that 43% of Australians aged 15 to 19 years identified mental health as the greatest national issue, a response rate double that seen in 2016.

Australian youths health statistics. The Australian Bureau of Statistics (2015) offers caution regarding unhealthy behaviours during childhood that increase the risk of maladaptive lifestyle patterns later in life. The 2014-15 National Health Survey (Australian Bureau of Statistics, 2015) revealed that one in four (27.4%) of Australian children aged five to 17 were classified as overweight or obese, with only 1 in 20 (5.1%) children aged 2 to 18 satisfying the Australian Dietary Guidelines for fruit and vegetable daily intake

(National Health and Medical Research Council, 2013). With adolescents reflecting reduced activity levels compared to their pre-adolescent counterparts (Nader, Bradley, Houts, McRitchie, & O'Brien, 2008), the strong positive correlation between activity levels and improved physical and mental health (Department of Health, 2016; Janssen & Leblanc, 2010) has led to the promotion of active lifestyles as a national health initiative in Australia (Department of Health, 2016).

Without question, the ways in which teenagers navigate adolescence has become increasingly complex. This is a consequence of the 21st century bringing about greater accessibility to rapidly advancing technologies (Department of Health, 2016), aggregating values and attitudes, and evolving interpersonal demands (Karoly & Panis, 2004). The World Health Organisation (2016) identifies increases in sedentary behaviour as a threat to public health and thus, schools are pressured to adapt accordingly and to combat these negative effects on students' physical, social and emotional development. It appears that OAE may offer educational policy makers one such opportunity to engage students in healthy practices that foster personal growth, learning, socialisation and healthy lifestyle behaviours, with the potential to extend well beyond adolescent years. This ambitious claim is a central premise of the present research.

Typical psychological trajectories. All teenagers are faced with the complex task of developing a working 'self', which enables them to be effective as adults. Adolescence, commonly referred to as a 'rite of passage', is a time of growth, exploration, discovery and adjustment, characterised by developmentally typical ebbs and flows in some of the most fundamental cognitive and affective operations. Blakemore, Burnett, and Dahl (2010) highlight that neuroendocrinological events associated with puberty impact "the perceptions, motivations, and behavioural repertoire of the individual" (p. 926), and are not restricted to changes in sexual maturation.

Yeung (2011) observed declines in perceptions of academic competencies, interest in academic subjects and effort goal orientation throughout high-school in a multicultural Australian youth sample. These findings are consistent with earlier research, showing a decline in self-concept (Marsh, Craven, & Debus, 1998) and intrinsic school motivation and engagement among school-aged students as they mature (Gillet, Vallerand, & Lafrenière, 2012; Gnambs & Hanfstingl, 2015; Lepper, Corpus, & Iyengar, 2005; Marsh, Parada, & Ayotte, 2004; Martin, 2007).

Conversely, students' locus of control (LoC) has been found to become increasingly internalised throughout adolescence (Knoop, 1981), with the greatest benefits occurring between grades 10 and 11 (Bachman, O'Malley, & Johnston, 1978; Chubb, Fertman, & Ross, 1997). This suggests that prior to age 15-16 (Grade 10) students hold more external control expectancies, a disposition associated with greater risk-taking (Özmen & Sümer, 2011), poorer health behaviours (Blaxter, 1990) and vulnerability to developing learned helplessness, a state associated with reduced motivation and disruptions to learning and emotional regulation (Petersen, Maier, & Seligman, 1995). Similarly, Yeung and McInerney (2005) observed that students aged 12-13 (Grade 7) displayed significantly greater effort orientation and career aspirations compared to students aged 14-15 (Grade 9), but not in comparison to older students in Grade 11. Together, these findings indicate that mid-adolescence is frequently typified by the greatest decrements in a range of positive psychological indicators.

In short, there are well-documented developmentally typical declines in key psychological variables throughout adolescence, with greatest 'dips' often observed in mid-adolescence. Such developmental patterns hold implications for internal school tracking, intervention protocols, and the interpretation and generalisation of research

findings derived from adolescent populations. But a greater understanding of the reasons for this decline is needed so as to guide appropriate interventions.

One explanation for downward developmental trends during teenage years is that more mature students begin to question the value of schooling and the personal resource investment involved (Wigfield & Eccles, 2000). Marsh et al. (1998) note that the decline in youths' self-concept may result from the originally unrealistic estimations of personal capabilities becoming more accurate with maturation; with this, they realise their personal weaknesses. Gnambs and Hanfstingl (2015) provide an alternative explanation, finding that diminished basic psychological needs satisfaction in respect of autonomy, relatedness and competence (see Chapter 2: Self Determination Theory) accounts for the reductions in high-school students' motivational tendencies.

Summary. Adolescence is a pivotal time to establish healthy lifestyles, which may promote lasting positive mental and physical health and protect against the risk factors that characterise this stage. The following section introduces the grounding theory for the present research as it relates to outdoor adventure education (OAE) and developmental outcomes in teenagers. It is intended that this theory will provide a strong basis to investigate the underlying mechanisms of change from the OAE program, a central aim of this research.

A Theory of Human Motivation and Development: Self Determination Theory

Self-Determination Theory (SDT; Deci & Ryan, 1985) provides an empirically derived theory of human motivation, behaviour and wellbeing. The framework proposes that all individuals are inherently active and intrinsically motivated, provided that their psychological needs of autonomy, competence and relatedness are satisfied (Deci & Ryan, 1985; Ryan & Deci, 2000b, 2000c, 2017). Individuals strive to satisfy these psychological needs, which are quintessential for adaptive self-development (Ryan & Deci, 2000b),

integrity, well-being (Ryan & Deci, 2000a, 2017) and optimal identity formation among school aged children (La Guardia, 2009).

Basic psychological needs. At the heart of SDT is the assumption that all humans have needs which they must fulfil if they are to cultivate and enrich their development and functioning (Ryan & Deci, 2000a, 2002). Just as an organism has physiological needs, SDT asserts likewise that all individuals have an inherent need for autonomy, competence and relatedness, which are conditional for optimal personality formation, cognitive functioning and psychological well-being (Ryan & Deci, 2002, 2017). Failure to satisfy any of these needs is theorised to impede development and wellbeing. Ryan and Deci (2017) propose "any factor or event that produces variations in needs satisfaction or needs frustration will also produce variations in wellness" (p. 243). The association between psychological needs and positive health outcomes is deemed partly due to the impact of past and present need satisfaction on an individual's ability to adaptively experience negative emotional states (Ryan & Deci, 2000a).

Autonomy. Autonomy involves feeling ownership over one's behaviour that is governed by volition, is self-initiated and involves freedom of choice (La Guardia, 2009). When students' actions are autonomous there is an alignment between their values and behaviour. Hansen and Jessop (2017) suggest that of SDT's three psychological needs, autonomy is especially crucial to support the emergence of teenagers' developmental potential. Autonomy-supportive environments have been linked with increased self-esteem, ego development, self-actualisation (Deci & Ryan, 1985) and internal LoC in adolescence (Ryan & Deci, 2002).

Relatedness. The psychological need satisfaction of relatedness is experienced by developing meaningful relationships and a genuine sense of belonging within one's social context (Ryan & Deci, 2000a). La Guardia (2009) identifies students' relationships with

teachers as particularly relevant in informing adolescents' developing self-concept, goals, and identity-orientated behaviours. In school settings, students' psychological need for relatedness can be fostered through the expression of individual attention and by establishing supportive interpersonal environments.

Competence. The psychological need for competence presumes the experience of personal effectiveness within one's social environment through the expression of personal capacities (Deci, 1975; Ryan & Deci, 2002). This does not imply achievement or skill development external to the individual's perspective, but rather is felt as a sense of mastery (Ryan & Deci, 2002). When students feel they can be successful at school they feel capable of undertaking new learning tasks (Ryan & Deci, 2002; Schunk & Pajares, 2005), show greater learning interest, apply greater effort and respond more adaptively to aversive experience (Bandura, 1997). Fulfilment of students' need for competence is associated with positive self-belief perceptions (Hughes, Galbraith, & White, 2011) and enhanced self-efficacy (e.g. Bandura, 1997; Shih, 2006; Usher & Pajares, 2006; Valentine, DuBois, & Cooper, 2004)—this latter being a construct pertinent to students' learning motivation, engagement and achievement (Pajares & Schunk, 2005).

Satisfaction of students' basic psychological needs has been linked with more internalised LoC and enhanced intrinsic motivation and wellbeing (Deci & Ryan, 1985; Ryan & Deci, 2000a). In contrast, in situations where they are faced with external control, leading to feelings of being pressured towards action, student behaviour is no longer self-determined, as autonomy is lost and competence dispossessed (Deci & Ryan, 1985). Social climates that fail to adequately fulfil students' psychological needs have been linked with adjustment problems, psychopathology such as disordered eating, negative affect, physical symptoms (Bartholomew, Ntoumanis, Ryan, Bosch, & Thogersen-Ntoumani, 2011), conduct disorders (Kasser, Ryan, Zax, & Sameroff, 1995), maladaptive patterns of

cognitive, affective and behavioural functioning (Deci & Ryan, 2000; Ryan & Deci, 2000a) and disruptions to personality formation (Ryan & Deci, 2000a; Ryan, Deci, & Grolnick, 1995).

In SDT there are two dominant motivational orientations which are directly influenced by the degree to which psychological needs are satisfied and subsequently play a pivotal role in healthy development (Ryan & Deci, 2000b). These are intrinsic and extrinsic motivation. While independent evaluation of these motivational forms is not within the scope of this thesis, measured constructs of LoC provide a means of inferring students' motivational orientation. This is due to research demonstrating the link between intrinsic motivation with internal LoC (e.g. Sundjoto, 2017) and similarly, extrinsic motivation with external LoC (e.g. Pelletier, Dion, Tuson, & Green-Demers, 1999). These motivational orientations will now be discussed.

Intrinsic and extrinsic motivational orientations. According to Ryan and Deci (2000b), to experience motivation is "to be moved to do something" (p. 54). The impetuses of one's motivation may change in orientation depending on one's attitudes and goals, while the magnitude of motivational levels may vary in degree (Ryan & Deci, 2000b, 2017). Within SDT, two forms of motivation engender action: namely, intrinsic motivation and extrinsic motivation (Ryan & Deci, 2000b).

Intrinsic motivation. Intrinsically-motivated behaviour is characterised by curiosity, interest and an inherent sense of enjoyment or satisfaction (Deci, 1975; Ryan & Deci, 2017). Ryan and Deci (2017) define the term as "a manifestation of our natural human propensities to assimilate and integrate knowledge" (p. 354). Such propensities play a critical role within healthy development (Ryan & Deci, 2000b, 2017), in accordance with SDT's fundamental assumption that all humans are naturally active and driven towards personal growth (Ryan & Deci, 2000b). Ryan and Deci (2000b) state "This natural

motivational tendency is a critical element in cognitive, social, and physical development because it is through acting on one's inherent interests that one grows in knowledge and skills" (p. 56). This notion has led the intrinsic orientation of motivation to receive significant attention within educational settings and in developmental literature. A growing body of evidence has linked this motivational tendency with greater learning, creativity (Deci, Vallerand, Pelletier, & Ryan, 1991; Ryan & Deci, 2000b; Sheldon, 1995), interest and task performance (Grolnick & Ryan, 1987), use of more-advanced learning strategies (Lam, Cheng, & Ma, 2009), persistence (Deci & Ryan, 1991), vitality (Nix, Ryan, Manley, & Deci, 1999), improved school adjustment (Mouratidis, Vansteenkiste, Lens, Michou, & Soenens, 2013), enhanced positive emotional experience and general well-being (e.g. Burton, Lydon, D'Alessandro, & Koestner, 2006; Bye, Pushkar, & Conway, 2007; Ryan et al., 1995).

Extrinsic motivation. In contrast, extrinsic motivation drives behaviour that is intended to generate outcomes separate from one's internal disposition (Ryan & Deci, 2000b). Extrinsic motivational is particularly relevant in school settings, where educators are often pressured to emphasise academic achievement instead of broader developmental outcomes. Ryan and Deci (2017) warn that "many schools fail to capitalise on students' intrinsic motivation and instead emphasize extrinsic motivators." (p. 351). The impact of low intrinsic motivation at school has been associated with poorer outcomes in education and wellness domains (Gottfried, Gottfried, Morris, & Cook, 2008).

Empirical research has frequently demonstrated how the clash of controlling versus autonomous regulation may support, or otherwise thwart, one's motivational orientation (e.g. Deci, 1971, 1972a, 1972b; Eisenberger & Cameron, 1996; Ryan & Deci, 2017). This pattern is especially relevant in school settings, where students' engagement will often not be intrinsically motivated. Leading SDT theorists thus established the Organismic

Integration Theory (OIT), which provides a framework for conceptualising the factors that drive an individual's actions when intrinsic interest is not presumed.

Organismic integration theory: Internalisation and self-integration. Self-Determination Theory views all people as being intrinsically moved to engage their environment in order to actualise their capabilities (Ryan & Deci, 2002). The Organismic Integration Theory posited within SDT highlights the operation of internalisation, defined as the "process of active learning and self-extension" (Ryan & Deci, 2017, p. 180) via the integration of external regulations such as values, beliefs and behaviours, to the self (Ryan & Deci, 2002, 2017). According to Ryan and Deci (2002), more autonomous regulation facilitates greater internalisation and thus self-integration, which is at the heart of self-determined action. Given that to a significant degree, students' school engagement cannot fundamentally be motivated by interest and enjoyment, understanding how extrinsic motivation may facilitate or otherwise undermine self-determined action, has become increasingly relevant in educational settings.

The application of SDT's principles of internalisation and self-integration has been subject to vast empirical evaluation due to the practical alignment of basic psychological needs satisfaction with the theoretical position of internalisation in mediation research. That is, by satisfying psychological needs, an intervention may foster more autonomous engagement in the experience, which in turn enables greater assimilation of treatment goals and health outcomes being realised. However, the interplay between environment and an individual's psychological needs satisfaction is reciprocal, and people strive to fulfil these needs. When psychological needs are adequately met, behaviour becomes self-determined and an individual is positioned as the agent of change. This agency motivates further pursuit and identification of psychological needs satisfaction. Little, Hawley, Henrich, and Marsland (2002) refer to this self-maintaining cycle as the causal action

sequence, a phenomenon that provides rationale for continual growth in outcome effects well after an intervention has concluded, namely ‘sleeper effects’. Jang, Kim, and Reeve (2012) observed that satisfaction of mid-semester autonomy led to an increase in classroom engagement and end of semester achievement among eighth grade students. Evans and Bonneville-Roussy (2016) applied SDT to Australian and New Zealand music students, finding that greater satisfaction of autonomy predicted higher quality and frequency of practise and the pursuit of learning more complex pieces of music. Despite the growing body of literature demonstrating that the principles of SDT constitute the critical mechanisms for change in intervention research, few studies have applied this theory to OAE.

Fostering Students’ Psychological Needs in Outdoor Adventure Education

The principles of SDT have been applied to OAE programs to explicate their claimed benefits (e.g. Daniel, Bobilya, Kalisch, & McAvoy, 2014; Wang et al., 2004). As emphasised in the Hattie (2009) review of educational research, the gains associated with OAE appear to be internalized by students resulting in rare sleeper-effects, where outcomes are stronger at follow-up than immediately post-intervention. Self Determination Theory explains this effect (Deci & Ryan, 1985; Ryan & Deci, 2000c), wherein the satisfaction of basic psychological needs leads students to more fully internalise program goals and therefore are more likely to realise healthy social, personal and physical development over time.

As Wang et al. (2004) state, successful OAE programs must meet the basic psychological needs of participants. This can be achieved by:

1. Providing participants with a meaningful rationale for their OAE involvement that promotes the values that students associate with the OAE experience;

2. Promoting autonomous action by including opportunities for choice and decision-making; and,
3. Building students' perceived competence by exposing them to positive learning experiences and providing them with encouraging feedback.

Scarf et al. (2017) and Scarf et al. (2018) demonstrated the important mechanism of relatedness in OAE programming when they observed that increases in youths' sense of belonging during a sailing expedition predicted gains in self-esteem and resilience, respectively. Hans (2000) elaborated on the importance of decision-making and choice availability during OAE, suggesting such opportunities enable participants "to more readily accept responsibility for their level of success in the program. As a result, they will feel more in control of self, necessitating that a certain degree of internality be present" (p. 41).

Summary. Self Determination Theory offers an evidence-based explanation of motivational orientations and basic psychological needs that are fundamental to healthy personality development and well-being (Deci & Ryan, 1985; Ryan & Deci, 2000). The principles underlying SDT have been repeatedly and successfully posited as the mediating mechanisms within intervention research but, surprisingly, very few studies have applied SDT to OAE investigations.

The following section provides a review of OAE literature, outlining research findings and rationales for these programs. Moderating factors which are known to influence the success of OAE will be discussed, followed by a review of three predominant constructs within OAE literature, namely self-concept, life-effectiveness and LoC. Lastly, the shortcomings of the body of OAE literature will be discussed.

Outdoor Adventure Education

Background and rationale. Outdoor Adventure Education has existed in educational curricula since 1941, when Kurt Hahn (1957) founded the first outward bound school. The success of the program, aiming to enhance self-reliance, independence, physical endurance and initiative, led to Hahn's continued involvement in the development of OAE schools across the world. His pioneering work in outdoor education led to the first Australian OAE program for youths being implemented in 1957, and contributed to the establishment of the esteemed Duke of Edinburgh Awards, which are offered in many schools across Australia to this day.

Hahn (1957) claimed that the aims of OAE are to "enthral and hold the young through active and willing Samaritan service, demanding care and skill, courage and endurance, discipline and initiative" (p. 10). Marsh et al. (1986a) report that OAE enhances self-regulation, goal setting and resilience skills by engaging participants in unfamiliar environments and difficult challenges. In doing so, an individual may "recognize and understand his own weaknesses, strengths, and resources and thus find within himself the wherewithal to master the difficult and unfamiliar" (Richards, 1977, p. 69). A recent study by Wilson and Sibthorp (2017) found that experiences of success and failure during OAE contributed to positive changes in participants' self-efficacy, coping and perseverance.

Outdoor education programs seek to foster holistic growth in participants through exposure to hands-on, meaningful learning experiences, referred to as experiential education. Dewey (1938) defines this as "education of, by and for experience" (p. 10), an integrative and systematic process consisting of concrete experience, reflection, abstract conceptualisation and subsequent active experimentation (Kolb, 1984). The rationale is that skills learnt in unfamiliar environments can transcend that context and be applied to

daily life (Holman, Pavlica, & Thorpe, 1997; Marsh & Richards, 1988; Priest, 1997), therefore integrating personal development programs based on exposure to unfamiliar situations, with the self-development and goal setting principles of Self Determination Theory.

Comparisons between OAE and other interventions, such as therapeutic, social, behavioural and academic, appear favourable (Bowen & Neill, 2013; Bowen, Neill, & Crisp, 2016; Hattie, 2009; Hattie et al., 1997). Empirical work by leading OAE researchers suggests that these programs generally have a positive impact on the most widely studied psychological outcomes (Marsh & Richards, 1988; Marsh et al., 1986a). The Hattie (2009) summary of 800 meta-analyses of educational outcomes stressed that OAE programs are the most successful of all interventions at enhancing self-esteem. From 96 studies, Hattie et al. (1997) found OAE to be most effective at providing participants with a sense of self-control, responsibility, self-regulation and self-assurance. This extensive review resulted in an overall average effect size of .34 at the end of the programs, with additional gains, namely sleeper effects, at follow-up (ES = .17)—this latter phenomenon being a remarkable and unusual finding in educational intervention research.

Other major OAE reviews have evidenced benefits to LoC (Hans, 2000), school attendance and adjustment (Cason & Gillis, 1994; Wilson & Lipsey, 2000), academic achievement (Marsh & Richards, 1988), self-esteem, multiple dimensions of self-concept (Cason & Gillis, 1994; Hattie et al., 1997; Marsh et al., 1986c), and the personal effectiveness areas of emotional stability, time-management, decision making, independence, task leadership, assertiveness, time management, and both personal and social development (Bettmann, Gillis, Speelman, Parry, & Case, 2016; Neill, 2008).

It is apparent that OAE can indeed provide powerful transformational experiences for teenagers which benefit them in behavioural, educational, social and personal resource

domains. However, few studies have examined the effects of school-integrated OAE programs, due to their being classified as too brief and less challenging, achieving fewer outcomes (see Hattie et al., 1997). Consequently, OAE programs in schools appear to be marginalised within the broader body of literature.

Classification and characteristics of OAE. Outdoor adventure programs are run in settings throughout the world and cater for clienteles of disparate types. Consequently, variance in philosophy, content, purpose and physical setting has engendered caution against the use of any formalised definition of OAE (e.g. Brookes, 2006; Neill, 2008). However, the identification and classification of program characteristics has proven useful to provide a broader picture of this expansive body of literature.

Outdoor Education programs vary in length from just a few days (e.g. Outward Bound: Inspiring Australians, 2016), to weeks (e.g. Sibthorp & Arthur-Banning, 2004), up to an entire year (e.g. Geelong Grammar School Exceptional Education, 2016). They cater for children (e.g. Scrutton, 2014), adolescents (e.g. Dolgin, 2014; Rahman, 2009) and adults (e.g. Vlamis, Bell, & Gass, 2011) from diverse cultural groups (e.g. Louw, Meyer, Strydom, Kotze, & Ellis, 2012; Ritchie, Wabano, Russell, Enosse, & Young, 2014; Wang et al., 2004) and of varying psychosocial, behavioural and intellectual functioning (e.g. Bowen, Neill, & Crisp, 2016; Bowen, Neill, Williams, et al., 2016; Larson, 2007; Vörösvári, 2016).

Typically, OAE programs take place in natural, outdoor settings such as bushlands, back-country and national parks, however they may also be run effectively indoors (Richards, Carpenter, & Harper, 2011). Some programs require overnight boarding, while others adopt a daily sessional approach, or a combination of both (Hans, 2000).

Priest and Gass (1998) typify four categories of OAE programs as:

(a) recreational programs, intended to evoke positive affect (e.g. enjoyment, excitement, relaxation);

(b) educational programs, enhancing knowledge and understanding;

(c) developmental programs, fostering personal development, adaptive behaviours and life skills; and

(d) therapeutic programs, improving dysfunctional behaviour.

In the Hattie et al. (1997) critical review of OAE literature, six universal features of OAE programming are identified:

> (a) wilderness or backcountry settings; (b) a small group (usually less than 16); (c) assignment of a variety of mentally and/or physically challenging objectives, such as mastering a river rapid or hiking to a specific point; (d) frequent and intense interactions that usually involved group problem solving and decision making; (e) a nonintrusive, trained leader; and (f) a duration of 2 to 4 weeks (p. 44).

Regardless of the variations in categories and features of OAE programs, they appear to share a grounding philosophy, in that "the most striking common denominator of adventure programs is that they involved doing physically active things away from the persons' normal environment" (Hattie et al., 1997, p. 44). The OAE program that constituted the intervention in this investigation is consistent with this description.

Drivers of outdoor adventure education program effectiveness. A major criticism of OAE research concerns the lack of clarity around the specific mechanisms through which program goals are achieved (e.g. Sibthorp & Morgan, 2011; Sibthorp, Paisley, & Gookin, 2007). Consequently, recent decades have seen an emergence of empirical studies looking to address this criticism, to generate insights into the factors which are instrumental to these programs' success.

Program duration and exposure. Meta-analyses and other major OAE reviews have consistently shown that longer programs are associated with more favourable

outcomes (Bettmann et al., 2016; Cason & Gillis, 1994; Hattie et al., 1997; O'Brien et al., 2011; Wilson & Lipsey, 2000). Hattie et al. (1997) found this to be especially the case at follow-up, where the average overall effect size for programs > 20 days (ES = .20) exceeded that of programs 19 days or less (ES = .13). Further supporting the notion that greater programmatic exposure is associated with greater OAE outcomes are the findings from Hans (2000). This meta-analysis successfully replicated findings from Hunter and Purcell (1984) by revealing that over-night boarding is associated with larger effects, compared to daily sessional OAE programs. In a similar vein, Marsh and Richards (1988) concluded that the operationalisation of OAE away from school settings was a critical factor in the enhanced academic self-concept among low-achieving high-school males. These findings support the central belief that effective OAE programs take place in unfamiliar settings.

Participant age. Meta-analytic reviews have repeatedly observed a positive correlation between participant age and OAE effect size (e.g. Bowen & Neill, 2013; Hattie et al., 1997). Bowen and Neill (2013) found that participant age moderated 6.8% of OAE short-term effects, students aged < 9 years yielding the smallest effect sizes (ES = .24), those aged 10 to 14 small to medium effects (ES = .37), and those aged 15-17 years large effects (ES = .50; see Chapter 4: General Method), while the largest effect sizes were found in participants of 18+ years (ES = .66). Hattie et al. (1997) observed students yielding mean effect sizes of .21 compared to .38, seen in adult participants at post-test. Conversely, Cason and Gillis (1994) found younger participants to show the greatest improvements. One explanation for this lies with the placement of mid-adolescence relative to the sample mean ages in each study. Hattie et al. (1997) included predominantly adult participants, whereas Cason and Gillis (1994) included participants from the age of 11 to those in first year of university. If adolescence is a time of stress, then the apparent

curvilinear age effect for declines in self-concept during mid-adolescence (Marsh, 1990a) may also explain an inhibited sensitivity to OAE effects during this stage of development.

Structure and philosophy. The influence of factors such as OAE program structure, philosophy and goals has gained considerable attention, due to the belief that the nature of each of these factors is crucial for program success (Dillon, 2013; Hans, 2000; Rickinson et al., 2004). While failing to replicate findings pertaining to OAE effects on measures of developmental tasks, Vlamis et al. (2011) suggest that careful alignment between programs' structure, philosophy and intended outcomes may be critical to programs' success. Meta-analytic reviews of LoC outcomes support this claim, evidencing that primary purpose therapeutic goals moderate effect sizes (Hans, 2000). Hattie et al. (1997) observed that OAE delivered by Outward Bound Australia led to greater effects when compared to other program providers. Thus, organisational philosophy may comprise a critical program-level factor that influences a program's effectiveness.

Within-person factors. Although the influence of individual differences on OAE learning has received increased attention over the past decade, it remains for the most part an area of research about which relatively little is known. Fry and Heubeck (1998) examined the influence of personality on individuals' affective state during OAE experiences, finding extroversion to be associated with more-positive emotional states. Exploring the nature of coping strategies on the OAE experiences of Australian adolescents, Neill and Heubeck (1998) found that adaptive coping strategies (e.g. problem-solving, working hard) were associated with more positive mental states, whereas maladaptive coping (e.g. avoidance, wishful thinking) constituted the greatest predictor of psychological ill-being.

Examining the influence of cognitive processes on OAE outcomes, Sibthorp, Paisley, and Gookin (2007) found that perceived personal empowerment while on-program

(e.g. “I had important responsibilities on this course”) was associated with positive outcomes in communication and leadership, small group interaction, judgement, outdoor skills and environmental awareness. An earlier study reported similar findings, whereby perceived empowerment mediated the relationship between pre-program expectations and personal effectiveness outcomes (Sibthorp & Arthur-Banning, 2004).

Comparison between effect sizes for normal-stream adolescents and those presenting with behavioural or social-emotional difficulties has yielded little evidence of systematic difference (Cason & Gillis, 1994; Hans, 2000; Neill, 2003; Wilson & Lipsey, 2000). Hattie et al. (1997) found that post-treatment effect sizes differed only marginally between delinquent (ES = .33) and normative (ES = .35) youth samples, while Neill (2003) found OAE programs to yield an overall ES = .31 for normal-stream adolescents, compared to ES = .33 seen in delinquent youth groups. These results indicate that throughout the OAE literature, variations in youths’ baseline functioning do not adequately account for the discrepancies in effect sizes and thus, program benefits do not appear to discriminate on the basis of participants’ pre-treatment psychosocial and emotional functioning.

Psychological needs satisfaction. Only a handful of OAE studies have examined how psychological needs satisfaction may act as a program effectiveness mechanism. The Daniel et al. (2014) review of autonomous OAE student experiences warns against minimising such experiences, due to the central role they play within the maturation process. Wang et al. (2004) applied the principles of SDT to examine motivational predictors of adolescent participation in OAE. They found that external regulations negatively predicted program satisfaction while, inversely, intrinsic motivation predicted greater program satisfaction.

Summary. Outdoor Adventure Education provides valuable learning experiences with the potential to benefit a breadth of positive psychological indicators during adolescence. As research uncovers new learnings on the factors influencing OAE programs' effectiveness, opportunities to refine and enhance these experiences emerge. Three predominant psychological constructs within OAE literature will now be reviewed.

Psychological constructs central to outdoor adventure education.

Self-concept and adolescence. Establishing a working self-concept capable of satisfying the demands of adulthood has long been recognised as a central task during adolescence (Erikson, 1968). Baumeister (1999) defines self-concept as "an individual's belief about himself or herself, including the person's attributes and who and what the self is" (p. 13). Adolescence is often depicted as a stressful and turbulent time (Erikson, 1968). This may be partially attributable to declines in numerous psychological variables such as self-concept that are typically observed from early to mid-adolescence (e.g. Janić et al., 2014; Marsh, 1990a; Molloy, Ram, & Gest, 2011; Watt, 2004), and that may account for the deflated OAE outcome effect sizes seen in this age group. Despite the developmental declines, the possibility for growth in self-concept domains following systematic intervention has attracted considerable interest within educational research (e.g. Marsh, 1990a; Marsh, Richards, & Barnes, 1986b; Marsh et al., 1986c; O'Mara et al., 2006).

Early scholars argued self-concept to be a unitary structure (e.g. Coopersmith, 1967); however, advancements in theory have led researchers to adopt a multifaceted and hierarchical conceptualisation of the construct (Marsh, 1990a; Marsh, Smith, Barnes, & Butler, 1983; Shavelson, Hubner, & Stanton, 1976). At the head of this hierarchy sits General Self-Concept, which subdivides into Non-academic, Academic English and Academic Mathematics domains (Marsh, 1990a). These domains divide further to reflect seven sub-facets of self-concept, as shown in Figure 1. The Marsh et al. (1983) analysis of

the stability and dimensionality of this model among young children demonstrated that student self-concept scores measure specific and separable aspects of self-concept, and that changes occur in some domains independent of change in others. The O'Mara et al. (2006) meta-analysis of 152 school-based self-concept interventions highlighted that those domains most closely aligning with intervention aims showed the greatest change.

The establishment of reliable and valid measures of self-concept (Marsh, 1988a) has resulted in widespread acceptance that this multi-faceted construct is positively associated with healthy adolescent development, with adjustment (Chubb et al., 1997; Fuentes, García, Gracia, & Lila, 2011; Schmidt, Valkanover, & Conzelmann, 2013), personal skills, reduced behavioural problems (Fuentes et al., 2011), educational outcomes such as achievement motivation (Harter, 2006), and both academic persistence and performance (Marsh, 1990a; Marsh, Byrne, & Shavelson, 1988; Marsh, Chanal, & Sarrazin, 2006; Marsh & Martin, 2011; Marsh et al., 1983).

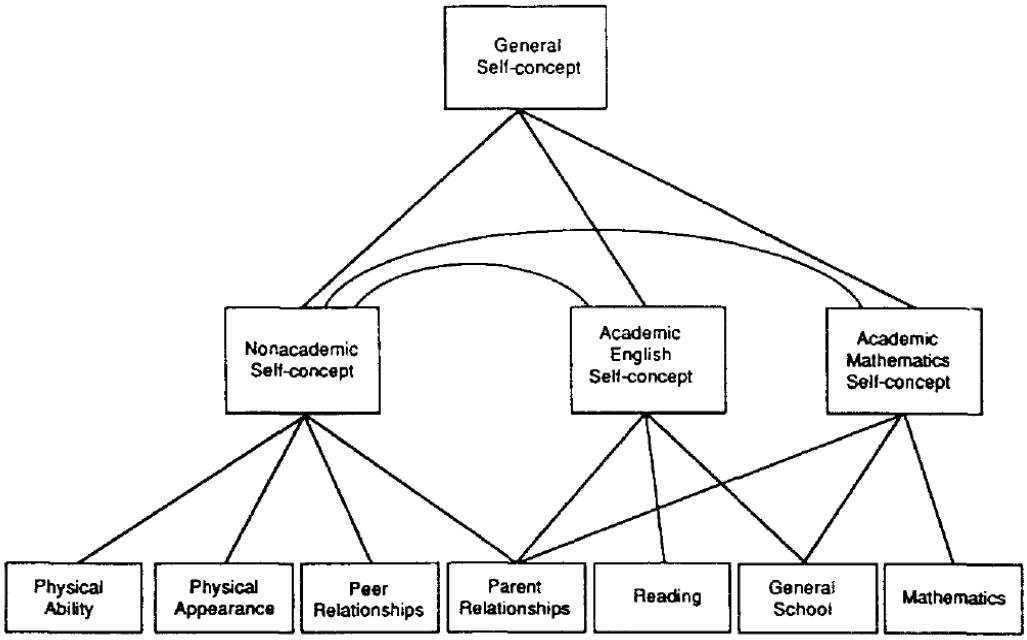


Figure 1. Marsh and Shavelson (1985) empirically derived hierarchical model of self-concept, revised from Shavelson et al. (1976).

Educational settings are crucial to students' developing sense of self (Erikson, 1968; Kaplan & Flum, 2012; O'Mara et al., 2006). It is therefore up to policy makers, educators and researchers to focus on effective treatment programs that embody the synergy between strong theoretical models and rigorous methodological designs, to nurture teenagers' self-development during this pivotal stage of life.

Outdoor adventure education and multidimensional self-concept. Self-concept is one of the most frequently researched constructs within OAE literature due to the construct's potential to benefit from such experiences (Bowen & Neill, 2013; Cason & Gillis, 1994; Hattie et al., 1997; Marsh et al., 1986b, 1986c). OAE programs frequently occur in rich and dynamic social contexts, drawing on strong community ties where cooperation is integral for group success, and where adult leaders encourage prosocial behaviours in students. These experientially dense opportunities, although temporary in nature, provide teenagers with a stage on which to trial behaviours and identity styles, and to observe how others respond. Cooley (1902) termed this the "looking glass self", a central process within identity formation.

Pioneering self-concept researchers, Marsh et al. (1986c), examined the effects of a twenty-six day residential program across thirteen facets of self-concept among adolescent and young adult participants. Post-treatment analysis observed significant enhancements in all domains, including General Academic, Problem Solving, Physical Ability, Same and Opposite Sex Relationships, Parent Relationships, Emotional Stability and General Self domains. The Gass, Gillis, and Russell (2012) review of nine meta-analyses of outdoor and wilderness interventions reported that favourable changes in self-concept was an emerging theme. The Bowen and Neill (2013) review of 197 studies noted that the greatest pre/post outcome effects where observed in self-concept, while Hattie et al. (1997) demonstrated that OAE fosters significant increases in the specific domains of independence, self-

efficacy, confidence and self-understanding. Supported by meta-analytic findings and established theory, the presence of self-concept analysis within OAE literature is well justified.

Personal effectiveness and life skills. Neill, Marsh, and Richards (1997b) define life effectiveness as an individual's capacity to adapt, survive and thrive in order to meet the demands of daily life. The term encapsulates a number of malleable and learnable skills relevant to functioning in key life domains, including personal abilities/self-belief, social performance, organisation and the 'energy' necessary for action-initiation (Richards, Ellis, & Neill, 2002). Nestled within these domains are the following areas, which make up the Review of Personal Effectiveness Scale (Richards & Neill, 2000a) that was specifically designed as a broad measure for experience-based intervention programs:

Active Involvement: The ability to use action and energy to produce outcomes

Cooperative Teamwork: The capacity to work effectively as part of a team

Leadership Ability: The ability to lead others effectively

Open Thinking: One's openness to ideas and cognitive adaptability

Quality Seeking: Ability to apply effort to achieve optimal results

Self Confidence: The belief in one's ability to succeed

Self-Efficacy: Belief one can manage situations and solve difficult problems

Social Effectiveness: Effectiveness in social situations

Stress Management: The ability to respond adaptively to stress and remain calm

Time Efficiency: The ability to use time well and plan efficiently

Coping with change: The capacity to adapt effectively to change

Overall Effectiveness: A person's overall effectiveness across all life domains

Generally, the necessity of developing life effectiveness skills has generated international recognition, and the logical alignment of these skills within interventions that aim to strengthen personal resources, as well as social and personal functioning, has been noted. Given that developing life effectiveness skills is a central aim of OAE (Richards, 1977), life effectiveness has become one of the most prolifically researched psychological variables in OAE literature (Rahman, 2009).

Developing life effectiveness through outdoor adventure education. Outdoor Adventure Education aims to challenge the social, physical and personal capacities of participants, and thus is a promising context for enhancing life skills. The efficacy of OAE in benefiting specific facets of life effectiveness has been widely evidenced within recent literature. Neill (2008) found significant beneficial and lasting changes in active initiation, emotional regulation, cognitive flexibility, self-confidence, social competence, leadership ability, time efficiency and overall effectiveness among adolescent and adult participants. Similar findings have been replicated in other major OAE reviews, which report OAE as benefiting leadership capacities, achievement motivation, emotional stability, and interpersonal domains such as social competence, communication and cooperation (Hattie et al., 1997). Louw et al. (2012) examined the effects of OAE on the life effectiveness of Native African teenagers, again reporting significant increases in self-reported social competence, intellectual flexibility, achievement motivation, emotional regulation, self-confidence and overall life effectiveness at the six-month follow-up. It is consistent with these findings that among teenage participants from an Australian private boys' school (McLeod & Allen-Craig, 2004), and a troubled youth population (Thomson & Burr, 2015), significant improvements in multiple domains of life effectiveness were again observed following OAE.

From the literature, it is clear that OAE has the potential to produce notable and lasting improvements in life effectiveness that span personal, behavioural, and psychosocial domains. This speaks to the notion that OAE fosters holistic growth in participants.

Locus of control during adolescence. The term LoC was first proposed by Rotter (1954) and refers to the generalised attribution of reinforcement through internal or external aspects of the self (Strickland, 1989). The construct is operationalised on a bipolar continuum, moderated by situational factors (Chubb et al., 1997). External LoC (ELoC) entails the perception that reinforcements are the product of factors separate from the self, such as luck, chance or another's will, whereas internal LoC (ILoC) suggests the belief that one's own efforts, capabilities or actions are responsible for the reinforcement (Chubb et al., 1997). Pelletier et al. (1999) denote that amotivation in the context of ELoC stems from the belief that one's actions will not lead to the outcome desired, and/or due to a lack of self-belief that one can successfully enact the behaviour to produce the desired outcome. They conclude:

> When individuals perceive themselves as competent, they express the desire to set optimal stimulating goals for themselves and believe that they have the capacity to pursue and attain those goals through engagement in a particular behavior (p. 2498).

LoC is thus argued to be most closely linked with SDT's basic psychological need for competence, in that more favourable LoC tendencies are positively associated with greater self-efficacy (Ryan & Deci, 2017). LoC as a psychological construct has gained widespread attention within educational literature, due to the influential role of students' perceived LoC on school engagement (Ekstrom, Goertz, Pollack, & Rock, 1986), responsibility, independence, self-control (Lefcourt, 1976), risk-taking (Özmen & Sümer, 2011), and positive adjustment across home, school and social settings (Nunn, 1987).

As adolescents mature, they become more responsible for initiating behaviours that influence their social, physical and emotional health. ILoC ultimately enables students to accept and act on this responsibility. School programs and youth-interventions may capitalise on this natural tendency to buffer against the maladaptive effects of external control expectancies during adolescence (Shehu & Mokgwathi, 2008).

Locus of control and outdoor adventure education. Locus of control is a commonly investigated construct in OAE research (e.g. Cason & Gillis, 1994; Hans, 2000; Hattie et al., 1997). The Cason and Gillis (1994) meta-analysis of OAE programs with adolescents cites thirteen studies reporting statistically significant increases in LoC, with an average effect size of .30. Hans (2000) and Hattie et al. (1997) replicated this finding, reporting average effect sizes for LoC of .38 and .30, respectively. Interestingly, the Davis-Berman and Berman (1994) review of longitudinal effects of OAE on LoC found that immediate post-treatment increases were lost at a four-month follow-up, but had returned one and two years later. Their research, while further exemplifying the potential for OAE to internalise LoC expectancies, highlights the critical need for follow-up evaluation in OAE research.

Shortcomings in the current body of literature. Although the findings of recent research are supportive of a wide range of positive effects of OAE, the body of literature has been scrutinised over manifesting several methodological shortcomings (e.g. Dillon, 2013; Marsh et al., 1986c; Rickinson et al., 2004; Scrutton & Beames, 2015). Rickinson et al. (2004) highlight the need to improve "the methodological rigour of outdoor learning research and evaluation" and the "research-based understandings of the outdoor learning process" (p. 56). They claim that the diversity of findings relating to program design and participant type inhibit convergence of the literature and outcome generalizability. Neill (2008) highlights the limiting factors in quantitative OAE research as including poor

statistical power, the over-reliance on inferential data, the absence of control groups and follow-up data, and the lack of investigation into independent variables.

The Scrutton and Beames (2015) review of twenty-two OAE studies, including six meta-analyses, highlights three problem areas:

1. use of small sample sizes;
2. inappropriate survey design and timing of administration; and,
3. lack of statistical control over extraneous variables.

They conclude that good methodological practice requires: (a) randomized assignment to participant groups; (b) careful selection of measures that meet the specific needs of the research; and (c) appropriate data analysis and clearly reported results. Bowen and Neill (2013) further recommend the use of multiple baseline and follow-up measures to complement pre/post-test evaluation.

Summary. There is little doubt that OAE programming provides valuable opportunities to benefit adolescents' psychological development. However, several significant limitations mean that the largely favourable body of literature has also been characterised as "inconclusive", "over-ambiguous" and "uncritical" (Barrett & Greenaway, 1995, p. 53). Consequently, there is a critical need for further research incorporating strong research designs, including controlled trials and well-defined and theoretically sound measures and statistical methods to address these shortcomings and examine OAE's claimed benefits.

A Call for Greater Psychometric Rigour

Recent decades have seen advancements in psychometric procedures that are designed to enhance measurement quality in applied research settings. Such procedures include the use of reliability testing to assess scales' internal consistency (McDonald, 1999); Structural Equation Modelling (SEM) to establish structural validity and reduce

measurement error (e.g. Marsh, 1994); and multitrait-multimethod (MTMM) analysis, used to examine the convergent and discriminant validity of latent constructs (Campbell & Fiske, 1959). A detailed review of literature on these methods is provided in Chapter 5: Study 1—Psychometric Analysis.

Summary

OAE appears to offer a promising framework for facilitating healthy psychological development in teenagers. However, research into these programs is questioned in regards to methodological rigour, which has prevented forward progress of OAE within Australian educational curriculum. The current research seeks to address these limitations by examining the outcomes from an OAE program that is incorporated into the curriculum of an independent boys' school in Sydney. The following chapter details the OAE program that was the intervention studied in this research.

CHAPTER 3: THE INTERVENTION PROGRAM

The Research Context: The Glengarry Program

Program background, goals and philosophy. Since 1988, The Glengarry Program (TGP) has provided an OAE program designed to build resilience and practical skills in Year 9 students. Encapsulating OAE's focus on the development of skills in a natural environment, TGP takes place well away from the students' family homes, in a bush setting that is highly restrictive of electronic media (mobiles, TV and computer games). Consistently with findings that longer OAE interventions have more-positive effects (Hattie, 2009), TGP is five-months duration. The program thus has a rich history of being well-integrated into the single-sex school's educational environment with the requisite strong commitment from the whole school community that is important for achieving program outcomes (Dillon, 2013). A consistent theme in OAE research is that the individual OAE program's philosophy and content are important moderators of their success (Dillon, 2013; Rickinson et al., 2004). Therefore, closely integrating OAE into the educational environment appears to be critically important to determine what effect sizes can be achieved. In the case of TGP, as stated in the TGP parent handbook (The Scots College, 2016), the OAE experience:

> Aims to foster growth by helping adolescent boys effectively navigate the transition to manhood through immersion in community, academic, spiritual and outdoor adventures. That is, to develop:
>
> 1) People who are competent, collaborative, creative, caring and dedicated;
>
> 2) Practices that are effective, safe, sustainable, and demonstrate environmental stewardship;
>
> 3) Student outcomes that promote the growth of responsible individuals who are critical and creative thinkers, cooperative team players, effective communicators and who demonstrate autonomous learning. (p. 4)

The Glengarry Program, embedded in experiential learning theory, constitutes best practice in terms of the design standards recommended in major OAE reviews (e.g. Cason & Gillis, 1994; Dillon, 2013; Hattie et al., 1997; Neill, 2008; Rickinson et al., 2004; Wilson & Lipsey, 2000). The following section details the core aspects of the program as they relate to key dimensions of adolescent development.

The OAE environment: An holistic approach to adolescent development.

The physical environment. The Glengarry campus includes five dormitories, each housing twenty-two students, which include a common room, sleeping area, equipment storage area, laundry and bathroom. Each dorm also contains a hardwired emergency phone with 24hr access to designated staff-on-duty and has a staff member residing in attached, but separable accommodation during the evenings. The library holds an extensive fiction collection aiming to cater to individual interests. A large multi-purpose hall includes a covered sports space (basketball, soccer etc.), a weights area, bike storage and indoor climbing walls. The campus has five formal classrooms with wireless internet access, dining hall and kitchen facilities, music practice rooms, a theatrette used for group meetings and a clinic staffed daily by registered nurses.

The academic program. The academic program at TGP aims to develop students' strengths using a breadth of experiential learning opportunities. The program combines the curriculum imperatives of the NSW Education Standards Authority with outdoor education and a long stay residential program. The academic timetable comprising of 4 x 1.5 hour periods, 5 days/week includes compulsory classes of English, Math, Science, History, Geography, Personal Development, Health and Physical Education (PDHPE), and Divinity. Students also choose two electives from a range of subject choices. Academic support is available twice weekly after school and class sizes are approximately 20 students.

Physical health and development. Central to TGP is a focus on physical activity and skill development, but also the establishment of healthy lifestyle behaviours. During the program, the students experience healthy eating and exercise. Students' set their own exercise regimens and long-term goals aiming to encourage healthy lifestyle behaviours (The Scots College, 2016). The program integrates a carefully designed dietary routine reviewed by independent health specialists, designed to correspond to the activities offered. A range of psychoeducational topics regarding mental health and physical development are integrated into the curriculum, while on-screen time is restricted to laptop access during academic classes and evening email time (i.e. students otherwise have no access to mobile phones, T.V., video games, social media platforms etc.).

While attending TGP, students choose from a range of physical activities. Interdorm run/ride challenges occur twice weekly, elective co-curricular activities once weekly, and overnight hikes take place on weekends. Emphasis is placed on students setting themselves personal performance goals in relation to their varied experiences, which include, but are not limited to, rock climbing and abseiling, camping and orienteering, canoeing/kayaking, hiking, caving, surfing, mountain biking and horse riding. In most cases, the outdoor activities are run in dorm groups (approx. 20 students), however, some orienteering challenges and hikes also take place in smaller groups of approximately four to seven students. When students are engaged in unsupervised wilderness activities, each group has with them a walkie-talkie with GPS tracking that is monitored by staff, and students are required to attend periodic check-in points. The level of challenge for wilderness activities is carefully graded to align with students' developing confidence and skill over the five-month experience.

The social climate. Establishing positive relationships with staff and peers is a central focus at TGP. The Scots College has a strong boarding community consisting of

rural and international students, so at TGP, students have the opportunity to mix with boys of different backgrounds. This cross-section of society exposes students to a wide range of cultural experiences and enables them to establish meaningful friendships. The Glengarry program teaches students to live with others in a community setting, while maintaining family bonds through regular letter and email correspondence, and scheduled family visits approximately twice during each school term. The residential environment is described as providing the opportunity to experience the challenges of living, studying and learning in teams and in a community setting in which the following qualities are emphasised:

- Personal development and leadership through the continual pursuit and achievement of "personal bests" and "community bests";
- Skills of teamwork and cooperation;
- The attainment of self-discipline;
- Values and traditions.

Pastoral care and student support. The Glengarry Program intends for all students to be challenged socially and emotionally during their five-month OAE experience, however, staff acknowledge the magnitude of challenge experienced by any individual student will vary substantially. A multi-level pastoral support system is thus maintained to ensure appropriate supports is available for different student needs. The residential Dorm Masters provide initial counsel to students and further counselling interventions serviced by the College's in-house counselling team, or otherwise through external referral streams, can be implemented on a needs-basis in consultation with parents. Such services are equally available for students attending the main-school campus. The students are supported by a dorm tutor who follows their progress and facilitates their learning in the important area of their growth. Students can also consult with the campus pastoral care coordinator or otherwise, their teachers, outdoor staff, or nursing and administration staff.

Evidently, all staff at TGP have a level of responsibility when it comes to student wellbeing. Additionally, a registered psychologist from the College Counselling team attends TGP on an approximately fortnightly basis.

The developing 'self', personal responsibility & life skills. Intentionally integrated into the program are opportunities for personal reflection. This pedagogical tool is intended to facilitate the articulation and integration of students' learning. This is often structured (e.g. as a weekly written/video journaling), or can otherwise take less-prescribed forms, as with the three-day solo hike, a final expedition where students spend a period of time autonomous in the bush during which they exercise resourcefulness and self-reliance.

At TGP, students are taught to recognise that their actions have consequences and thus the program highlights the notion of accountability, with the aim of fostering a strong sense of self-discipline. Students are encouraged to develop an expanded world view without the use of mobile phones, video games and TV. Doing so helps students to take ownership over all aspects of their development through goal setting and the establishment of healthy routines.

TGP aims to enable students to face their fears and extend themselves beyond their perceived capacities (The Scots College, 2016). Student adopt daily responsibilities for campus sustainability by undertaking domestic tasks and lifestyle activities, including indoor/outdoor cleaning, food preparation, wilderness environmental maintenance and cleaning classrooms, recreational and personal sleeping/living areas.

Fostering students' basic psychological needs. As Wang et al. (2004) state, in order to meet the basic psychological needs of students, OAE programs must provide students a meaningful rationale for their participation, opportunities to pursue interests, and the experience of growth in competence.

Analysis of TGP suggests that the program is well equipped to meet the psychological needs of participants. Students are provided with elective opportunities and encouraged to believe they can drive the benefits of the program for themselves. This focus on autonomy is further emphasized by the leadership focus in the program and by exposing students to the philosophy and rationale for TGP upon enrolment. Students' competence is built using tasks situated in the natural environment that challenge students physically, socially, emotionally and academically. Staff provide positive feedback on students' progress aiming to challenge their personal capacities in support of goal attainment and mastery of new skills. Finally, there is a strong emphasis on social support, where students learn to live in a community setting with their peers and campus staff. Student-parent relationships are also supported through parents' inclusion in the program. It is, however, important to note that TGP was not explicitly designed with SDT in mind. The principles of basic psychological needs were applied to the existing OAE program to provide a means of investigating the underlying mechanisms of change following participation in the outdoor experience.

Summary

TGP offers a best-practice OAE program which appears to be designed to meet students' basic psychological needs. However, like many OAE programs, it has never been empirically tested in relation to its goals and to the moderating role of these psychological need-related supports. This research explores the potential for TGP to foster positive psychological outcomes in youths, and the mechanisms through which they come about. The following chapter details the methodology that was used to conduct the present investigation.

CHAPTER 4: GENERAL METHOD

Introduction

This investigation endeavours to examine the link between Outdoor Adventure Education (OAE) and positive mental health outcomes among male youth. The present investigation, adhering to the recommended design standards outlined in leading OAE reviews (see Chapter 2: Literature Review), embraces a synergy between theory, empirical research and a strong intervention model to achieve this research objective. The purpose of this chapter is to provide a comprehensive description of the overarching methodology of this investigation, including research sample, design and statistical processes, instrumentation, and data collection schedule. There were variations in procedure, shaped to accomplish the Study 1 and Study 2 objectives, but this chapter specifies the methodology that is consistent across the whole project. Each study entails a method, embracing procedures necessary to satisfy its objectives (see Chapter 5: Study 1 Psychometric Analysis, and Chapter 6: Study 2 Intervention).

Participants

A total of 413 male students (waitlist control $N = 202$; treatment condition $N = 211$) enrolled in the 2015 and 2016 Year 8 cohorts at an urban high school were invited to participate in this research project. All students were attendees at a non-selective New South Wales independent boys school located in the Sydney metropolitan region. As reflected in Table 1, the participation rates differed across the data-points, as some students were either absent or otherwise elected not to participate in some instances. The sample size was fixed by the design of the study and could not be modified.

Table 1 *Participation Frequency Distribution Based on Treatment Condition at Each Data-Wave*

Time	Research Sample Frequency					
	Intervention		WLC		Total	
	N	%	*N*	%	*N*	%
T0	182	90	164	81	346	84
T1	194	92	184	91	378	92
T2	192	91	188	93	380	92
T3	156	74	--	--	156	74

Note. The left column signifies the data collection time-wave: T0 = Extended baseline (10 weeks prior to pre-test), T1 = Pre-test, T2 = Post-test, and T3 = Follow-up. WLC = Waitlist Control Group. T3 relates only to treatment group participant data, as the WLC at that point was no longer viable to act as a control sample as they had embarked on the OAE program.

A post hoc power analysis was conducted using G*Power3 (Faul, Erdfelder, Lang, & Buchner, 2007) to test the difference between two independent group means using a two-tailed test, a small effect size (ES = .25), and an alpha of .05. Results with a total sample of 346 participants (treatment = 182; WLC = 164) achieve a strong power of .996.

Materials

The item pool of 134 items was developed using the Qualtrics survey development tool (2015) to derive the online and hardcopy questionnaire formats (see Appendix 2: The Research Survey Instrument). Each instrument was considered on the basis of its evidenced suitability to youth samples, and feedback from key personnel involved in the implementation of the OAE program. This ensured a selection of outcome measures appropriate to the research aims and OAE program philosophy and goals. As part of this process, some scales were adapted for their relevance to the secondary school environment. All modifications are specified in the description of quantitative measures provided later in this chapter.

Research Design

A quantitative approach was adopted to test all the research hypotheses. This was deemed a vital design aspect to fill a much-needed gap in OAE literature and complement

an already established body of qualitative research (e.g. Bell & Holmes, 2011; Dolgin, 2014; Gass, Garvey, & Sugerman, 2003; Zink, 2010). Study 1 conducted a detailed psychometric evaluation of all instrumentation, while Study 2 followed an experimental design with a randomly assigned waitlist control condition (WLC) to empirically measure the short- and long-term effects of the 5-month OAE program. A detailed account of the methodology for each study is provided in the respective Chapter 5 and Chapter 6 sections. Figure 2 illustrates the research survey administration schedule and timeline for the project.

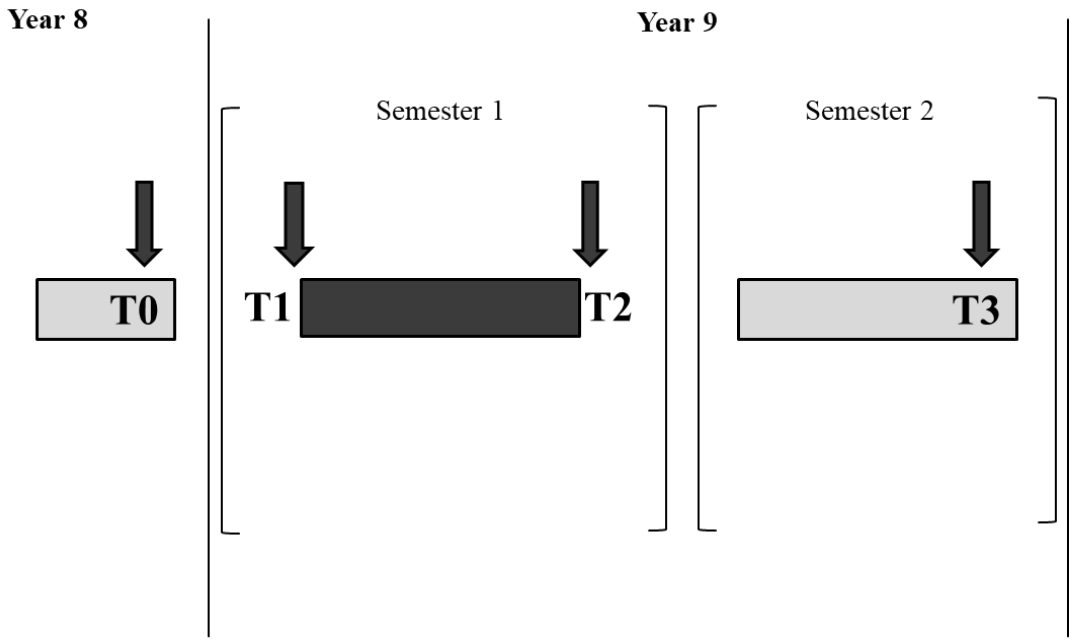


Figure 2. Graphical representation of data collection points in relation to high-school semester and student grade.

Note. Rectangles represent school semesters, while the text and vertical arrows indicate the four data collection points: T0 = Extended-baseline, T1 = Pre-test, T2 = Post-test, T3 = follow-up. The period of the 5-month OAE intervention for the treatment group has been shaded dark grey, while light grey rectangles indicate usual school timetable.

Ethics Procedures

Consent to undertake the present research was obtained from the Australian Catholic University Human Research Ethics Committee (Ethics Register Number: 2015-245H; see Appendix 1), and The Scots College internal research department and executive board. As all participants were under the age of 16 when the research was conducted, parental/guardian permission was required for all students. Consent procedures followed opt-out methods, whereby parents/guardians were required to specify if they did not wish their child to participate. Students with parental/guardian permission were invited to participate voluntarily in the study. This procedure was completed for all participants prior to the administration of each questionnaire.

Data Analysis

The following section overviews the data analytic methods which were applied to undertake the present investigation.

Statistical software. Initial data screening and preliminary analysis (reliability, frequencies, descriptive statistics etc.) was undertaken using SPSS Statistics 22.0 (IBM Corp., Released 2013). More advanced statistical processes were conducted using Mplus Version 8.0. (Muthén & Muthén, 1998-2017) with a maximum likelihood robust estimator (MLR), protecting against violations of normality assumptions. The specific application of each software package is detailed in the respective Chapter 5 and Chapter 6 method sections.

Deriving latent factor scores. The present investigation utilised factor score regression weights rather than raw composite scores to reflect students' responses for any single scale at each of the four waves (T0, T1, T2 and T3). These factor scores reflect an aggregation of indicators, or items, in a scale that is weighted according to the item loadings, as described by Rowe and Rowe (1999). This method minimises the contribution

of measurement error to each composite score by the respective indicators, and thus enhances the reliability and validity of computed scores. The present investigation derived factor scores for each latent construct, following the structural equation modelling (SEM) procedures outlined in Chapter 5: Psychometric Analysis and applied these to all analyses in Study 2.

Treatment of missing data. A common issue in longitudinal research is the matter of missing data, which may result from sampling attrition, error on the part of the investigators, or chance oversights by respondents during survey administration. Given the presence of missing data in the present investigation (see Table 2), some scales received fewer responses than others. To impute missing data, Full Information Maximum Likelihood (FIML) estimator was selected. This decision was based on FIML, which comprises the most robust approach in SEM procedures (Enders, 2008, 2010; Graham, 2003) and arguably as the most suitable identification method in applied Exploratory Structural Equation Modelling (ESEM) research (Marsh, Morin, Parker, & Kaur, 2014).

Table 2 *SDs and Average % of Missing Data for the 134-Items at Each of the Four Data Collection Points*

Time-Wave	Response Rate Frequency	
	SD	%
T0	1.21	4.55
T1	1.18	4.56
T2	1.18	4.66
T3	1.16	4.71
Mean	1.17	4.63

Note. Each time-wave indicated in the column on the left reflects the response rate frequencies for the combined 2015 and 2016 Year 8 cohorts.

The FIML method makes full use of a data-set by attempting to recover missing information (Enders, 2010); a capability not apparent in more traditional methods such as listwise and pairwise deletion during SEM analysis (Allison, 2003). For these reasons, FIML was employed for the management of missing data in the present investigation.

Pilot Survey and Instrument Validation

A pilot administration (T0) of all quantitative measures was conducted with the full research sample prior to the pre-treatment (T1) administration. This step was considered imperative given that some scales had not been adequately validated in longitudinal research involving representative adolescent participants. Additionally, consideration of theory and substantive factors led to the adaptation of some scales' original form. The main aim of the pilot survey therefore was to determine the suitability of the survey length and to conduct preliminary psychometric assessment of each instrument.

The Quantitative Measurement Tool

The present investigation draws on psychometrically established and theoretically sound quantitative scales to achieve the research aims and explicate OAE's claimed benefits (see Appendix 2: The Research Survey Instrument). A detailed description of these scales and their application in the present investigation will now be provided.

Self-concept. The Self-Description Questionnaire-II (SDQ-II; Marsh, 1990b) was selected to measure self-concept in the present research. The instrument was established with the intention of offering a reliable and valid assessment of self-concept in adolescents aged 12 to 18 and was derived from the Shavelson et al. (1976) multidimensional and hierarchical representation of self-concept. The SDQ-II originally consists of 11 factors and 51 items, 20 of which are negatively worded, however the present research selects only five factors (see Table 3) from this instrument, on the basis of their careful alignment with the aims, philosophy and setting of the OAE intervention. This decision was guided

by meta-analytic findings from 152 school-based self-concept interventions (O'Mara et al., 2006) showing that factors most closely aligning with program aims reflect the greatest change, demonstrating the important interplay between theoretical and substantive factors in research design. Applying this alignment in the present investigation reduced survey completion time, cognitive demands and associated fatigue, while overcoming criticisms relating to inappropriate selection of measurement tools within OAE research contexts (e.g. Marsh et al., 1986c; Scrutton & Beames, 2015). As the research sample was male only, the intentional exception to this alignment was the inclusion of a sixth SDQ-II measure of Opposite-Sex Relations, a scale used to act as a control measure for post-treatment biases, as described by Marsh et al. (1986c). This is elaborated on in the Chapter 6: Study 2--Intervention method section.

In addition to the six SDQII scales, the original eight item Problem Solving scale was adapted from the SDQIII, as shown in Table 3 (Marsh, 1988b). The SDQIII is the final instrument in the SDQ series, and was designed to measure self-concept in late adolescence-early adulthood. The *a priori* structural model for the SDQ instrument used to measure self-concept in this research is represented in Figure 3.

All items were scored on a 6-point Likert response scale, as used in the original SDQ instrument protocol. The response descriptors consisted of the following labels in order to standardise the response process within the survey tool:

1 = *Disagree a lot* 2 = *Disagree* 3 = *Disagree a little*

4 = *Agree a little* 5 = *Agree* 6 = *Agree a lot*

The SDQ instruments have been subject to substantial psychometric analyses and construct validation, leading the scales to be deemed among the best measures of multidimensional self-concept (e.g. Byrne, 1996; Gilman, Laughlin, & Huebner, 1999; Leung, Marsh, Yeung, & Abduljabbar, 2015; Marsh, 1990a; Marsh et al., 1983).

Furthermore, the instruments have been used frequently to evaluate the effectiveness of school-based settings (see Hattie, 2009; Hobby, 2015; Schmidt et al., 2013), due to the distinct multidimensionality of the measure. This provides a strong basis for assessing interventions targeting specific domains of self-concept in adolescents.

Table 3 *Summary Description of the Self-Description Questionnaire II/III (SDQII/SDQIII) Scales*

Scale	Description	Example Item
General Academic	Student rating of their skills and ability in school subjects in general	"I get good marks in most school subjects"
Physical Ability	Student ratings of their skills and interest in sports and physical activities	"I am good at things like sport, gym, and dance"
Global Esteem	Student ratings of themselves as effective, capable individuals, who are proud and satisfied with the way they are	"A lot of things about me are good"
Problem Solving	Student ratings of their ability to solve novel problems and challenging tasks	"I can figure out unusual answers to new problems easily"
Parent Relations	Student ratings of how well they get along with their parents and the quality of their interactions	"My parents treat me fairly"
Opposite-Sex Relations	Student ratings of their popularity with members of the opposite sex and ability to make opposite-sex friendships	"I have lots of friends of the opposite sex"
Same-Sex Relations	Student ratings of their popularity with members of their own sex and ability to make same-sex friendships	"It is difficult to make friends with members of my own sex"

Note. The Problem Solving scale was taken from the SDQIII instrument, while all others came from the SDQII. Using a six-point Likert scale, 1 = *Disagree a lot* – 6 = *Agree a lot*, participants indicated to what extent they felt each item applied to them.

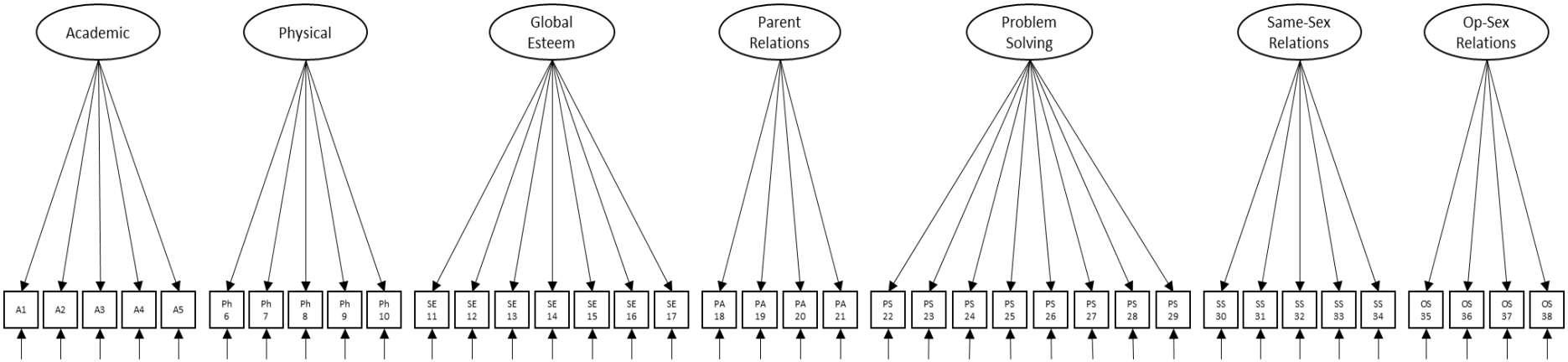


Figure 3. Factor structure of the SDQII/III subscales used in the present investigation.

Note. Academic = General Academic, Physical = Physical Ability, Op-Sex Relations = Opposite-Sex Relations. All factors are from the SDQII instrument, apart from the Problem Solving domain, which is from the SDQIII.

Reliability analysis of the SDQII instrument among a normative Australian adolescent sample (n = 5,494) from the greater metropolitan Sydney region evidenced internal consistency estimates for General Academic, Physical Ability, Global Esteem, Parent Relations, Opposite-Sex relations and Same-Sex Relations ranging between alpha = .85 to .90, while the SDQIII instrument for late adolescence aged 15 years and up demonstrated internal consistency for the Problem Solving scale at alpha = .84 (Marsh, 1989).

Life effectiveness and locus of control. The term "life effectiveness" refers to one's capacity to adaptively meet the demands of daily life in order to be successful (Neill, Marsh, & Richards, 1997a) and encapsulates acquirable skills spanning key areas of personal functioning. Outdoor adventure education is widely endorsed as aiming to enhance individuals' personal resources, development and psychosocial functioning (Richards, 1977), generating a natural alignment between life effectiveness constructs and OAE research such as the present investigation.

The Review of Personal Effectiveness with Locus of Control (ROPELOC) scale was developed by Richards et al. (2002) with the intent of providing a robust measure of multiple facets of life effectiveness in adventure and experiential-based settings. The scale was adapted from the Review of Personal Effectiveness Scale (ROPE; Richards & Neill, 2000b), which was preceded by the Life Effectiveness Questionnaire (LEQ-H; Neill et al., 1997b). The ROPELOC differs from its predecessor in having been extended to include a bi-dimensional scale of locus of control (see Table 4).

The 45-item ROPELOC scale contains 14 factors and a control scale. However the control items were excluded from the research instrument, due to the inclusion of the Opposite-Sex Relations self-concept scale, which performed an equivalent function. The External Locus of Control (ELoC) scale was reverse-scored to align with all other

instrument scales. Students responded to each item using the original 8-point Likert scale, ranging from 1 = "*False / Not like me*" to 8 = "*True / Like Me*".

Table 4 *Summary Description of the ROPELOC Scales*

Scale	Description	Example Item
Self Confidence	Confidence and belief in one's ability to be succeed	"I am confident in my ability to be successful"
Self-Efficacy	Ability to handle things well and overcome difficult situations	"No matter what the situation is I can handle it"
Stress Management	Self-control and calmness in stressful situations	"I am calm in stressful situations"
Open Thinking	Openness and adaptability with thinking and ideas	"I am open to new thoughts and ideas"
Social Effectiveness	Competence and effectiveness when operating in social situations	"I communicate effectively in social situations"
Cooperative Teamwork	Cooperating in team situations	"I like cooperating in teams"
Leadership Ability	Leadership capability	"I am seen as a capable leader"
Time Efficiency	Efficient planning and use of time	"I plan and use my time efficiently"
Quality Seeking	Applying effort to achieve the best possible results	"I try to get the best possible results when I do things"
Coping with Change	Ability to cope well with change	"I cope well with changing situations"
Active Involvement	Using action and energy to produce results	"I like being active and energetic"
Overall Effectiveness	The overall effectiveness of a person in all aspects of life	"Overall, in my life I am an effective person"
Internal Locus of Control	Extent to which students agree that life determinants exist within their control	"My own efforts and actions are what will determine my future"
External Locus of Control	Extent to which students believe life determinants reside outside of their control	"Luck, other people and events control most of my life"

Facets of life effectiveness, such as those featuring in the ROPELOC instrument, are frequently found in research concerning the impact of OAE programs on adolescent life-skills (e.g. Louw et al., 2012; McLeod & Allen-Craig, 2004; Neill, 2008; Thomson & Burr, 2015). Evaluation of internal consistency estimates of ROPELOC domains in an Australian school sample demonstrated Cronbach's alphas ranging between .71 to .89 (Ellis, Marsh, & Craven, 2009). As such, the ROPELOC was selected as an appropriate evaluative tool for the key outcomes of life effectiveness and LoC in the present investigation.

Basic psychological needs satisfaction. Central to Self Determination Theory (SDT) is the assumption that all humans have innate and universal basic psychological needs that are essential to human thriving (Ryan & Deci, 2000a, 2002). The Basic Psychological Needs Satisfaction Frustration Scale (BPNSFS) was developed by Chen et al. (2015) to provide a structurally valid and reliable cross-cultural measure of psychological need satisfaction, as posited within SDT (Deci & Ryan, 2000; Ryan & Deci, 2000a). The original instrument comprises 24 items, 12 of which measure satisfaction, while the other 12 measure frustration of SDT's three basic psychological needs, of autonomy, relatedness and competence (see Table 5).

Psychometric evaluation of psychological needs satisfaction compared to need frustration has led to general agreement that these constructs are independent of one-another (e.g. Bartholomew et al., 2011; Bartholomew, Ntoumanis, & Thøgersen-Ntoumani, 2010; Chen et al., 2015; Rodríguez-Meirinhos, Antolín-Suárez, Brenning, Vansteenkiste, & Oliva, 2019).

Table 5 *Summary Description of the BPNSFS Satisfaction Scales*

Scale	Description	Example Item
Autonomy	To have willingness, volition and self-determination when undertaking an activity	"At school, I feel a sense of choice and freedom in the things I undertake"
Relatedness	The experience of interpersonal intimacy and genuine connection to others	"At school, I feel close and connected with other people who are important to me"
Competence	To feel effective and capable of achieving desired outcomes	"I feel confident that I can do things well at school"

The present investigation used the three basic psychological need satisfaction scales comprising 12-items in total. These were adapted by reframing item wordings to reflect the school context. For example, "I feel capable at what I do" was worded, "At school, I feel capable at what I do" (see Appendix 2: The Research Survey Instrument). Students were asked to, "consider your last school term" and respond using a 6-point Likert scale, ranging from 1 = "*Disagree a lot*", through to 6 = "*Agree a lot*". Given the OAE program entails a compulsory school program operating within the academic-term format, it seemed reasonable that the use of the word 'school' would also apply to the OAE program for attendees of the OAE program. The *a priori* factor structure of the BPNSFS satisfaction scales is presented in Figure 4.

The evaluation of psychological needs satisfaction among adolescent samples is well documented (see Gnambs & Hanfstingl, 2015; Thomaes, Sedikides, van den Bos, Hutteman, & Reijntjes, 2017). A large university study examining the structural validity and reliability of the BPNSFS demonstrated structural cross-cultural equivalence across independent samples from USA, China, Peru and Belgium, while internal consistency estimates for Autonomy, Relatedness and Competence ranged from alpha = .74 to .88 (Chen et al., 2015).

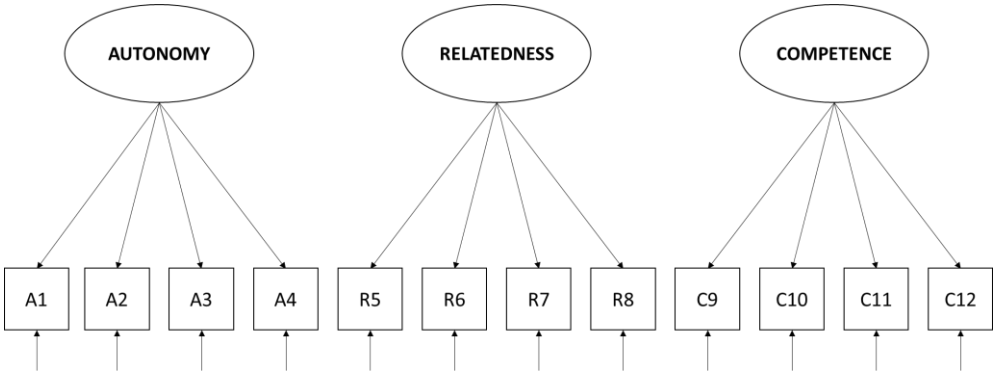


Figure 4. A priori factor structure of the BPNSFS (Chen et al., 2015); satisfaction items only.

Psychological illbeing. The General Health Questionnaire-12 (GHQ-12; Goldberg, 1972) was selected to measure psychological distress in the adolescent sample. The GHQ-12 concerns individuals' symptomatology of the past few weeks, so to screen for breaks in normal functioning, rather than to detect long-term pathology. Such breaks in functioning measured by the GHQ-12 may be associated with symptoms of anxiety, stress, and/or depression. While the factor structure of the GHQ-12 has generated some debate, a two-factor solution (see Figure 5) of general psychological health for Positive (social emotional functioning) and Negative (psychological distress) symptoms of psychological ill-being poses a dominant and structurally validated model (e.g. Abubakar & Fischer, 2012; Andrich & Van Schoubroeck, 1989; Makikangas et al., 2006; Molina, Rodrigo, Losilla, & Vives, 2014; Shevlin & Adamson, 2005; Ye, 2009). A large international study (N = 26,120) evidenced these two factors as accounting for 58% of total model variance explained (Werneke, Goldberg, Yalcin, & Ustun, 2000).

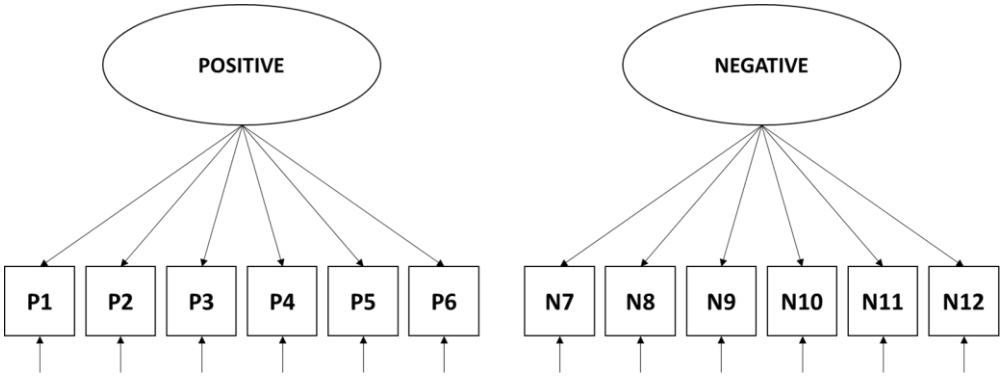


Figure 5. Two-factor structure of the GHQ-12.

Students responded to the GHQ-12 items using the original 4-point multiple choice response scale (see Appendix 2: The Research Survey Instrument). Items were scored by following the Likert scale scoring method (0-1-2-3), given that this approach provides a more smooth score distribution that allows for greater assessment of severity, compared to the original GHQ-12 system (0-0-1-1; Goldberg et al., 1997). Reverse scoring methods were applied to both scales, and students were imputed an overall weighted score for each of the two factors ranging between 0-18, whereby higher values indicated reduced psychological distress.

Personal Resource Variables

OAE programs typically focus on character development and on building mental toughness as major outcomes (Hattie et al., 1997; Neill, 2008; Rahman, 2009; Wilson & Lipsey, 2000). They are congruent with positive psychological theory in that OAE typically avoids emphasising a participant's weaknesses or pathology, and instead aim to enhance positive health outcomes, such as well-being, resilience, and personal growth (Kashdan & Ciarrochi, 2013). In addition to quantitatively examining key outcome variables, the present investigation aimed to evaluate changes in personal psychological resources including motivation and engagement, life satisfaction, resilience, well-being and gratitude. The instruments used to measure these outcomes are described below.

School motivation and engagement. The Motivation and Engagement Scale (MES) was developed as a partner scale to the theoretically integrative multidimensional Motivation-Engagement framework pertaining to behavioural and cognitive aspects of academic motivation and engagement in adolescents (Martin, 2007, 2009). Atop the MES factor structure are four higher-order factors, including cognitive and behavioural Boosters (adaptive cognitive and behaviour), Mufflers (impeding cognitive) and Guzzlers (maladaptive behaviours). Below these sit 11 lower-order factors, each consisting of four items (see Figure 6), resulting in a 44-item instrument.

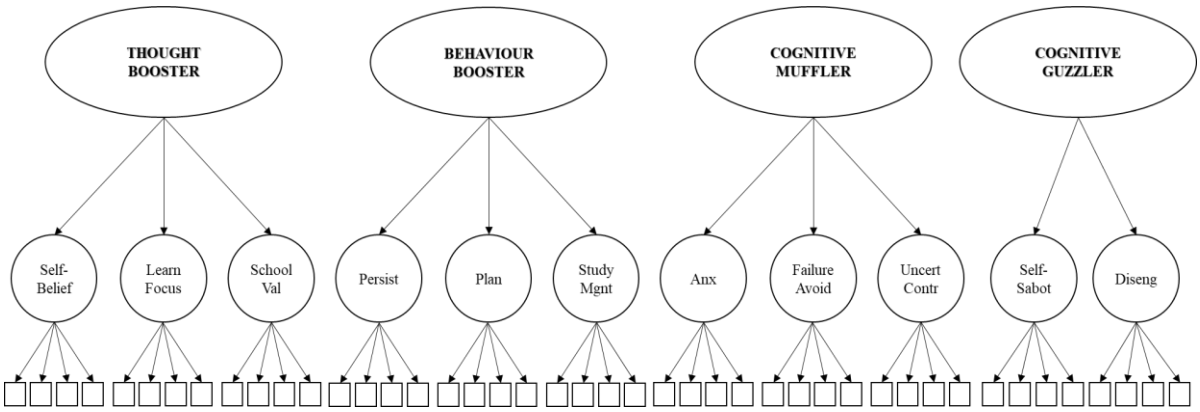


Figure 6. MES *a priori* higher-order factor structure.
Note. Learn Focus = Learning Focus, School Val = Valuing School, Persist = Persistence, Plan = Planning, Study Mgnt = Study Management, Anx = Anxiety, Failure Avoid = Failure Avoidance, Uncert Contr = Uncertain Control, Self-Sabot = Self-Sabotage, and Diseng = Disengagement.

Despite substantial support for the underlying theory of the MES, a validation approach to the MES *a priori* measurement model, using an adolescent sample, revealed an inadequate fit to the population (Reschly, Betts, & Appleton, 2014). This research thus refines the original MES form to an 11-item, four-factor model, whereby each original higher-order factor represents first-order factors, and the original 11 first-order factors are represented by a single item selected from the original pool of 44-items. This refined *a*

priori model is illustrated in Figure 7, and is firmly grounded in the empirical foundation of the Martin (2009) MES instrument (see Appendix 2: The Research Survey Instrument).

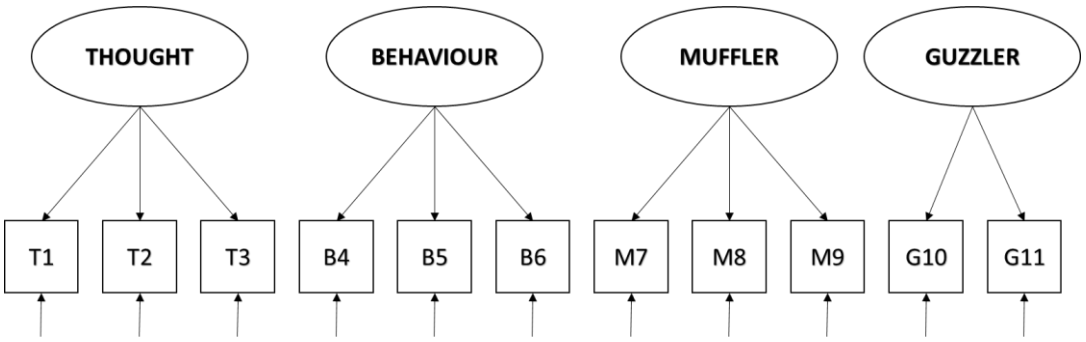


Figure 7. Revised four-factor structure for the MES.

Overall satisfaction with life. Students' life satisfaction was measured using a single overall satisfaction with life (OSL) indicator taken from the Personal Well-being Index-School Children scale (PWI-SC, Cummins & Lau, 2005). The item asked participants to indicate "How happy are you with your life as a whole", responding using a 10-point end-labelled scale ranging from 1 = "Very Sad" to 10 = "Very Happy". Measures of OSL have been applied frequently within youth research (e.g. Casas, Bălţătescu, Bertran, González, & Hatos, 2012; Casas & Rees, 2015; Casas et al., 2011; Savahl et al., 2017). Campbell, Converse, and Rodgers (1976) note the importance of incorporating single-item measures of OSL in well-being research, with such indicators having utility within convergent validity assessment (Casas & Rees, 2015).

Lucas and Donnellan (2012) applied multivariate latent state-trait methods to a large sample of longitudinal data for the single-item OSL measure to estimate univariate and bivariate models, reporting reliabilities of .61 and .74, respectively.

Unidimensional well-being. Recent decades have brought about a surge of interest in, and awareness of, community health. This has led to greater attention being

placed on well-being initiatives in educational settings. One predominant instrument used to evaluate subjective well-being and psychological functioning is the Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS; Stewart-Brown et al., 2009), a seven-item unidimensional measure scored on a 5-point Likert scale (see Figure 8). The SWEMWBS is an abbreviated form of the WEMWBS (Tennant et al., 2007) that examines positive aspects of mental health including eudemonic (e.g. healthy functioning) and hedonic (e.g. happiness) forms of well-being.

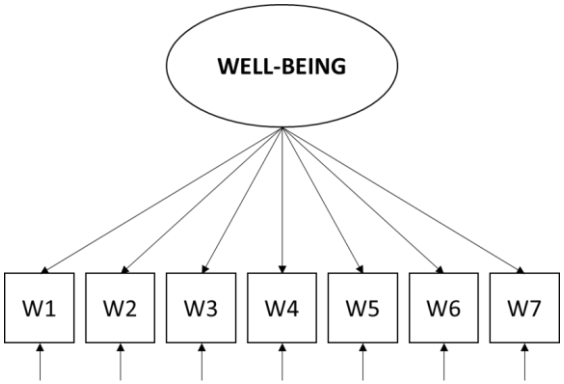


Figure 8. *A priori* factor structure of the SWEMWBS

The SWEMWEBS instrument has been validated in adolescent populations (see Ringdal, Bradley Eilertsen, BjΓërnsen, Espnes, & Moksnes, 2018; Stewart-Brown et al., 2011), including a large Australian sample where the single-factor scale evidenced high internal consistency of alphas = .86 (Hunter, Houghton, & Wood, 2015).

Gratitude. Gratitude is described as a positive affective state measured by the frequency and intensity with which one recognises and responds to people or situations with grateful emotions (Langer, Ulloa, Aguilar-Parra, Araya-Veliz, & Brito, 2016). The construct has gained growing attention in positive psychology research, due to the construct's inverse relationship with neuroticism (McCullough, Emmons, & Tsang, 2002; Simon, 2016), potential to buffer against negative symptoms of stress and depression, and

association with increased perceived social support (Wood, Maltby, Gillett, Linley, & Joseph, 2008), prosocial behaviour, social integration and satisfaction with life in adolescents (Froh, Bono, & Emmons, 2010).

The Gratitude Questionnaire (GQ-6) presents a psychometrically strong unidimensional measure of gratitude (McCullough et al., 2002). The original scale contains six items. However, a psychometric validation study by Langer et al. (2016) demonstrated a five-item version of the GQ-6 showed best fit within an adolescent sample, with internal consistency reaching alpha = .83.

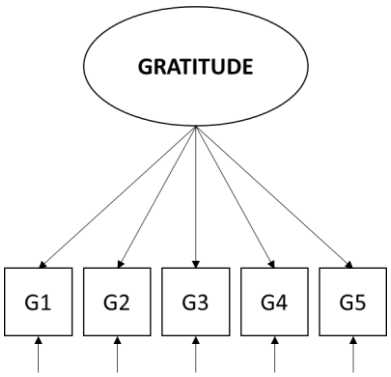


Figure 9. Adapted factor structure of the GQ-6.

This five-item adaption was used (see Figure 9) applying a six-point Likert scale for consistency with other measures within the survey instrument (1 = "*Disagree a lot*", through to 6 = "*Agree a lot*"). The items included were:

1. I have so much in life to be thankful for
2. If I had to list everything that I felt grateful for, it would be a very long list
3. When I look at the world, I don't see much to be grateful for (reverse-scored)
4. I am grateful to a wide variety of people
5. As I get older I find myself more able to appreciate the people, events, and situations that have been part of my life history

Resilience. The Academic Resilience scale, developed by Martin and Marsh (2006), provides a psychometrically robust unidimensional measure of academic resilience in adolescence. Using an Australian high-school sample, the Academic Resilience scale evidenced an excellent internal consistency estimate of alpha = .89 (Martin & Marsh, 2006).

The present research maintains the six-item structure seen in the Academic Resilience scale (see Figure 10), with item wordings reframed however, to generalise to a general-life context. For example, “I believe I’m mentally tough when it comes to exams” was worded, “I believe I am mentally tough when it comes to overcoming life challenges” and “I don’t let study stress get on top of me” was adapted to “I don’t usually let life stresses get on top of me” (see Appendix 2: The Research Survey Instrument). Given the evidenced psychometric properties of the Academic Resilience scale within an Australian representative sample, an adapted form was selected as an appropriate indicator of general resilience.

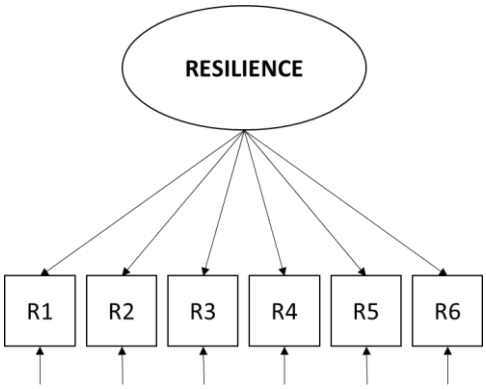


Figure 10. A priori factor structure of the Academic Resilience scale.

Data Collection Procedure

During the initial survey session, students were provided with an information sheet by the research team, emphasising voluntary involvement, the confidentiality of their individual responses, and that the data would be used for research purposes only, and would not be provided to the school. It was also however explained that an exception to this would be any circumstance where a student indicated a risk of harm to themselves or another person. In such cases, students were advised, the school psychologist would be notified so that appropriate follow-up could occur. Each testing session commenced with a brief set of instructions on how to access and complete the survey. All students were encouraged to seek assistance from a staff or research team member if they were unsure of any of the survey questions.

Questionnaires were administered via one of three modes. For students based at the suburban main school campus, surveys took place online via an email link accessed on student laptops, and were completed in a large theatrette during a timetabled class period. For questionnaires administered to students embarking on the outdoor program, hardcopy surveys, pens and blank envelopes were provided. These were completed in a large auditorium within the main college campus, immediately prior to their departure to the Glengarry program. The post-intervention questionnaire took place in a surf club hall and was completed offline via iPads or using the previously described hardcopy format. All surveys completed through laptops or iPad were developed using the Qualtrics (2015-17) electronic survey development tool. Students were asked to complete the questionnaire on their own and where applicable, to provide the completed form in a sealed envelope to the researcher when they had finished. Each administration of the questionnaire took approximately 25-35 minutes.

The survey was administered on four occasions over a period of one year (five consecutive school terms). The first instance surveyed all participants at T0 during Term 4 of Year 8 (Pilot/extended baseline control) and on three occasions during Year 9 at T1 (pre-test for intervention group, baseline control for the WLC), T2 (immediate post-test for the intervention group, baseline control for WLC), and T3 (long-term follow-up for the intervention group). Identical testing procedures were used for the consecutive 2015 and 2016 Year 8 cohorts.

Administration dates were as follows:

1. December 2015/16: End of the school year (T0)
2. January 2016/17: First day of the school year (T1)
3. June 2016/2017: Last week of Semester 1, term 2 (T2)
4. November/December 2016/2017: Last week of the school year (T3)

Summary

The research design and statistical procedures used in this investigation overcome several limitations within the broader body of literature to offer best-practice methodology (see Scrutton & Beames, 2015). The following chapter details the statistical procedures that were used to establish the psychometric properties of the survey instrument.

CHAPTER 5: STUDY 1—PSYCHOMETRIC ANALYSIS

Introduction

Establishing robust psychometric properties in instruments employed in applied research, is essential to protect the integrity of inferences drawn from them. With regard to Outdoor Adventure Education (OAE) research, a lack of appropriately selected scales and well-defined measurement of outcomes has long being a point of criticism (Dillon, 2013; Marsh et al., 1986c; Rickinson et al., 2004; Scrutton & Beames, 2015). This chapter commences by defining the research problem and hypotheses for the Study 1 investigation. An overview of methodology and results follows, with the intent of examining and enhancing measurement quality throughout this OAE investigation and to establish reliable evaluation of the impacts of OAE on key research outcomes.

Each hypothesis is underpinned by established theory and empirical evidence (see Chapter 2: Literature Review). As OAE is a developing field of research and is still characterised by gaps in evidence and methodological weaknesses, psychosocial as well as substantive factors were used to develop the research hypotheses.

Each hypothesis is numbered according to both study and hypothesis number (e.g. hypothesis 1.2 refers to Study 1, hypothesis 2).

Statement of Problem and Overarching Aim

Do the instruments employed in this research demonstrate psychometrically robust measurement within the male secondary school sample? That is, are students' responses to the instruments reliable? Do participant responses adequately represent *a priori*, or otherwise specified factor structures, and do the instruments demonstrate measurement consistency and convergent/discriminant validity over time? This statement of problems is applied to the following instruments, as detailed in Chapter 4: General Method.

1. The Self-Description Questionnaire-II (SDQII), as a measure of six domains of adolescent self-concept (General Academic, Physical Ability, Same-Sex Relations, Opposite-Sex Relations, Parent Relations, and Global Esteem) and the Self-Description Questionnaire-III (SDQIII) as a single-domain addition (Problem Solving);
2. Review of Personal Effectiveness with Locus of Control (ROPELOC), as a measure of 12 facets of life-effectiveness (Active Involvement, Cooperative Teamwork, Leadership Ability, Open Thinking, Quality Seeking, Self Confidence, Self-Efficacy, Social Effectiveness, Stress Management, Time Efficiency, Coping with Change, and Overall Effectiveness) and a bi-dimensional measure of students' perceptions of internal versus external regulation (Internal and External LoC);
3. Basic Psychological Needs Satisfaction and Frustration Scale (BPNSFS; adaptation), as a measure of adolescents' satisfaction of the three basic psychological needs posited within Self Determination Theory (Relatedness, Competence and Autonomy);
4. Motivation and Engagement Scale (MES), as a multidimensional measure of adolescent motivation and engagement (Thought, Behaviour, Mufflers, and Guzzlers);
5. General Health Questionnaire-12 (GHQ-12), as a clinical measure of psychological ill-being (Psychologic Distress);
6. Personal Well-being Index-School Children (PWI-SC), a single item used as a measure of student global satisfaction with life;
7. The Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS), as a unidimensional measure of students' subjective well-being;
8. The Gratitude Questionnaire (GQ-6; adaptation), as a five-item measure of adolescent proneness to experience gratitude in daily life.

9. Academic Resilience Scale (adaptation), as a six-item unidimensional measure of students' capacity to effectively respond to setbacks, challenges and adversity in general life.

The overarching aim of Study 1 was to test the survey instruments, to ensure they provide psychometrically-robust measurement of the outcomes being investigated. The following hypotheses specify the procedures through which this was achieved.

Statement of Hypotheses

Hypothesis 1.1: Reliability. The assessment of measurement scales' reliability will reflect acceptable internal consistency estimates (McDonald's omega coefficient >.7).

Hypothesis 1.2: Factorial structure. Confirmatory Factor Analysis (CFA) and Exploratory Structural Equation Modelling (ESEM) of each instrument's *a priori* hypothesised factor structure will reflect satisfactory fit on select indices (Comparative Fit Index (CFI) > .9, Tucker-Lewis Index (TLI) > .9; Root Mean Square Error of Approximation (RMSEA) < .08).

Hypothesis 1.3: Factorial invariance over time. Longitudinal invariance testing of research instrumentation will evidence measurement consistency over time.

Hypothesis 1.4: Structural validity of the combined instrument model. Each instrument's factor structure, as determined by Hypothesis 1.2, will remain stable and reflect acceptable fit when combined into a single unified model.

Hypothesis 1.5: Convergent/discriminant validity. The research instrument will show strong convergent and discriminant validity in relation to stability over time (T0 to T1) using the Campbell and Fiske (1959) multitrait-multimethod (MTMM) paradigm.

Rationale

In recent decades, educational policy makers have been placing greater emphasis on the need for evidence-based interventions residing within the school context, a

powerful setting to foster healthy development in adolescents (e.g. Kaplan & Flum, 2012; O'Mara et al., 2006). Findings from OAE research have yielded largely favourable results (e.g. Bowen & Neill, 2013; Hattie, 2009; Hattie et al., 1997), however, criticisms over inappropriate survey design (Scrutton & Beames, 2015) leave many of the claimed benefits considered little better than anecdotal, or described as "inconclusive", "over-ambiguous" and "uncritical" (Barrett & Greenaway, 1995, p. 53). Such issues highlight the critical need for strong research designs using well-defined and psychometrically strong measures to address these shortcomings and bolster the empirical foundation of OAE's claimed benefits.

Each of the above stated hypotheses was postulated in regards to an existing body of literature that demonstrates the psychometric suitability of the selected instrumentation within an adolescent population (see Chapter 4: General Method).

Method

The following section provides a comprehensive description of methodology specific to Study 1's research hypotheses, design and statistical analytic procedures.

Participants

The participant sample (N = 346) used to undertake the psychometric analyses described in Study 1 is consistent with the description provided in the Chapter 4: General Method section. That is, the participants were students aged 13 to 14 enrolled in the 2015 and 2016 Year 8 cohorts at an urban all-male high school.

Research Design

Study 1 set out to investigate the psychometric properties of each measurement instrument used in the present research. A total of 346 secondary school students from two consecutive Year 8 groups were administered the questionnaire battery mid-way through Term 4 of Year 8, and again on three occasions during Year 9.

Data Analysis

Consistent with Study 1 aims, reliability estimates, CFA, ESEM, longitudinal invariance analysis and MTMM were undertaken, to thoroughly investigate the psychometric properties of each measure within the survey instrument. The SPSS data analytic software (IBM Corp., Released 2013) was used to calculate descriptive information and inter-variable correlations for MTMM analysis, whereas factor reliability estimates, CFA, ESEM and longitudinal invariance testing were carried out in Mplus Version 8.0 (Muthén & Muthén, 1998-2017).

When reliability (hypothesis 1.1.) and structural validity (hypothesis 1.2. and 1.4.) were tested, the four-wave data set (T0-T3) was stacked, or arranged in "long" format (*N* = 1,384 sets of responses by 346 individuals) to maximise the number of cases and increase the statistical power of the analysis. When stacked data was applied, because 346 participants contributed responses at T0, T1, T2 and T3, the Mplus complex design option was used to adjust standard errors and account for the fact that each participant had contributed multiple sets of responses (Muthén & Muthén, 1998-2017).

ESEM analysis used an oblique target rotation, as this method is deemed most appropriate for complex ESEM models represented by multiple factors and item-loadings, and when *a priori* measurement models have been clearly defined (Asparouhov & Muthén, 2009; Marsh et al., 2014). Marsh et al. (2014) explain that in these circumstances, using target rotation provides stronger *a priori* models, allows for greater control when specifying model parameters, and provides clarity when interpreting output. In accordance with recommendations by Howitt and Cramer (2005), factor loadings greater than .30 were considered to be meaningful and were determined as contributing adequately to their respective variables.

Evaluating the Survey Instrument's Psychometric Properties

Reliability of the measurement scales, omega (ω). Reliability refers to the degree to which a scale demonstrates consistent measurement, in that scores remain unaffected by measurement error (Kline, 2005). In general, reliability is estimated through the computation of internal consistency, referring to the 'agreement' between items or the stability of item responses on any given scale. One dominant method of deriving the consistency with which an item-set measures a single latent factor, is to use the omega coefficient, as suggested by McDonald (1999). This reliability estimate reflects the proportion of test variance among a set of observed scores that is accounted for by a general factor. Where the widely applied alpha coefficient (Cronbach, 1951) assumes item-loadings to be equal for single-factor models, the omega estimate indexes the proportion of variance that a common factor shares with its respective items, and therefore provides important information when item-saturation from a latent factor varies (Revelle & Zinbarg, 2009). In doing so, the omega coefficient overcomes biases noted with the lower-bound alpha coefficient when scale items are not tau-equivalent; for this reason it is widely endorsed as a more appropriate index of internal consistency in applied research (Dunn, Baguley, & Brunsden, 2014; Revelle & Zinbarg, 2009; Zinbarg, Revelle, & Yovel, 2007; Zinbarg, Yovel, Revelle, & McDonald, 2006).

Establishing the structural validity of the measurement scales. Structural equation modelling (SEM) methods have become increasing popular across social science disciplines, as they allow a researcher to examine complex multivariate relationships among observed and latent variables (Marcoulides & Yuan, 2016). In applied SEM research, a range of indicators can be used to estimate a model's goodness of fit; that is, how well the data from a research sample represents the instrument's factor structure. General agreement has been reached regarding the appropriate use of the

Tucker–Lewis Index (TLI; Tucker & Lewis, 1973), the Comparative Fit Index (CFI; Bentler, 1990), and the Root Mean Squared Error of Approximation (RMSEA; Marsh, Hau, & Grayson, 2005; Steiger, 1989). The TLI (see Marsh, Balla, & Hau, 1996; Schumacker, 2010) and CFI (Hu & Bentler, 1999) values vary along a 0 to 1 continuum, where values of >.90 reflect adequate fit and values of >.95 suggest excellent fit to the data (Cheung & Rensvold, 2002). According to Browne and Cudeck (1992), an RMSEA of less than 0.05 is an indication of good model fit, while values up to .08 reflect reasonable errors of approximation. A further index to determine goodness of fit is the overall chi-squared statistic, whereby a low chi-squared value relative to the degrees of freedom (*df*) producing a non-significant value represents good fit to the data (Marsh, 1994). The chi-squared value has, however, been criticized for its known sensitivity to sample size, thus leading much of applied SEM research to focus on indices such as the CFI (Cheung & Rensvold, 2002), TLI and RMSEA, which are relatively sample-size independent (Hu & Bentler, 1999; Little, Card, Slegers, & Ledford, 2007). Given the relatively restricted sample size in the present investigation, CFI, TLI and RMSEA indices will be examined as the primary determinant of models' goodness of fit, along with a detailed evaluation of parameter estimates in relation to *a priori* predictions.

Marsh (1994) suggests the following steps to establish the structural validity of a measurement model:

1. Establish that model parameters and estimates are justified by substantive *a priori* rationale and common sense,
2. Determine whether the proposed solution converges and is well defined,
3. Evaluate changes in various indices between systematically competing models.

The following section outlines two advanced SEM methods recommended by (Marsh, 1994), namely CFA and ESEM, as they were used in the present investigation to

establish whether *a priori* measurement structures converged and reflected acceptable fit. Generally speaking, ESEM models are preferred when their fit indices exceed those of the corresponding CFA model, however when this is not the case, the CFA model should be retained on the basis of parsimony (Marsh, Nagengast, & Morin, 2013).

Confirmatory Factor Analysis (CFA). Confirmatory Factor Analysis (CFA) was applied to each measure within the research instrument to evaluate *a priori* hypothesised factor structures within the research sample. This structural validity method holds the assumption that variance between the observed scores for scale-items is determined by the influence of an underlying hypothesised, or latent, construct plus residual measurement error. In the Strauss and Smith (2009) review of construct validity, theory and methodology, it is noted:

> A major advantage of CFA in construct validity research is the possibility of directly comparing alternative models of relationships among constructs, a critical component of theory testing. (p. 15-16)

When administering CFA, a researcher proposes an *a priori* structural model that reflects a set of relationships between specified observable variables, such as item scores, to an underlying construct, also referred to as a factor (Wood, 2008). In this example, the factor would be hypothesised to influence the respective item responses to group together, as they are assumed to measure the same underlying construct. The *a priori* hypothesized, or otherwise specified, structural models for the research instruments have been detailed in the preceding Chapter 4: General Method section.

CFA is traditionally employed as a first step in evaluating structural validity, whereby the *a priori* structure has been identified, however the method is not without its limitations. CFA has been found to under-represent models' goodness of fit (Marsh, Hau, et al., 2005), resulting in biased parameter estimates resulting from misspecification of zero loadings, in that each item can only load on a single factor. In acknowledgement of

this, it is recommended that researchers employ both CFA and ESEM methods for structural validity assessment and that for each comparison, the model reflecting preferred fit on specified indices should be retained for ongoing analysis (Marsh et al., 2014; Marsh et al., 2013). ESEM will now be introduced as a complementary structural validity procedure to CFA.

Exploratory Structural Equation Modelling (ESEM). Exploratory structural equation modelling provides an innovative and integrative framework that draws on Exploratory Factor Analysis (EFA), CFA and SEM methods. By doing so, ESEM offers the flexibility of EFA, while maintaining the breadth of application seen in CFA (Marsh et al., 2014). A fundamental virtue of incorporating exploratory methods to *a priori* model identification is the avoidance of misspecifications resulting from ambiguous and weak theoretical bases for *a priori* measurement models (Browne, 2001).

Further justification for exploratory solutions arises from the fact that constructs measured in applied research frequently contain cross-loadings which hold a substantive premise or may represent an alternative source of measurement error (Asparouhov & Muthén, 2009). Asparouhov and Muthén (2009) add that while CFA solutions can bring about strategic advantages, their non-targeted and constrained zero loadings generate inflated factor correlations and biased SEM estimates. Marsh, Liem, Martin, Morin, and Nagengast (2011) demonstrated how ESEM solutions for multidimensional measures generally reflect better fit to the data, given that they are substantially less restrictive, despite encompassing reductions in parsimony. Furthermore, the ESEM solution's allowance for indicator variable cross-loadings has been reported to generate substantially less-inflated factor correlations, a critical indicator when examining instrument multidimensionality (Marsh et al., 2011).

It is widely acknowledged that ESEM is a complementary method to CFA, rather than a replacement (Lee & Lam, 2014). As such, it is recommended both methods be administered and the respective fit indices compared (Marsh et al., 2014). In alignment with recommendations for best-practice psychometric procedures in structural validation (Lee & Lam, 2014; Marsh et al., 2014) , the present investigation compares CFA and ESEM solutions for each measure within the research instrument.

Measurement stability over time. Establishing the relational stability of latent factors and their respective items over time has been identified as a crucial requirement in repeated-measures research (Marsh et al., 2011), and is richly attested to within SEM literature (Jöreskog & Sörbom, 1979; Meredith, 1993; Mora et al., 2011). This procedure, referred to as longitudinal invariance testing (LIT), examines whether an instrument's measurement in a particular sample is equivalent when re-administered. When this is the case invariance is supported, in that mean differences and correlations between time-waves are free of extraneous factors.

LIT is conducted in a sequential manner, requiring the equivalence of increasingly restricted paired-model parameters to be simultaneously contrasted using data fitted to two or more time-waves. As described by Chen (2007), the first model, named the configural model, is unconstrained; whereby all parameters are freely estimated. This provides a test for overall model fit and a baseline representation for subsequent comparison. Secondly, the metric model places constraints on factor loadings that represent the strength of relationship between factors and their respective items. When item-loadings on a latent factor are equal on different occasions, the factor is inferred to have the same unit or interval—a necessary determinant when comparing change-scores in longitudinal research. The third model in LIT takes place at the intercept level and is named the scalar model. Chen (2007) explains that intercepts represent the origin of a

scale and that invariance at this level reflects group measurement on different occasions having equivalent measurement (loadings) and the same origin (intercepts).

Cheung and Rensvold (2002) and Chen (2007) note that if the decrease in fit for more parsimonious models evidenced by incremental fit indices is less than .01, then factor loadings are established as invariant across time, and acceptance of the more parsimonious model is justified. In such cases, following comparison between the configural and metric models, weak invariance has been established. This calls next for the metric model to be contrasted to the scalar model. Adhering to the same index-change guidelines, when the more parsimonious scalar model is accepted, strong measurement invariance has been ascertained.

Following Little et al. (2007), the chi-squared difference was not interpreted as an indicator of invariance due to the index's relative sensitivity to sample size. Instead, changes in CFI and TLI (> -.01) and RMSEA (> .01) were indicative of violation to invariance in contrasting the configural, metric and scalar models fitted to the T0 and T1 time-waves. Finally, LIT included correlated uniquenesses of matching items, so that their inclusion did not positively bias the test-retest correlation estimates (Marsh & Hau, 1996).

Discriminant and convergent validity of research constructs. The multitrait multimethod (MTMM) paradigm proposed by Campbell and Fiske (1959) is a widely endorsed construct validation design in psychological research (e.g. Byrne, 1996; Marsh, Ellis, Parada, Richards, & Heubeck, 2005; Marsh, Martin, & Hau, 2006). This approach investigates convergent and discriminant validities of latent constructs through the measurement of multiple traits (e.g. research outcomes) when applying multiple methods (e.g. multiple instruments or measurement occasions; Campbell & Fiske, 1959). Referring to an *a priori* factorial structure correlation matrix, the classic MTMM

paradigm requires comparisons of convergent validities, known as monotrait-heteromethod correlations (MTHM; matching traits test-retest correlations), heterotrait-heteromethod correlations (HTHM; different traits on different occasions), and heterotrait-monomethod correlations (HTMM; nonmatching traits on the same occasion). Convergent validity is supported when MTHM correlations are significant and substantial in size, while discriminant validity is supported when these convergent validities are notably larger than the other HTHM and HTMM correlations.

Operationalizing multiple methods across multiple occasions through MTMM analysis provides a strong approach to evaluating the stability of responses for multidimensional instruments (Campbell & Connell, 1967; Marsh, Martin, & Jackson, 2010). This is especially pertinent to the present research, as inadequate computer memory and computing capacity of the statistical programs, prevented the integrated latent model containing all instruments to be tested simultaneously for longitudinal invariance. As such, MTMM procedures were run using weighted factor scores derived from the combined instrument battery containing the best fitting CFA, or otherwise ESEM configuration, for each instrument. Extending the original Campbell and Fiske (1959) guidelines, the analysis generated a large correlation matrix, including latently derived factor scores representing all combinations of traits and methods. Doing so has been argued to achieve a more heuristic evaluation of convergent and discriminant validity compared to that provided by manifest scores (Marsh, Martin, et al., 2006; Marsh et al., 2014).

Summary

The previous section has overviewed the psychometric procedures that were pertinent to testing Study 1 hypotheses and established the process of reliable and valid measurement of research outcomes. Such procedures include assessment of reliability,

baseline structural models' goodness of fit, testing longitudinal measurement invariance, and employing MTMM procedures to assess convergent and discriminant validity in relation to stability over time. The results from these analyses are provided below.

Results

The previous section offered an introduction to and explanation of, arguably the most advanced and innovative statistical methods in reliability and validation research. Each instrument in the present investigation was systematically subjected to psychometric analysis in accordance with Study 1 hypotheses. The results from this analysis will now be presented.

Psychometric Properties of the Research Instrumentation

Hypothesis 1.1: Omega (ω) as an estimate of internal consistency.

Overview. Hypothesis 1.1 postulates that all scales within the research instrumentation will demonstrate acceptable reliability scores for the total research sample. As described earlier in this chapter, the omega coefficient (ω) was reported as the reliability estimate for all instrument component scales.

SDQ II/III. Reliability estimates for the SDQ factors shown in Table 6 reveal the subscales reached internal consistency estimates in the good to excellent range (ω = .77 to .93.), with a median ω of .89.

ROPELOC. Reliability estimates for the 14 ROPELOC factors shown in Table 7 demonstrated all scales reflecting internal consistency estimates in the good to excellent range (ω = .78 to .91.), with a median ω of .86.

BPNSFS. Reliability analysis for the three BPNSFS factors shown in Table 8 demonstrate all scales reflecting omega estimates in the excellent range (ω =.86 to .90.), with a median ω of .86.

Table 6 *Internal Consistency Coefficient Omega, CIs and Number of Items for the SDQ Scales When Applying Stacked Data to the Total Research Sample*

SDQ II/III	Year 8 2015/2016 (n = 380)		
	Omega (ω)	95% CI [LL, UL]	No. of Items
Subscales			
General Academic	.91	[.904, .921]	5
Physical Abilities	.89	[.880, .902]	5
Same-Sex Relations	.83	[.811, .847]	5
Opposite-Sex Relations	.77	[.740, .793]	4
Parent Relations	.83	[.816, .854]	4
Global Esteem	.93	[.918, .934]	7
Problem Solving	.93	[.920, .935]	8
Median reliability score	.89		
Mean reliability score	.87		

Table 7 *Internal Consistency Coefficient Omega, CIs and Number of Items for the ROPELOC Scales When Applying Stacked Data to the Total Research Sample*

ROPELOC	Year 8 2015/2016 (n = 380)		
	Omega (ω)	95% CI [LL, UL]	No. of Items
Subscales			
Overall Effectiveness	.88	[.872, .897]	3
Cooperative Teamwork	.86	[.840, .875]	3
Self Efficacy	.88	[.863, .891]	3
Leadership Ability	.89	[.872, .901]	3
Active Involvement	.81	[.793, .835]	3
Open Thinking	.81	[.785, .827]	3
Quality Seeking	.84	[.825, .859]	3
Self Confidence	.84	[.822, .856]	3
Social Effectiveness	.91	[.895, .919]	3
Stress Management	.87	[.852, .881]	3
Time Efficiency	.87	[.860, .889]	3
Coping with Change	.89	[.878, .903]	3
ILoC	.84	[.823, .859]	3
ELoC	.78	[.762, .808]	3
Median reliability score	.86		
Mean reliability score	.86		

Note. ILoC = Internal Locus of Control, and ELoC = External Locus of Control.

Table 8 *Internal Consistency Coefficient Omega, CIs and Number of Items for the BPNSFS Scales When Applying Stacked Data to the Total Research Sample*

BPNSFS	Year 8 2015/2016 (n = 380)		
	Omega (ω)	95% CI [LL, UL]	No. of Items
Subscales			
Autonomy	.86	[.849, .878]	4
Competence	.90	[.893, .913]	4
Relatedness	.86	[.844, .876]	4
Median reliability score	.86		
Mean reliability score	.88		

Motivation Engagement Scale (MES). Internal consistency analysis of the four MES factors, namely Thought, Behaviour, Muffler, and Guzzler, demonstrated explicit reliability issues, whereby two scales failed to achieve acceptable limits of >.7. Estimates ranged from ω = .60 to .84 (Median ω = .68; see Table 9). This suggests that the variance observed in the true scores for the Muffler and Guzzler factors was impacted by measurement error, independent of variance explained by the observed scores. Consequently, further analysis of the instrument's measurement properties will be undertaken in the structural validation section to follow, where latent factors will be used to control for measurement error and model configuration will be adjusted as necessary.

GHQ-12. Reliability analysis for the GHQ-12 demonstrated both scales as reflecting omega estimates in the acceptable range (ω = .74 to .84.), with a median ω of .79 (see Table 10).

Table 9 *Internal Consistency Coefficient Omega, CIs and Number of Items for the MES Scales When Applying Stacked Data to the Total Research Sample*

MES	Year 8 2015/2016 (n = 380)		
	Omega (ω)	95% CI [LL, UL]	No. of Items
Subscales			
Thought	.72	[.688, .752]	3
Behaviour	.84	[.825, .857]	3
Muffler	.60	[.531, .674]	3
Guzzler	.65	[.609, .686]	2
Median reliability score	.68		
Mean reliability score	.70		

Table 10 *Internal Consistency Coefficient Omega, CIs and Number of Items for the Two GHQ-12 Clinical Scales When Applying Stacked Data to the Total Research Sample*

GHQ-12	Year 8 2015/2016 (n = 380)		
	Omega (ω)	95% CI [LL, UL]	No. of Items
Subscales			
Positive	.74	[.717, .767]	6
Negative	.84	[.822, .853]	6
Median reliability score	.79		
Mean reliability score	.79		

Unidimensional Instruments. Reliability analysis was conducted for each unidimensional measure. The results demonstrated all scales reflecting omega estimates in the acceptable to excellent range (ω = .75 to .93.), with a median ω of .86 (see Table 11).

Table 11 *Internal Consistency Coefficient Omega, CIs and Number of Items for the Four Unidimensional Measures When Applying Stacked Data to the Total Research Sample*

One-Factor Scales	Year 8 2015/2016 (n = 380)		
	Omega (ω)	95% CI [LL, UL]	No. of Items
WEMWBS	.86	[.845, .871]	7
GQ-6	.75	[.731, .776]	5
Resilience	.93	[.920, .935]	6
Life Satisfaction	N/A		1
Median reliability score	.86		
Mean reliability score	.85		

Conclusion. Reliability estimates for all component scales comprising the SDQ, ROPELOC, BPNSFS, GHQ-12 and unidimensional instruments exceeded the minimum acceptable limits and thus, hypothesis 1.1 was supported and accepted. Consequently, these instruments were deemed to be reliable measures within the total research sample. An exception to this was noted for the four-factor MES scale, which revealed unacceptable omega estimates on two subscales. It was decided that closer examination of the MES' structural integrity would be undertaken by administering CFA and ESEM procedures, to determine the most appropriate course of action to rectify this issue.

Hypothesis 1.2: Factor structure and variable correlations

Overview. Hypothesis 1.2 predicted that each instrument would be adequately represented by the specified *a priori* factor structure, as detailed in Chapter 4: General Method. Hypothesis 1.2 was tested by applying both CFA and ESEM representations to each measurement model and then retaining the stronger configuration based on an evaluation of goodness of fit.

SDQ II/III. CFA was applied to the restrictive *a priori* seven factor SDQ structure, whereby items could only load on their respective factor. The statistical indices used to evaluate model fit (described earlier in this chapter) included CFI, TLI, and RMSEA. The

analysis found the model showing acceptable fit to the data, with CFI = .907, TLI = .898 and RMSEA = .048, 90% CI [.046, .049]. The ESEM representation was next applied, whereby items were allowed to cross-load to other factors using an oblique target rotation. This model showed fit indices to falling comfortably in the acceptable range, CFI = .953, TLI = .928 and RMSEA = .04, 90% CI [.038, .042] (see Table 12).

Table 12 *Model Fit Statistics When Administering CFA and ESEM Procedures to the Seven SDQ Scales Using Stacked Data*

SDQ Model	χ^2	*df*	CFI	TLI	RMSEA	90% CI
CFA	3165.87	642	.907	.898	.048	.046, .049
ESEM	1725.73	456	.953	.928	.040	.038, .042

Note. *df* = Degrees of freedom for the model; CFI = comparative fit index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square of Approximation.

Comparing CFA and ESEM set-ups, the ESEM model represented substantially improved fit to the data. To examine model parameters in more detail, factor loadings and inter-factor correlations were next examined. As expected, CFA resulted in marginally higher loadings across a majority of items. All loadings using the CFA were found to be statistically significant and sufficient in size, with values ranging between .44 and .91, however ESEM saw a single item loading fall below the recommended .30 cut-off on the same-sex relations subscale, all loadings ranging from .28 to .94. Exploration of factor relationships between the seven sub-scales when applying CFA found correlations to range from .18 to .65 (Mean *r* =.38; Median *r* = .38), which was unsurprisingly higher than those seen using ESEM, with values ranging from .03 to .62 (Mean *r* =.30; Median *r* = .30). These results are somewhat consistent with, although slightly higher than, the SDQ II factor correlations reported by Marsh (1989), which ranged from near-zero to

moderately positive (SDQII; Median *r* = .15). It is, however, important to note that only 6 of the 11 SDQII factors used in this paper were incorporated into this investigation.

When running CFA and ESEM analysis for the SDQ instrument, there appeared to be a conceptual justification for model modifications based on wording similarities: to correlate the errors for items 3 ('I make friends easily with boys') and 4 ('I make friends easy with girls'); and items 7 ('I do not get along very well with boys') and 8 ('I do not get along very well with girls'). Modification indices further supported these adjustments, which were accepted and included into ongoing SDQ model specifications.

The ESEM representation of the seven factor SDQ instrument was selected for ongoing analysis, given the substantially improved overall fit and lower factor correlations, supporting multidimensionality and distinctiveness between factors. Table 13 provides the item loadings and factor correlations when applying ESEM methods. The same-sex relations factor requires mention, due to one item loading poorly on to the factor. It is considered that residual error relating to wording effects for the same and opposite-sex relations factors may have contributed to this, as indicated by the increased item cross-loadings seen between these subscales. However, in light of the remaining four items making up the same-sex relations factor loading highly (Mean =.65; Median = .74), it was concluded the factor was adequately defined.

ROPELOC. Both CFA and ESEM representations were applied to the 14-factor ROPELOC instrument, and overall fit indices were computed. Set-ESEM was applied to the ESEM representation, whereby the two LoC scales were modelled separately using CFA configuration, so to not share non-target loadings. Given the later addition of the LoC scales to the Life Effectiveness Questionnaire, to complement the life effectiveness domains by including a separable measure of LoC, there appeared to be substantive and theoretical justification for making this model adjustment.

Table 13 *ESEM Factor Loadings and Correlations for the SDQ Scales When Applying Stacked Data*

	Acad	Phys	Sm Sx	Op Sx	Prnt	Esteem	Prob
Factor Loadings							
1	.71	.68	.70	.50	.88	.82	.64
2	.83	.81	.32	.86	.90	.80	.76
3	.89	.94	.78	.88	.85	.77	.73
4	.78	.84	.80	.36	.45	.77	.80
5	.81	.65	.28			.76	.79
6						.80	.83
7						.75	.90
8							.80
Factor Correlations							
Acad	*1.00*						
Phys	*.25*	*1.00*					
Sm Sx	*.09*	*.16*	*1.00*				
Op Sx	*.14*	*.40*	*.27*	*1.00*			
Prnt	*.18*	*.18*	*.26*	*.16*	*1.00*		
Esteem	*.56*	*.55*	*.19*	*.39*	*.46*	*1.00*	
Prob	*.50*	*.36*	.03	*.34*	*.28*	*.62*	*1.00*

Note. All parameters were estimated using stacked data, and are reported in standardised format. SDQ factors are: Acad = General Academic, Phys = Physical Abilities, Sm Sx = Same-Sex Relations, Op Sx = Opposite-Sex Relations, Prnt = Parent Relations, Esteem = General Esteem, and Prob = Problem Solving. Each factor was inferred on the basis of four to eight measured variables, labelled 1-8 in the upper left column. All factor loadings are significant at $p < .001$. Significant factor correlations have been italicised.

The CFA overall fit statistics suggested acceptable fit to the data, with CFI = .922, TLI = .907, and RMSEA = .045, 90% CI [.043, .046]. In contrast, the ESEM reflected substantially improved fit, indices falling in the acceptable to excellent range (CFI = .958, TLI = .921, and RMSEA = .041, 90% CI [.039, .043]; see Table 14).

An assessment of CFA factor loadings observed each to achieve statistical significance and be sufficient in size, with values ranging between .70 and .89. However, factor inter-correlations showed four scales to be highly correlated with the Overall Effectiveness scale (r >.8). This places the notion of multidimensionality and the

Table 14 *Model Fit Statistics When Administering CFA and ESEM to the 14 ROPELOC Scales Using Stacked Data*

ROPELOC Model	χ^2	*df*	CFI	TLI	RMSEA	90% CI
CFA	3265.60	728	.922	.907	.045	.043, .046
ESEM	1831.60	460	.958	.921	.041	.039, .043

Note. df = Degrees of freedom; CFI = comparative fit index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square of Approximation.

distinctiveness of these factor pairs in question. All correlations for the CFA model ranged from .00 to .89 (Mean *r* =.60; Median *r* = .65). Factor loadings for the set-ESEM configuration (see Table 15) saw the Overall Effectiveness scale as poorly defined, with loadings ranging from .05 to .35 (Mean *r* =.19; Median *r* = .15). Closer examination indicated that this was due to the variance from the Overall Effectiveness scale items being absorbed into other factors. This did not seem unusual, given that this scale represents a theoretically global indicator of the remaining ROPELOC factors. Set-ESEM factors were adequately distinct from one another, indicated by only modest factor correlation estimates (*r*s = .01 to .71; Mean *r* =.39; Median *r* = .42); a substantial improvement from the CFA configuration.

The set-ESEM configuration was retained for the ROPELOC instrument on the basis of notably improved overall fit and greater distinctiveness between factors. Table 15 provides the item loadings and factor correlations resulting from applying this method. It was, however, determined appropriate to exclude the Overall Effectiveness scale from ongoing analysis and the Study 2 investigation, as the scale was inadequately defined. With this adjustment, set-ESEM loadings ranged from .28 to 1.17 (Mean =.7; Median = .76), with a single loading below .3 on the Open Thinking scale. However, this factor was deemed adequately represented by the remaining two indicator loadings, which were substantial in size: .51 and .90.

Table 15 *Set-ESEM Factor Loadings and Correlations Among the ROPELOC Scales When Applying Stacked Data*

	CT	SF	LA	AI	OT	QS	SC	SE	SM	OE	TE	CH	IL	EL
							Factor Loadings							
1	.73	.61	1.03	.80	.90	.39	.32	.90	.76	.37	.66	.76	.79	.70
2	.83	1.17	.85	.47	.51	.75	.33	.84	.62	.15	.77	.81	.84	.75
3	.86	.48	.61	.34	.28	.67	.41	.83	.80	.05	.77	.69	.77	.78
							Factor Correlations							
CT	1.00													
SF	.56	1.00												
LA	.68	.67	1.00											
AI	.63	.61	.69	1.00										
OT	.41	.47	.42	.51	1.00									
QS	.36	.43	.40	.53	.57	1.00								
SC	*.23*	*.42*	*.36*	*.31*	*.19*	*.31*	1.00							
SE	.62	.59	.63	.58	.40	.40	*.38*	1.00						
SM	.49	.67	.52	.45	.43	.28	*.30*	.56	1.00					
OE	*.10*	*.21*	*.10*	*.23*	*.20*	*.19*	*.09*	.14	*.17*	1.00				
TE	.47	.52	.46	*.42*	.36	.38	*.42*	.50	.46	*.08*	1.00			
CH	.52	.67	.53	.46	.48	.40	*.37*	.55	.71	*.03*	.58	1.00		
IL	.46	.63	.58	.70	.59	.71	*.49*	.51	.47	*.21*	.41	.53	1.00	
EL	*.03*	*.11*	*.04*	*.16*	*.12*	*.28*	*.17*	*.01*	*.02*	*.16*	*.03*	*.05*	.33	*1.00*

Note. All parameters were estimated using stacked data and are reported in standardised format. ROPELOC factors are: CT = Cooperative Teamwork, SF = Self Efficacy, LA = Leadership Ability, AI = Active Involvement, OT = Open Thinking, QS = Quality Seeking, SC = Self Confidence, SE = Social Effectiveness, SM = Stress Management, OE = Overall Effectiveness, TE = Time Efficiency, CH = Coping with Change, IL = Internal Locus of Control, and EL = External Locus of Control. Each factor was inferred on the basis of three measured variables, labelled 1-3 in the upper left column. All factor loadings are significant at $p < .001$. Significant factor correlations have been italicized.

BPNSFS. CFA and ESEM were administered on the three-factor SDT instrument to establish structural validity. Overall fit indices for the CFA model showed excellent fit to the data (CFI = .988, TLI = .984 and RMSEA = .032, 90% CI [.026, .038]). The ESEM representation similarly evidenced excellent fit to the data, whereby CFI = .988, TLI = .985 and RMSEA = .031, 90% CI [.023, .039] (see Table 16).

Table 16 *Model Fit Statistics When Administering CFA and ESEM Procedures to the BPNSFS Needs Satisfaction Scales When Using Stacked Data*

BPNSFS Model	χ^2	*df*	CFI	TLI	RMSEA	90% CI
CFA	140.54	51	.988	.984	.032	.026, .038
ESEM	86.40	33	.993	.985	.031	.023, .039

Note. df = Degrees of freedom; CFI = comparative fit index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square of Approximation.

The overall fit indices for each SEM model indicated comparably excellent fit to the data. An assessment of factor loadings and inter-factor correlations was conducted. Using CFA, all loadings were statistically significant and sufficient in size, values ranging from .65 to .87 (Mean =.80; Median = .82). Results from the ESEM analysis were equivalent, values ranging from .60 to .92 (Mean =.79; Median = .80), with mean loadings between the two representations varying only at the second decimal place. Similarly, factor correlations for CFA analysis ranged from *r* = .66 to .72 (Mean *r* = .68; Median *r* = .68), differing only from the ESEM representation in mean factor correlation at the third decimal place, values ranging from *r* = .65 to .72 (Mean *r* = .68; Median *r* = .68).

The CFA representation for the three factor SDT instrument was retained on the basis of parsimony, given the equivalency of fit indices, loadings and factor correlations to the ESEM model. Table 17 provides the item loadings and factor correlations after applying CFA analysis.

Table 17 *Item Loadings and Factor Correlations for the BPNSFS When Applying CFA to Stacked Data*

	Factor Loadings		
	Autonomy	Competence	Relatedness
1	.76		
2	.87		
3	.79		
4	.73		
5		.82	
6		.87	
7		.83	
8		.83	
9			.77
10			.87
11			.84
12			.65
	Factor Correlations		
Autonomy	*1.00*		
Competence	*0.72*	*1.00*	
Relatedness	*0.68*	*0.66*	*1.00*

Note. All parameters were estimated using stacked data and are reported in standardised format. Each factor was inferred on the basis of four measured variables labelled 1-12 in the upper left column. Significant factor correlations have been italicized.

MES. As previously reported in the results for the reliability analysis, the four-factor MES failed to reveal acceptable internal consistency values on two scales ($\omega < .7$). CFA and ESEM procedures were applied to test whether modification of the instrument's factor structure was required to satisfy psychometric requirements for ongoing use in the project.

Goodness of fit statistics were compared for the CFA and ESEM representations. Overall fit indices for the CFA model suggested poor model fit (CFI = .892, TLI = .843, RMSEA = .074, 90% CI [.068 - .081]). Inspection of loadings found the Muffler scale to be poorly represented (Item 7, "I get quite anxious about school work and tests"; loading = .24), while high factor correlations between the Guzzler and Muffler scales (r = .86)

suggested that multicollinearity problems may be contributing to poor model fit. The ESEM representation was subsequently applied, however this model was unable to converge and did not produce indicators for model fit (see Table 18).

Table 18 *CFA and ESEM Goodness of Fit Statistics for the Four-Factor MES When Applying Stacked Data*

MES	CHISQ	*df*	CFI	TLI	RMSEA	CI
CFA/SEM	402.11	38	.892	.843	.074	.068, .081
ESEM	N/A	N/A	N/A	N/A	N/A	N/A

Note. The ESEM model could not converge, due to excessive factor cross-loadings and a negative residual variance for Item 1.

The CFA and ESEM procedures concluded that the four factor MES was not adequately defined. The *a priori* factor structure for the instrument in this research, while firmly grounded in empirical theory (see Chapter 4: General Method), adopts structural adaptations that have not been subject to rigorous psychometrically testing. Consequently, Exploratory Factor Analysis (EFA) was deemed to be an appropriate solution for identifying a workable factor structure for the MES. Doing so retained a reflective approach when deriving the specifications of the motivation and engagement measurement model, in that observed items manifest from the predetermined latent construct.

EFA – Constructing a revised measure of motivation and engagement. EFA with maximum likelihood extraction and geomin rotation was run in Mplus Version 7 (Muthén & Muthén, 1998-2017) to yield item-correlations, factor loadings, communalities, and variance percentages for the 11 MES items. A .5 target value was specified for the rotation formula. Extraction of two, three and four factor solutions was specified—a decision informed by substantive reasoning—for school motivation and engagement to be a multidimensional construct. As expected, EFA determined a four-factor solution to be insufficiently identified and unable to converge; a recurring result after trialling several

rotation methods. Factor extraction based on eigenvalues exceeding one identified three components (eigenvalues = 3.80, 1.94, and 1.05). Unsurprisingly, the analysis found the two-factor solution to reflect poor fit (CFI = .882, TLI = .809, and RMSEA = .082, 90% CI [.077, .089]), however a three-factor solution saw all overall fit indices falling in the acceptable to excellent range (CFI = .972, TLI = .938, and RMSEA = .047, 90% CI [.038, .055]; see Table 19).

Table 19 *Summary of Fit Indices When Deriving Two, Three and Four Factor Solutions using EFA for the MES*

EFA Factor Solution	χ^2	*df*	CFI	TLI	RMSEA	90% CI
2-Factors	429.432	34	.882	.809	.082	.075, .089
3-Factors	119.050	25	.972	.938	.047	.038, .055
4-Factors	NA	NA	NA	NA	NA	NA

Note. df = Degrees of freedom; CFI = comparative fit index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square of Approximation. Geomin rotation was applied, specifying two, three and four factor extraction. NA reflects unobtainable index, estimated due to failure of model to converge.

As illustrated in Table 20, all loadings for the three-factor solution were significant and substantial in size, indicating that each factor was well identified. Component 1 was represented by Items 1 to 3 (original Thought items; loading range = .44 - .71) and Items 4 to 6 loaded on component 2 (original Behaviour items; loadings = .47 to .88). The third component consisted of the six negatively worded items that were originally specified to represent the Muffler and Guzzler scales (loadings = .35 to 70). Factor correlations were weak to modest (*r*s = .12 to .41, Mean = .25) and indicative of construct multidimensionality.

Table 20 *Factor Loadings Following EFA using a Geomin Rotation for the 11-Item MES Instrument*

	Factors		
Items	1	2	3
MvE_1	.71		
MvE_2	.59		
MvE_3	.44		
MvE_4		.72	
MvE_5		.88	
MvE_6	.37	.47	
MvE_7r			.35
MvE_8r			.53
MvE_9r			.59
MvE_10r			.60
MvE_11r			.70

Note. Loadings < .3 have been removed from the table. All reported loadings are significant at $p < .05$.

In light of EFA procedures deriving a structurally valid three-factor representation for the MES, it was determined appropriate to re-examine reliability estimates and subject the new model to CFA and ESEM procedures. Factors 1 and 2 continue to be respectively labelled 'Thought' and 'Behaviour', while the new, third factor comprising the five negatively worded items originating from the Muffler and Guzzler scales, is renamed 'Hampering'. The results from this analysis are presented below.

Reliability and structural validity for the revised three-factor MES. In alignment with hypothesis 1.1, reliability testing was conducted on the revised MES instrument, revealing all internal consistency estimates falling in the acceptable range (ω =.71 to .84; median = .72; see Table 21). Thus, hypothesis 1.1 was supported, and the revised three-factor structure of the MES was determined a reliable measure of motivation and engagement within the research sample.

Table 21 *Internal Consistency Coefficient Omega, CIs and Number of Scale Items for the Revised MES Instrument When Applying Stacked Data*

Three-Factor MES	Year 8 2015/2016		
	Omega (ω)	95% CI [LL, UL]	No. of Items
Scales			
Thought	.72	[.688, .752]	3
Behaviour	.84	[.825, .857]	3
Hampering	.71	[.683, .731]	5
Median reliability score	.72		
Mean reliability score	.76		

The CFA and ESEM representations of the revised MES model were next compared, to determine whether an acceptable fit to the research sample had been achieved. The CFA model overall fit indices suggested poor model fit (CFI = .875, TLI = .832, and RMSEA = .077, 90% CI [.071, .083]), however, of significant contrast was the ESEM representation. This model evidenced acceptable to excellent overall fit (CFI = .972, TLI = .938, and RMSEA = .047, 90% CI [.038, .055]; see Table 22) and represents an equivalent model to the EFA configuration, differing only in regard to rotation. Analysis of item loadings further substantiated the model's suitability, all proving to be statistically significant and adequate in size, values ranging from .37 to .91 (Mean = .61; Median = .61).

Table 22 *Model Fit Indices for the CFA and ESEM Representations of the Revised Three-Factor MES*

MES Model	χ^2	*df*	CFI	TLI	RMSEA	90% CI
CFA	460	41	.875	.832	.077	.071, .083
ESEM	119	25	.972	.938	.047	.038, .055

Note. Df = Degrees of freedom; CFI = comparative fit index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square of Approximation.

Correlations between the three factors ranged from *r* = .22 to .50 (Mean *r* = .35; Median *r* = .34), reflecting construct differentiation. Figure 11 shows the model's final structural representation, while Table 23 shows the loadings and factor correlations for the revised MES instrument scales. In light of the above, the three-factor MES ESEM configuration was supported.

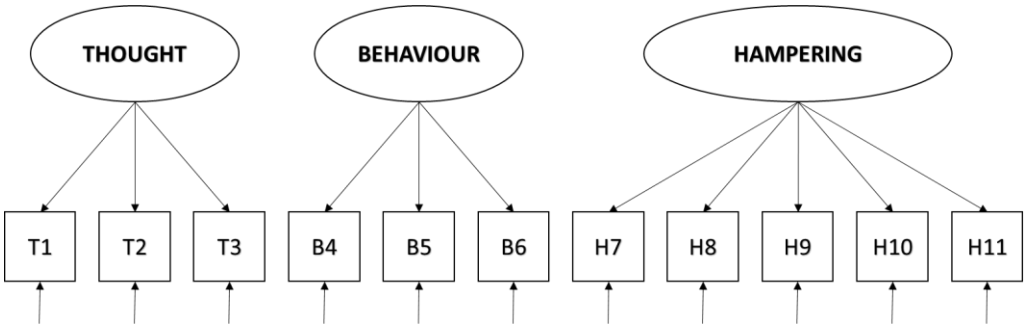


Figure 11. Revised three-factor structure of the MES.

GHQ-12. The two-factor GHQ-12 was subjected to CFA and ESEM analysis to assess structural validity. The CFA showed excellent fit (CFI = .960, TLI = .950, and RMSEA = .042, 90% CI [.036, .048]). The ESEM further indicated excellent fit to the data (CFI = .965, TLI = .946, and RMSEA = .043, 90% CI [.037, .050]; see Table 24); this was comparable to the CFA on each index. All item-loadings for the CFA model were statistically significant and sufficient in size, values ranging from .46 to .77 (Mean = .63; Median = .61). The correlation between the GHQ-12 Positive and Negative factors (see Chapter 4: General Method for descriptions) was *r* = .70. The ESEM representation evidenced near equivalent loadings, all reaching statistical significance and ranging from .48 to .76 (Mean = .62; Median = .63), while the factors correlated at *r* = .67.

Table 23 *Item Loadings and Factor Correlations for the Revised MES When Applying ESEM to Stacked Data*

	Factor Loadings		
	Thought	Behaviour	Hampering
1	.71		
2	.58		
3	.42		
4		.75	
5		.91	
6		.47	
7			.37
8			.55
9			.61
10			.62
11			.72
	Factor Correlations		
Thought	*1.00*		
Behaviour	*0.50*	*1.00*	
Hampering	*0.34*	*0.22*	*1.00*

Note. All parameters were estimated using stacked data and are reported in standardized format. Each factor was inferred on the basis of three to five measured variables, labelled 1-11 in the upper left column. Significant factor correlations have been italicized.

Table 24 *Comparison Between the CFA and ESEM Fit Indices for the GHQ-12*

GHQ-12 Model	χ^2	*df*	CFI	TLI	RMSEA	90% CI
CFA	217.51	53	.960	.950	.042	.036, .048
ESEM	185.42	43	.965	.946	.043	.037, .050

Note. df = Degrees of freedom; CFI = comparative fit index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square of Approximation.

Based on these findings, the CFA representation for the two-factor GHQ-12 instrument was selected for ongoing analysis, given the model's relative equivalence in overall fit, loadings and factor correlation to the ESEM representation. It is on the basis of

parsimony that the more restricted CFA was preferred. Table 25 provides the item loadings and factor correlations for this model.

Table 25 *Item Loadings and Factor Correlations for the CFA Representation of the GHQ-12 Positive and Negative Symptom Scales Using Stacked Data*

Factor Loadings		
	Positive	Negative
1	.46	
2	.61	
3	.56	
4	.59	
5	.59	
6	.65	
7		.61
8		.62
9		.61
10		.74
11		.77
12		.72
Factor Correlations		
Autonomy	*1.00*	
Competence	*0.70*	*1.00*

Note. All parameters were estimated using stacked data and are reported in standardised format. Each factor was inferred on the basis of six measured variables, labelled 1-12 in the upper left column. Significant factor correlations have been italicized.

Unidimensional Measures. To determine whether the single-factor measures used in the research instrument demonstrated acceptable fit to the participant sample, an integrative four factor model (see Chapter 4: General Method) comprising the SWEMWBS, GQ-6, Resilience and Life Satisfaction scales was subjected to SEM validity procedures.

The CFA overall fit statistics suggested excellent fit to the data (CFI = .970, TLI = .965, and RMSEA = .037, 90% CI [.033, .040]). The alternative ESEM representation

further reflected excellent fit (CFI = .985, TLI = .978, and RMSEA = .030, 90% CI [.025, .034]), with improvements seen on all indices (see Table 26). Exploration of item-loadings for the CFA model found each to be statistically significant, but a single variable for the five-item GQ-6 scale fell below the target limit of > .3, all loadings ranging from .22 to .90 (Mean = .73; Median = .76). For the GQ-6, the other four items reflected loadings of substantial size, values ranging from .75 to .90. As such, the factor was deemed to be adequately identified. As expected, each of the factors within the CFA model showed statistically significant correlations, values ranging between r = .18 and .67 (Mean r = .42; Median r = .40).

Table 26 *Comparison of Fit Indices Between the CFA and ESEM Models for the Combined Single-Factor Scales*

Single-Factor Scales	χ^2	*df*	CFI	TLI	RMSEA	90% CI
CFA	499.22	147	.970	.965	.037	.033 .040
ESEM	297.19	117	.985	.978	.030	.025 .034

Note. df = Degrees of freedom; CFI = comparative fit index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square of Approximation.

In line with previously detailed strengths and limitations of ESEM compared to CFA, the ESEM representation for the four-factor model demonstrated marginally reduced loadings (.25 to .88; Mean = .71; Median = .75), but smaller factor correlations also (r = .03 to .62; Mean r = .35; Median r = .33). Again, a single item within the GQ-6 scale loaded below the .3 cut-off, however the other four items sufficiently accounted for this and it was decided that the factor was adequately identified, the remaining loading values ranging from .61 to .82 (Mean = .71; Median = .69).

The ESEM representation for the four-factor model containing the research unidimensional instruments was selected for ongoing analysis on the basis of the model's

substantially better fit to the data, acceptable loadings and lower factor correlations compared to running CFA. Table 27 provides the item loadings and factor correlations.

Table 27 *Item Loadings and Factor Correlations for the ESEM Representation of the Combined Single-Factor Instruments Model When Applying Stacked Data*

Factor Loadings				
	Well Being	Gratitude	Resilience	Life Satisfaction
1	.57			
2	.76			
3	.69			
4	.77			
5	.75			
6	.61			
7	.62			
8		.76		
9		.82		
10		.25		
11		.61		
12		.63		
13			.75	
14			.76	
15			.88	
16			.86	
17			.84	
18			.86	
19				1.00
Factor Correlations				
Well Being	*1.00*			
Gratitude	*0.16*	*1.00*		
Resilience	*0.62*	*0.03*	*1.00*	
Life Satisfaction	*0.62*	*0.15*	*0.50*	1.00

Note. All parameters were estimated using stacked data and are reported in standardised format. Each factor was inferred on the basis of up to seven measured variables, labelled 1-7 in the upper left column. Significant factor correlations have been italicized.

Conclusion. Each measure used in the research instrument was compared in relation to CFA and ESEM configurations, to determine the model best fitting the research

sample. Results from the analysis demonstrated all instruments finding an acceptable SEM representation, as evidenced by overall fit indices, scale loadings and factor correlations. With regard to the MES instrument, neither SEM set-up adequately represented the *a priori* four factor model. EFA was thus applied, and an adapted three-factor model was supported. Furthermore, the ROPELOC Overall Effectiveness scale was deemed not to be a distinct factor and is not interpreted henceforth. The scale showed high correlations with other theoretically related ROPELOC scales and was poorly defined by its items. Thus, it was presumed that the variance from the Overall Effectiveness factor had been absorbed into the other scales. Results saw the GHQ-12 and SDT instruments retained the CFA representation, while the SDQ, ROPELOC, MES and integrated unidimensional measures favoured the less restrictive ESEM configuration. With all instruments reflecting acceptable structural validity, hypothesis 1.2 was supported and accepted.

Hypothesis 1.3: Measurement invariance over time.

Overview. Hypothesis 1.3 postulated that each measure within the research instrument would reflect measurement stability over time. To test this, a series of longitudinal invariance models were evaluated, using the T0 and T1 data-waves for the total participant sample. Each instrument retained its preferred CFA or ESEM configuration from hypothesis 1.2. Model 1 (configural) represented the baseline model with no constraints on parameter estimates. Model 2 (metric) comprised the weak factorial model, whereby factor loadings were constrained to be equal. Model 3 (scalar), the strong factorial model, constrained both factor loadings and indicator intercepts to be invariant, making it the most parsimonious and restrictive model. In each of the three models, uniquenesses of parallel items were correlated, as were the uniquenesses of similarly worded items.

Following recommendations by Cheung and Rensvold (2002) and Chen (2007), the CFI, TLI and RMSEA were used to systematically compare each model, given their superiority to chi-squared values in measurement invariance. These authors suggest that a decrease in CFI of ≤ -.01 (Cheung & Rensvold, 2002) or an RMSEA increase of ≥ -.01 for more parsimonious models indicates that the null hypothesis for invariance should be accepted, and result in rejection of the more constrained model. Generally speaking, imposing such constraints on a model should have negative effects on fit (Yin & Fan, 2003), however, Marsh, Hau, and Wen (2004) caution that a penalty for lack of parsimony on TLI and RMSEA indices can lead to improved fit with more restrictive models and therefore that such cut-offs should be considered as guidelines, rather than being set in stone.

SDQ II/III. Evaluating factor invariance across the T0 and T1 time-waves involved comparing the TLI, CFI and RMSEA indices across three models. As shown in Table 28, the unconstrained configural model reflected generally good fit (CFI = .916; TLI = .894; RMSEA = .044), however the TLI fell marginally below the generally accepted limit of .90.

Table 28 *Longitudinal Invariance of the SDQ Instrument Between T0 and T1 When Applying ESEM*

Model	χ^2	*df*	CFI	TLI	RMSEA
1 Configural	4034.44	2261	.916	.894	.044
2 Metric	4292.52	2478	.914	.901	.043
3 Scalar	4318.91	2509	.914	.903	.042

Note. Configural = Parameters Freely Estimated; Metric = Factor loadings constrained to be equal; Scalar = Factor loadings and Intercepts constrained as equal. In each model, correlated uniquenesses of parallel items were constrained to be equal.

To examine invariance at the factor loading level, constraints were imposed on all loadings (metric). Interestingly, the effect of constraining factor loadings saw marginal improvements in TLI at the second decimal place, leading all fit indices to exceed cut-offs for goodness of fit (CFI = .914; TLI = .901; RMSEA = .043). The marginal improvement in TLI is reflective of the penalty for lack of parsimony, which can result in a better fit than would a less restrictive model.

Following recommendations of Cheung and Rensvold (2002) and Chen (2007), changes in CFI, TLI and RMSEA were examined to determine whether the more parsimonious model (metric) should be retained. Compared to the baseline configural model, the metric model values varied by discrete amounts (△ CFI = -.002; △ TLI = .007; △ RMSEA = -.001; see Table 28), which did not exceed cut-off limits. Therefore, it was concluded that factor loadings were adequately invariant across time.

The scalar model was next examined by holding factor loadings and intercepts invariant. Results showed all fit indices falling in an acceptable range (CFI = .914; TLI = .903; RMSEA = .042). Comparison to the less parsimonious metric model revealed little change in fit statistics (△ CFI = .000; △ TLI = .002; △ RMSEA = -.001; see Table 28) following the imposition of additional constraints to intercepts. As such, the most parsimonious scalar model was retained and strong measurement invariance across time was achieved.

ROPELOC. As illustrated in Table 29, the ROPELOC instrument demonstrated acceptable fit across configural (CFI = .931; TLI = .902; RMSEA = .044), metric (CFI = .922; TLI = .900; RMSEA = .042), and scalar (CFI = .923; TLI = .904; RMSEA = .042) models.

Table 29 *Longitudinal Invariance of the ROPELOC Instrument Between T0 and T1 When Applying Set-ESEM*

Model	χ^2	*df*	CFI	TLI	RMSEA
1 Configural	4153.28	2454.00	.931	.902	.044
2 Metric	4683.65	2742.00	.922	.900	.042
3 Scalar	4684.93	2784.00	.923	.904	.042

Note. Configural = Parameters Freely Estimated; Metric = Factor loadings constrained to be equal; Scalar = Factor loadings and Intercepts constrained as equal. In each model, correlated uniquenesses of parallel items were constrained to be equal.

A comparison of the configural and metric models saw marginal drops in index values which did not exceed cut-off limits (△ CFI = -.009; △ TLI = -.002; △ RMSEA = -.002; see Table 29). Therefore, invariance at the factor-loading level was established.

A final comparison was made between the metric and scalar models. The analysis revealed no changes to indices (△ CFI = .001; △ TLI = .004; △ RMSEA = .000; see Table 29), leading to the most parsimonious scalar model being accepted and strong measurement invariance over time inferred.

BPNSFS. As illustrated in Table 30, the CFI, TLI and RMSEA indices reflected excellent fits to the data for the unconstrained configural (CFI = .975; TLI = .969; RMSEA = .040), the metric (CFI = .975; TLI = .971; RMSEA = .039); and the scalar (CFI = .975; TLI = .972; RMSEA = .038) representations.

Table 30 *Longitudinal Invariance of the BPNSFS Instrument Between T0 and T1 When Applying CFA*

Model	χ^2	*df*	CFI	TLI	RMSEA
1 Configural	371.30	225	.975	.969	.040
2 Metric	376.62	234	.975	.971	.039
3 Scalar	386.71	243	.975	.972	.038

Note. Configural = Parameters Freely Estimated; Metric = Factor loadings constrained to be equal; Scalar = Factor loadings and Intercepts constrained as equal. In each model, correlated uniquenesses of parallel items were constrained to be equal.

Compared to the baseline configural model, the metric model values varied only marginally (△ CFI = .000; △ TLI = .002; △ RMSEA = -.001; see Table 30) and did not approach cut-off values. Furthermore, the chi squared difference test was non-significant, again suggesting acceptance of the more parsimonious model, giving rise to the conclusion that factor loadings across time were adequately invariant.

The scalar model was contrasted to the metric model. As shown in Table 30, the results again showed little change in overall fit statistics (△ CFI = .000; △ TLI = .001; △ RMSEA = -.001), leading to the most parsimonious scalar model being accepted.

MES. Table 31 illustrates the overall fit indices for each model falling in the acceptable to excellent range. The unconstrained configural model fit was CFI = .940, TLI = .908, and RMSEA = .051, while the metric model evidenced values of CFI = .947, TLI = .930, and RMSEA = .045, and the scalar model CFI = .949, TLI = .936, RMSEA = .043. Compared to the baseline configural model, the metric model values saw marginal improvements for all indices (△ CFI = .007; △ TLI = .022; △ RMSEA = -.006; see Table 31) and a non-significant value for the chi squared difference test.

Table 31 *Longitudinal Invariance of the MES Instrument Between T0 and T1 When Applying ESEM*

Model	χ^2	*df*	CFI	TLI	RMSEA
1 Configural	312	151	.940	.908	.051
2 Metric	317	175	.947	.930	.045
3 Scalar	319	183	.949	.936	.043

Note. Configural = Parameters Freely Estimated; Metric = Factor loadings constrained to be equal; Scalar = Factor loadings and Intercepts constrained equal. In each model, correlated uniquenesses of parallel items were constrained to be equal.

The more parsimonious metric model was accepted and subsequently contrasted to the scalar representation. Results revealed little change in index values (△ CFI = .002; △

TLI = .006; △ RMSEA = -.002; see Table 31), leading the most parsimonious scalar model to be retained and the three-factor MES to evidence strong measurement invariance over time.

GHQ-12. The results following longitudinal invariance testing for the GHQ-12, reflected in Table 32 suggested acceptable overall fit for each configuration. The unconstrained configural model evidenced CFI = .919, TLI = .905, and RMSEA = .049, while the metric model revealed index values of CFI = .921, TLI = .910, and RMSEA = .047, and the scalar model CFI = .919, TLI = .912, RMSEA = .047. Compared to the configural model, the metric model values saw few variations in index values which did not exceed cut-off limits (△ CFI = .002; △ TLI = .005; △ RMSEA = -.002; see Table 32. The metric and scalar models were contrasted and revealed only marginal changes in two index values (△ CFI = -.002; △ TLI = .002; △ RMSEA = .000). On this basis, the most parsimonious scalar model was accepted and the GHQ-12 was determined as evidencing strong measurement invariance over time.

Table 32 *Longitudinal Invariance Analysis of the GHQ-12 Two-Factor Solution Using CFA*

Model	χ^2	*df*	CFI	TLI	RMSEA
1 Configural	459.846	234	.919	.905	.049
2 Metric	466.024	244	.921	.910	.047
3 Scalar	3078.303	276	.919	.912	.047

Note. Configural = Parameters Freely Estimated; Metric = Factor loadings constrained to be equal; Scalar = Factor loadings and Intercepts constrained as equal. In each model, correlated uniqueness's of parallel items were constrained to be equal.

Unidimensional Instruments. The goodness of fit indices shown in Table 33 show that the configural (CFI = .923, TLI = .900, RMSEA = .041), metric (CFI = .936; TLI = .925; RMSEA = .036); and scalar models (CFI = .939; TLI = .930; RMSEA = .035) were adequately represented by the research sample.

Comparison between the configural and metric models saw marginal improvements across each index value, and so the metric model was accepted (△ CFI = .013; △ TLI = .025; △ RMSEA = .005). Comparison of this to the scalar model again saw improvements in fit indices (△ CFI = .003; △ TLI = .005; △ RMSEA = -.001; Table 33). As such, the most parsimonious scalar model was accepted and strong measurement invariance established.

Table 33 *Longitudinal Invariance Analysis for the Combined Unidimensional Scales using ESEM*

Model	χ^2	*df*	CFI	TLI	RMSEA
1 Configural	1596.94	946	.923	.900	.041
2 Metric	1586.70	1046	.936	.925	.036
3 Scalar	1590.23	1071	.939	.930	.035

Note. Configural = Parameters Freely Estimated; Metric = Factor loadings constrained to be equal; Scalar = Factor loadings and Intercepts constrained as equal. In each model, correlated uniquenesses of parallel items were constrained to be equal.

Conclusion. Hypothesis 1.3 was evaluated by running systemic comparisons of increasingly constrained structural models for each instrument, to establish their equivalent longitudinal measurement. In each instance, the final and most constrained model (scalar) demonstrated acceptable goodness of fit on CFI, TLI and RMSEA indices. As such, all instrumentation evidenced strong measurement invariance within the research sample, wherein slopes and intercepts were comparable over time.

Hypothesis 1.4: Structural integrity of the combined instrument model.

Overview. Hypothesis 1.4 predicted that the structural integrity of the combined research instrument would be maintained when all instruments were administered in a single battery. To evaluate this, all measures were integrated into a single model while retaining their CFA (or otherwise ESEM) configuration from the hypothesis 1.2 analysis.

Stacked data were used simultaneously for the 32 scales and their respective 131 items, to produce estimates of overall model goodness of fit (see Table 34).

Table 34 *Model Fit Indices for the Combined Research Instrument When Applying Stacked Data*

Full Model	χ^2	*df*	CFI	TLI	RMSEA	90% CI
CFA/ESEM	15434.953	7750	.935	.925	.024	.023, .024

Note. df = Degrees of freedom; CFI = comparative fit index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square of Approximation.

Results. The combined model produced overall fit indices of CFI = .935, TLI = .925, and RMSEA = .024, demonstrating acceptable to excellent structural validity to the research sample. See Appendix 3 for scale-loadings and descriptive statistics, and Appendix 5 for model syntax, goodness of fit statistics and item loadings.

Conclusion. Structural validation procedures using the combined instrument battery evidenced acceptable goodness of fit to the research sample. This encouraging result suggests that individual measures within the broader research instrument were adequately differentiated, and maintained their structural integrity when embedded with other scales.

Hypothesis 1.5: Multitrait-multimethod analysis.

Overview. Following recommendations by Marsh, Ellis, et al. (2005), MTMM analysis was administered as the final psychometric procedure, to ensure within-construct validation issues had been addressed. In the present study, traits are represented by the 32 outcome scales contained in the survey tool, while variation in method is reflective of the T0 and T1 time-waves, resulting in a 64 x 64 matrix of all latent constructs (32 traits by two methods). The analysis was based on factor scores which were derived from the hypothesis 1.4 combined structural model analysis.

As previously noted, application of the Campbell and Fiske (1959) guidelines to a latent matrix of correlations overcomes widely noted limitations on the use of these

heuristic guidelines. As previously stated, the ROPELOC Overall Effectiveness scale was excluded from MTMM analysis.

Convergent validity. Support for convergent validity specifies that the correlations between the 32 MTHM convergent validities (same traits at different time-waves) should be statistically significant and substantial. Appendix 4 shows these highlighted in dark grey in the central diagonal of the matrix within the bold box. There is clear support for this criterion in that all 32 convergent validities were statistically significant and substantial (*r*s = .473 to .780; *M r* = .639).

Discriminant validity. According to the Campbell and Fiske (1959) guidelines, there are two main criteria used to assess discriminant validity. The first stipulates that convergent validities (the same traits measured on different occasions) are higher than correlations between non-matching traits, the heterotrait-heteromethod correlations (HTHM; different traits and different methods; see Appendix 4 for the off-diagonal values in the bolded box). There is clear support for this criterion in that the 992 HTHM correlations (*r*s = .002 to .665, *M r* = .344) were substantially smaller than the convergent validities *r*s = .473 to .780; *M r* = .639.

Campbell and Fiske (1959) also propose that each convergent validity should be compared with all the other HTHM correlations involving the same trait. In the present investigation, each of the 32 convergent validities was compared with the 62 HTHM traits, a total of 32 x 62 = 1,984 comparisons. Inspection of Appendix 4 shows that this criterion was met for 1,979 of the 1,984 comparisons (99.7%). The ROPELOC instrument saw two HTHM correlations exceeding their MTHM equivalent, observed between the T0 Coping with Change and T1 Stress Management scales (*r* = .588; MTHM *r* = .571); and the T0 Open Thinking and T1 Quality Seeking scales (*r* = .508; MTHM *r* = .495). Additionally, the GHQ-12 Positive symptoms score at T0 reflected a marginally stronger correlation

with T1 Wellbeing (r = .617; MTHM r = .579), while the T0 score for the Resilience scale correlated more highly with the T1 ROPELOC Stress Management (r = .595) and SWMWBS Wellbeing (r = .608) scales compared to their MTHM correlation (r = .585).

The second criterion of discriminant validity specifies that convergent validities are higher than heterotrait-monomethod correlations (HTMM; different traits and same methods). Appendix 4 shows these for the values below the main diagonal in the T0 and T1 triangular submatrices shaded in light grey. Again, there is good support in that on average, the convergent validities (M r = .639) are larger than the 496 HTMM correlations at T0 (rs = .000 to .858; M r = .409) and 496 correlations at T1 (rs = .001 to .874; M r = .443); a total of 976 criteria met from 992 comparisons (98.4%).

Within the framework it is useful to test for halo effects within each method (time-wave). This is identified by HTMM correlations being systematically higher than HTHM correlations. Results saw the T0 (M r = .409) and T1 (M r = .443) HTMM correlations comfortably exceeding the HTHM correlations (M r = .344). Of interest, both T0 and T1 found eight of their respective 496 HTMM correlations to exceed .80, suggesting factors were comparatively differentiated at each time-wave.

Conclusion. MTMM procedures provided strong support for both the convergent and discriminant validity of participant responses, in relation to stability over time. In particular, all 32 convergent validities were substantial, and support for the two criteria of discriminant validity was met for 1,979 of the 1,984 comparisons.

Summary of Results

Chapter 5 set out to establish the robustness of all measures used in the present investigation by subjecting participant data to a battery of psychometric evaluations. Reliability testing saw all scales exceeding recommended estimates (> .7) for internal consistency, while SEM procedures saw CFA and ESEM model configurations compared

using each instrument's *a priori* factor structure, with each measure achieving a configuration of acceptable fit to the data. The originally proposed MES four-factor model however was unable to converge, prompting the use of EFA and a newly derived three-factor alternative, which exceeded all indicators of acceptable fit. For each instrument, SEM configurations evidenced strong measurement invariance over time, while MTMM analysis using latently derived factor scores offered strong support for convergent and discriminant validity. The research administration schedule required participants to complete all outcome measures in a single battery, reflecting the need to establish whether each instrument's latent SEM structure holds when administered in conjunction to the others. SEM confirmed this to be the case, and thus the full integrated model containing all instruments was used to derive weighted factor scores for all scales at each time-wave. These factor scores subsequently formed the basis for the Study 2 analysis and results. The following chapter turns to this Study.

CHAPTER 6: STUDY 2—INTERVENTION

Introduction

Since the mid-19th century, outdoor adventure education (OAE) programs have provided an acclaimed framework for engaging youths in enriching experiences. This has led to widespread agreement that OAE leads to a variety of health benefits to individuals of diverse backgrounds. However, the notable lack of empirical grounding for these claims led Neill (2002) to warn against an over-reliance on anecdotal indicators of OAEs effectiveness that is not supported by strong empirical research.

In spite of Neill's warning, only a small percentage of OAE programs are designed to undergo empirical evaluation (Bowen & Neill, 2013). Consequently, the lack of robust evaluation of OAE outcomes has placed the value of OAE programs in question (Dillon, 2013; O'Brien et al., 2011). Frequent criticisms surround the lack of appropriate comparison groups and follow-up measures. In addition, poor measurement and inappropriate statistical models further undermine the largely favourable findings emerging from OAE investigations.

The content, aims, and design of OAE programs vary substantially between contexts and it is of no surprise that sizable discrepancies in effect sizes have been observed for equivalent outcomes (Bowen & Neill, 2013). This has led to the recognition that research needs to consider the underlying mechanisms on which the effectiveness of OAE is founded, to consolidate what has been described as a fragmented body of literature. As noted in Chapter 2: Literature Review, Self Determination Theory (SDT) provides such a framework, and has been frequently applied to youth interventions (Jang et al., 2012; Jang, Reeve, Ryan, & Kim, 2009), including OAE programs (Rahman, 2009; Wang et al., 2004). This investigation draws on SDT theory as a mediating concept of and facilitating mechanism for OAE program effectiveness.

The influence of individual differences on how students learn through OAE is poorly understood, however is a pertinent consideration for the development of effective, tailored OAE programming (Neill, 2002). One such area of difference concerns individual's pre-treatment aptitude for social and emotional functioning and raises the question of whether variations in such functioning influence students' sensitivity to benefit in goal-congruent OAE domains. Leading OAE meta-analyses report clinical scales concerning the measurement of psychopathology indicators produce greater effect sizes compared to non-clinical scales, however comparisons of effect sizes derived from clinical measures yield comparable effect sizes in normative versus delinquent adolescent samples (e.g. Bowen & Neill, 2013; Cason & Gillis, 1994; Hattie et al., 1997; Neill, 2003). Thus, elevated effect sizes appear to be a function of construct measurement or malleability, rather than being determined by participants' baseline functionality. Aligning with this notion are findings pertaining to self-concept/self-esteem and LoC outcomes, where OAE meta-analyses again produced comparable effect sizes for normative (Hans, 2000; Neill, 2003) versus juvenile/clinical (Wilson & Lipsey, 2000) adolescent samples. Despite this suggestion that OAE programs perhaps offer similar benefits to attendees regardless of pre-treatment functioning, this has not been rigorously tested. Thus, this research quantitatively examines students' sensitivity to benefit from the OAE intervention, in relation to their baseline aptitude on each of the 32 outcomes investigated.

Employing an experimental randomized controlled trial longitudinal design, supported by psychometrically strong instrumentation (see Chapter 5: Psychometric Analysis) and sufficient sample size (see Scrutton & Beames, 2015; also see Chapter 7: Discussion and Conclusion, Methodological contributions and Chapter 4: General Method, power analysis), it is anticipated that this research will contribute much-needed evidence

regarding the effects of OAE on teenagers' psychological development, to determine how the most effective OAE learning experiences can be achieved.

Rationale for Research Outcomes

The alignment between OAE program structure, philosophy and aims provides a crucial determinant of the outcomes one might expect to see from OAE experiences (Dillon, 2013; Hans, 2000; Rickinson et al., 2004). The constructs selected for this investigation were derived from extensive analysis of the OAE intervention program and its specific aims (see Chapter 3: The Intervention Program).

The five-month OAE program studied in this research aims to provide a comprehensive experiential learning experience that fosters academic, physical/mental health, inter- and intra-personal competencies. Students learn about themselves and the external environment in a nurturing and self-reflective climate that encourages individuals to extend their capacities. The notion of personal responsibility is a dominant theme throughout the OAE program, where students engage in challenges that promote independence, personal effectiveness and enhanced life skills. Taking place in a residential bushland setting, the program teaches students to live with peers from various backgrounds and to cooperate, lead, work as team members and support one another in a social climate that emphasises camaraderie instead of competitiveness. Additionally, the campus is devoid of media technology (e.g. mobiles, TV, video game devices), while family bonds are enriched through regular letter-writing and two scheduled 'parent-hike' weekends.

Building on the well-defined content, structure and aims of the OAE program (see Chapter 3: The Intervention Program), this investigation selected psychological measures of self-concept, life effectiveness and locus of control (LoC) as the primary research outcomes. Secondary outcomes included measures of wellbeing/ill-being, resilience, life satisfaction, gratitude, and school motivation and engagement.

Statement of Problems and Overarching Aims

Compared to waitlist control participants, does the OAE intervention program lead to statistically significant post-treatment gains in the research outcomes, and if so, will these effects be maintained at six-month follow-up? Are the benefits from the OAE program equivalent for all participants, irrespective of baseline aptitude on each outcome, and does students' satisfaction of basic psychological needs for autonomy, relatedness and competence mediate the treatment gains? This statement of problems relates to the following self-report measures, as detailed in Chapter 4: General Method:

1. The Self-Description Questionnaire-II (SDQII), as a measure of six domains of adolescent self-concept and the Self-Description Questionnaire-III (SDQIII) as a single-domain addition (Problem Solving);
2. Review of Personal Effectiveness with Locus of Control (ROPELOC), as a measure of 12 facets of life-effectiveness and a two-factor measure of students' perceptions of internal versus external control expectancies (LoC);
3. Basic Psychological Needs Satisfaction and Frustration Scale (BPNSFS; adaptation), as a measure of adolescents' satisfaction with the three basic psychological needs posited within Self Determination Theory;
4. Motivation and Engagement Scale (MES), as a multidimensional measure of adolescent motivation and engagement;
5. General Health Questionnaire-12 (GHQ-12), as a clinical indicator of psychological ill-being;
6. Personal Wellbeing Index-School Children (PWI-SC), a single item used as a measure of student global satisfaction with life;
7. The Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS), as a unidimensional measure of students' eudemonic and hedonic wellbeing;

8. The Gratitude Questionnaire (GQ-6; adaptation), as a five-item measure of adolescence proneness to experiencing gratitude in daily life.
9. Academic Resilience Scale (adaptation), as a six-item unidimensional measure of students' capacity to effectively respond to setbacks, challenges and adversity in general life.

The overarching aim of Study 2 was to investigate the immediate and lasting effects of the OAE intervention program in adolescent development, and to explore the underlying mechanisms through which these transpire. The following hypotheses detail the procedures used to achieve this aim.

Statement of Hypotheses and Rationale

Hypothesis 2.1: Experimental effects of the OAE intervention. Major OAE reviews and meta-analytic research has found small to moderate short-term effects for a variety of positive psychological variables in non-clinical adolescent samples (e.g., Bowen & Neill, 2013; Cason & Gillis, 1994; Hattie et al., 1997; Neill, 2008). Within the literature, areas of self-concept, LoC and personal effectiveness have received particular attention. This is due to these constructs' close theoretical alignment with the rich psychosocial experiences that OAEs' experiential modality and challenge-based learning provide. Additionally, the measured benefits of employing clinical measures in OAE research consistently exceed those detected using positive psychological measures (e.g., Bowen & Neill, 2013; Cason & Gillis, 1994; Neill, 2003). Accordingly, it was hypothesised that students attending the OAE intervention will display statistically meaningful beneficial changes in domains of self-concept, life effectiveness, LoC and secondary personal resource variables (see Chapter 4: General Method) at the end of the outdoor program (T2) when compared to WLC participants. Furthermore, that the greatest benefits will be

detected on clinical outcomes of psychological distress (negative scale) and social-emotional dysfunction (positive scale) measured by the GHQ-12.

Hypothesis 2.2: Lasting effects of the intervention (T2 to T3). Over the past decade, OAE research has responded to criticisms of the absence of longitudinal investigations (see Chapter 2: Literature Review). The Bowen and Neill (2013) synthesis of OAE meta-analytic research observed that treatment gains across a variety of psychological and prosocial outcomes were maintained at follow-up, a finding similarly reported in the Gillis and Speelman (2008) analysis of 12 OAE controlled studies. Hattie et al. (1997) found lasting benefits from OAE programs, with the greatest follow-up effect sizes occurring in self-concept. SDT explains these effects through the operation of internalisation (Ryan & Deci, 2017; see Chapter 2: Literature Review, Organismic integration theory). Accordingly, it was predicted that post-treatment experimental effects (T2) will be maintained at six-month follow-up (T3) and that statistically significant additional gains will be observed over the post-treatment interval, as explained by SDT sleeper effects.

Hypothesis 2.3: Aptitude-treatment interaction (T1 to T2). Following recommendations by Neill (2002), OAE meta-analytic research has attempted to better understand how different participant characteristics influence outcomes and effect sizes from OAE experiences. Cason and Gillis (1994) found no significant differences in the benefits experienced by normal youth populations compared to individuals with emotional or behavioural problems. Hattie et al. (1997) replicated these findings, whereby post-treatment effect sizes differed only marginally between delinquent (ES = .33) and normative (ES = .35) youth samples, while Neill (2003) found OAE programs yielded an overall ES = .31 for representative adolescents, compared to ES = .33 seen in delinquent client groups. Similar findings have been shown using self-concept and LoC measures

(Hans, 2000; Neill, 2003; Wilson & Lipsey, 2000). These results suggest that variations in participants' baseline functioning do not adequately explain the likely benefits that one may attribute to OAE programming. Accordingly, it was hypothesised that research participants' pre-treatment aptitude on each dependent variable would not predict the degree of change observed in that outcome over the treatment interval (T1 to T2). That is, no Aptitude-Treatment Interactions (ATI) from pre-treatment (T1) to post-treatment (T2) will be observed for research outcomes.

Hypothesis 2.4: Mediation of the experimental effects. Self Determination Theory provides an empirically derived framework which describes three universal psychological needs that are held crucial for optimal human development and functioning (e.g., Deci & Ryan, 1985; Ryan & Deci, 2000b; Ryan & Deci, 2000c, 2017). In this theory, satisfaction of an individual's psychological needs for autonomy, relatedness and competence facilitates more intrinsically motivated behaviour, which in turn fosters greater internalisation of learning experiences. Thus, supporting students' basic psychological needs is argued to be critical to the success of the experiential treatment protocol (see Chapter 2: Literature Review). Intervention research has frequently examined the mediating role of SDT concepts within treatment contexts, with a growing body of literature suggesting that psychological needs satisfaction constitutes a vital mechanism through which interventions achieve their goals (e.g., Jang et al., 2012; Jang et al., 2009; Wang et al., 2004). It was therefore hypothesised that the OAE program will foster significant enhancements in SDT's basic psychological needs of autonomy, relatedness and competence compared to WLC participants at T2, and that psychological needs satisfaction will mediate the effects of the OAE program on T2 research outcomes.

Summary. Meta-analyses have demonstrated that OAE programs supersede a range of alternative treatments in their capacity to benefit areas of psychosocial and emotional

functioning (Bowen & Neill, 2013). However, the substantial economic and human capital required to implement such programs has led to the need for a strong, empirically-founded justification. The notable lack of robust research designs using appropriate comparison groups and follow-up measures has contributed to a decline in OAE programs in the Australian curriculum, despite generally consistent findings that these programs benefit a variety of desirable psychological outcomes (e.g. Cason & Gillis, 1994; Hattie et al., 1997; Neill, 2008).

The hypotheses presented in this study address a critical gap within OAE literature by drawing on a strong randomized controlled trial and longitudinal design to contribute much-needed hard evidence regarding the effects of outdoor education on healthy adolescent development.

Method

Participants

A total of 413 male students were invited to participate in the Study 2 investigation (treatment group = 211; WLC = 202; see Chapter 4: General Method for sample *n*s at each time-wave).

Research Design

Hypotheses 2.1-2.4 were tested using a pre-test post-test longitudinal randomized controlled trial design with an extended baseline and waitlist control condition (see Figure 12). By including multiple pre-treatment measures for all participants, confounds and pre-treatment biases could be controlled for. This design aspect further allowed for rigorous preliminary analysis of randomization procedures. Hypothesis 2.2, concerning long-term treatment gains, used WLC post-test (T2) data to act as the reference group for T3 follow-up effect analyses. This was in in lieu of a viable true no-treatment control group at T3, as the WLC students had by this time embarked on the OAE program and were no longer a

viable control comparison. This is further detailed in the H2.2 results section and will be explored in Chapter 7: Discussion and Conclusion in regards to research strengths and limitations.

A randomised controlled trial. A common issue in human research is the impractical, or otherwise unethical, nature of assigning participants to treatment conditions. Consequently, very few OAE studies include controlled comparisons. The Gillis and Speelman (2008) synthesis of controlled OAE studies published between 1986-2006 identified just six education-focused OAE studies involving middle or high-school adolescents.

Consistently with the ambitious aims of this research and adhering to recommended design standards (e.g., Cason & Gillis, 1994; Dillon, 2013; Wilson & Lipsey, 2000), this investigation benefited from a randomly assigned waitlist-control condition (WLC): that is, a group expecting to receive the intervention subsequent to the treatment condition. At the approximate commencement of the final Year 8 term, allocation procedures saw students randomly assigned to attend the compulsory OAE program in either Semester 1 (treatment condition) or Semester 2 (WLC) of Year 9. As such, WLC participants provided a strong representative control group at the T0 (extended-baseline), T1 (Pre-test) and T2 (Post-test) time-waves, after which they were introduced to the treatment protocol (prior to T3 follow-up).

Data Analysis

Descriptive data were analysed using SPSS Statistics 22.0 (IBM Corp., Released 2013), while inferential statistics and more complex analyses, including all regression models, were computed using latently derived weighted factor scores in Mplus Version 8.0. (Muthén & Muthén, 1998-2017). Statistical significance was inferred using a nominal

p-value of < .05, however a more conservative valuation of significance can be derived by observing the Estimate (Est.) to Standard Error (S.E.) ratio, which exceeded 2.56 ($p < .01$).

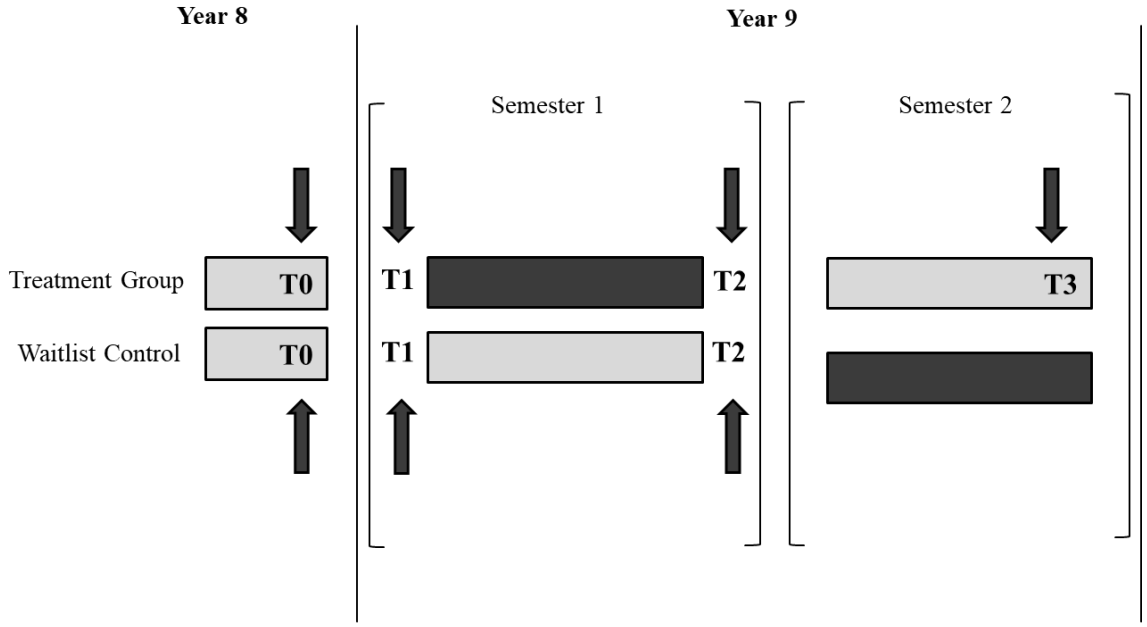


Figure 12. Graphical representation of the project data collection schedule that was replicated for the 2016 and 2017 Year 9 cohorts.

Note. Rectangles represent school semesters, while the text and vertical arrows indicate the four data collection points: T0 = Extended-baseline, T1 = Pre-test, T2 = Post-test, T3 = follow-up. The occurrence of the five-month OAE intervention for the treatment group has been shaded dark grey, while light grey rectangles indicate usual school timetable.

Reverse-scored scales. External LoC, Hampering, and GHQ-12 Positive and Negative factors were reverse-scored, so to align with all other positive psychological variables. This adjustment means that larger effect sizes in a positive direction indicate favourable effects from the OAE intervention across all 32 outcomes.

Effect sizes. The reporting of effect sizes (ESs) is now an expected standard in empirical research, to allow for quantitative comparison of intervention effects (American Psychological Association, 2010). The Cason and Gillis (1994) meta-analysis of OAE

research defines ES as “a measure of the amount of change experienced following an adventure experience” (p. 41). The standardised metric for estimating ESs in this investigation is implicit within the latently calculated weighted factor scores.

When interpreting an ES it is recommended that consideration be given to the research context, rather than adopting a universal method. Wolf (1986) differentiates between practical versus educational significance when referring to ES estimates—the latter generally carrying greater meaning at lower limits. Another consideration includes the relative sensitivity of research outcomes to change following intervention. For example, a small ES reflecting change in a practically valuable and generally robust outcome should be deemed more impressive than a larger ES for one that is less so. This research takes on such recommendations and applies the following suggested labels for ES cut-off limits when interpreting outcome effect sizes (Cohen, 1988):

ES = .1 (Small)

ES = .3 (Medium)

ES = .5 (Large)

Mediation analysis. To test hypothesis 2.4, the multiple mediation methods proposed by Preacher and Hayes (2008) were applied. Post treatment data (T2) was used for all participants so to effectively measure the OAE programs influence on students’ basic psychological needs satisfaction. This approach stipulates that investigation of multiple mediation should first determine whether the set of mediators (M) transmits the effects from X to Y, and then explore specific hypotheses regarding the individual mediators (i.e. specific indirect effects; see Figure 13).

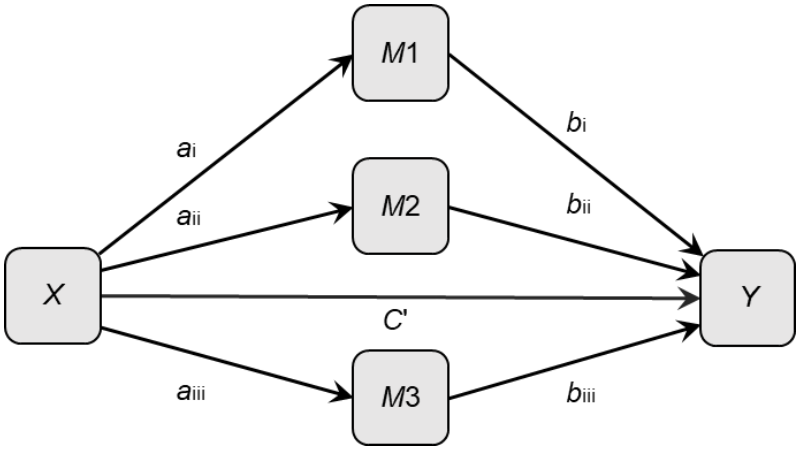


Figure 13. Multiple mediation model proposed by Preacher & Hayes (2008).

Multiple regression analyses were run through Mplus Version 8.0. (Muthén & Muthén, 1998-2017) to assess each component of the mediation model (see Figure 14). Indirect effects were computed for each of the 5,000 bootstrapped samples, using bias-corrected 95% confidence intervals. This approach was selected as it provides the most powerful method of obtaining CIs for indirect effects during complex mediation analysis, indicated by the least biased CIs, greatest sensitivity for detecting non-zero effects, and most accurate Type 1 error (Briggs, 2006; Preacher & Hayes, 2008; Williams & Mackinnon, 2008). In each model, the analysis controlled for pre-treatment group differences in each outcome (T0 and T1), and controlled for baseline psychological needs satisfaction (Autonomy, Relatedness and Competence) at T1.

Controlling for Pre-/Post-treatment Biases

With a growing body of research endeavouring to rigorously capture the effects of outdoor education programs, threats to measurement have been identified as residing within the assessment schedule. That is, participants can be prone to reporting inflated, otherwise diminished self-report scores, due to psychological and affective factors in play.

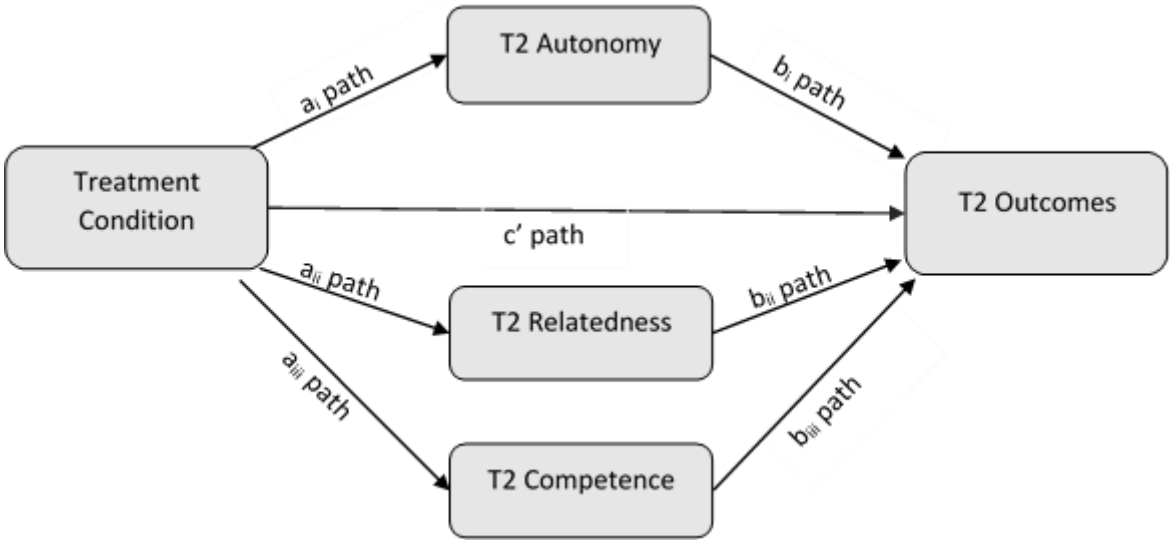


Figure 14. A priori hypothesised mediation model of SDT psychological needs satisfaction on the relationship between the OAE treatment and the dependent variables at T2 (post-test).

For example, in reference to pre-treatment biases, Hattie et al. (1997) note "It is possible that the anticipation, sense of excitement, and/or trepidation of confronting something so different and challenging may lead to depressed scores on many measures" (p. 51).

To minimise the impact of these naturally occurring psychosocial processes from impeding measurement validity, it is recommended that baseline measurement takes place several weeks prior to the intervention. A second recommendation includes the use of a control group, which allows a researcher to examine pre-treatment group differences and control for these using statistical methods.

A second source of bias was described by Marsh et al. (1986c) as post-group euphoria (PGE), and refers to the favourable biases in self-report measures due to excitement, elation and accomplishment at the completion of an OAE experience. To detect whether PGE is influencing post-treatment effects, Marsh et al. (1986c) recommend the use of control scales: that is, measures which do not align with the OAE program goals or content, and therefore should not alter as a result of the experience. If significant increases in these unrelated variables are detected at post-test, PGE may be present. A

second recommendation by Marsh et al. (1986c) to protect from PGE bias is to incorporate follow-up measures against which post-treatment results may be contrasted.

Adhering to best-practice design recommendations in leading OAE reviews, this investigation accounts for pre-treatment measurement bias by conducting an extended baseline assessment of all outcomes, screening for pre-test group differences, and controlling for baseline group effects in subsequent hypothesis testing. PGE biases, as described by Marsh et al. (1986c) were considered by the inclusion of a self-concept measure of opposite-sex relations (Marsh, 1990b) to act as a control scale. As the OAE program in this research is male only, we would not expect the participants to display genuine gains in opposite-sex relations self-concept compared to WLC students. Finally, the present investigation conducted a 6-month follow-up assessment of all outcome variables.

Preliminary Analysis

Overview. Prior to Study 2 hypothesis testing, screening of the T0 to T1 baseline period was conducted. A series of 2 (Time) by 2 (Group) factorial regression models were conducted for the 33 research outcomes. ESs with 95% CIs are presented in Table 35**,** and descriptive statistics (Ms, SDs) for both groups at T0, T1, T2 and T3 are provided in Table 36.

Results. Analysis of the T0 to T1 baseline period revealed that of 32 dependent variable outcomes, the Treatment and WLC conditions were equivalent on 29. A significant main effect for Group was observed for the ROPELOC Active Involvement (β = -.335, p = .033) scale, reflecting higher mean scores for the WLC condition at T0. A significant Time X Group interaction (β = .215, p = .011) was also found, represented by a reduction in group differences which became non-significant at T1 (p > .05). For General Academic self-concept, a statistically significant Group effect (β = .537, p = .002)

occurred in favour of the treatment condition at T0, however again a significant interaction (β = -.190, p = .033) resulted in a narrowing of group differences and non-significant effect at T1 (p > .05). A final main effect for group was observed for Physical Ability self-concept (β = -.411, p = .011), with higher mean scores occurring in WLC participants at T0. However, again a statistically significant Time X Group interaction (β = .218, p = .006) resulted in a reduction in group differences at T1, where a non-significant effect was observed (p > .05).

Conclusion. Preliminary analysis examined group and interaction effects for the treatment versus WLC condition on all outcomes over the T0 to T1 baseline period. The analysis found strong support for group equivalence, where no significant differences were observed at T0 on 29 of 32 outcomes and the remaining three differences were non-significant at T1 (pre-test). These results are highly indicative of balanced groups and successful randomization procedures. Furthermore, non-significant interactions for Group over Time were observed on 29 of 32 outcomes, suggesting a likely absence of pre-treatment anticipatory cognitive-affective biases. Despite this, subsequent hypothesis testing statistically controlled for group differences in the outcome variables at T0 and T1, thus adhering to strict guidelines that further protect the experimental analysis from confounding factors, while increasing the power of the statistical analyse.

Table 35 *Effect Sizes and Descriptive Statistics Following Repeated-Measures Analysis of the T0 to T1 Baseline Interval Showing Main Effects for Group, Time, and Interaction (Group X Time) on all Dependent Variables*

	Group-Effect		Time-Effect		Interaction (Group X Time)	
Scale	*β*	*SE*	*B*	*SE*	*β*	*SE*
ROPELOC						
Self Confidence	.070	.174	.039	.068	.046	.092
Time Efficiency	.040	.175	.037	.066	-.069	.093
Active Involvement	-.335*	.157	-.133*	.061	.215*	.085
Coop Teamwork	-.194	.176	-.070	.073	.085	.097
Open Thinking	.039	.179	-.017	.080	-.025	.104
Quality Seeking	.154	.164	-.035	.069	-.101	.092
Self-Efficacy	.031	.178	.021	.067	-.005	.094
Social Effectiveness	-.120	.163	-.061	.064	.056	.089
Stress Management	.023	.183	.027	.066	-.044	.095
Leadership Ability	-.206	.180	-.119	.073	.102	.099
Coping with Change	.047	.185	.007	.076	-.060	.101
External LoC	.097	.165	.159*	.070	.063	.090
Internal LoC	.004	.169	.012	.069	.029	.094
BPNSFS						
Autonomy	-.104	.146	.051	.060	.016	.081
Relatedness	-.051	.135	.004	.076	.013	.076
Competence	.119	.131	.008	.055	-.057	.071
SDQ						
General Academic	.537**	.173	.044	.067	-.190*	.089
Physical Ability	-.411*	.161	-.116	.059	.218**	.080
Global Esteem	.087	.178	.038	.076	.002	.096
Parent Relations	.180	.187	.065	.078	-.026	.103
Problem Solving	-.037	.184	-.077	.079	.048	.109
Sm-Sex Relations	-.048	.183	-.001	.078	.003	.102
Op-Sex Relations	-.198	.162	.074	.064	.085	.088
MES						
Thought	.152	.165	.020	.068	-.062	.089
Behaviour	-.057	.173	.011	.069	-.040	.092
Hampering	.291	.161	.147*	.063	-.051	.083

Note. β = standardised effect sizes. ROPELOC scales are: Coop Teamwork = Cooperative Teamwork, External LoC = External Locus of Control, and Internal LoC = Internal Locus of Control. SDQ scales are: Sm-Sex Relations = Same-Sex Relations, Op-Sex = Opposite-Sex Relations.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 35 *(continued)*

	Group-Effect		Time-Effect		Interaction (Group X Time)	
Scale	*β*	*SE*	*B*	*SE*	*β*	*SE*
GHQ-12						
Positive	-.018	.048	.034	.019	.014	.025
Negative	.002	.094	.099**	.034	.018	.048
One-Factor Scales						
Wellbeing	-.234	.154	.047	.058	.158	.079
Life Satisfaction	-.400	.373	.119	.134	.368	.192
Gratitude	.279	.174	.068	.073	-.122	.099
Resilience	.032	.196	.106	.081	-.029	.104

Note. β = standardised effect sizes. * $p < .05$, ** $p < .01$, *** $p < .001$.

Results

This section outlines the results from Study 2, which investigated the effects of an OAE program on adolescent self-concept, life effectiveness and LoC. Hypotheses 2.1, 2.3 and 2.4 used data collected on three occasions: approximately 10-weeks pre-program (Time 0; T0), first day (Time 1; T1), and last day (Time 2; T2; see Chapter 4: General Method). The analyses examined post-treatment experimental effects, the influence of baseline aptitude on treatment gains, and the mediating mechanisms of Self Determination Theory (SDT) basic psychological needs satisfaction on intervention effectiveness. Hypothesis 2.2 examined the long-term effects of the program by comparing treatment-group six-month follow-up data (T3) to the WLC post-test (T2) data. As described in Chapter 4: General Method, all analyses used factor score regression weights to represent participant scores on each scale at all time-waves. The results are presented in accordance with the Study 2 hypotheses described earlier in this chapter.

Table 36 *Overall Ms and SDs for the Treatment and WLC Conditions at T0, T1, T2 and T3 for ROPELOC, SDQ, BPNSFS, MES, GHQ-12, and One-Factor Scales*

		T0 (*N* = 346)		T1 (*N* = 378)		T2 (*N* = 380)		T3 (*N* = 156)	
Scale		*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
ROPELOC									
Self Confidence	Treat	6.12	1.25	6.81	1.16	6.48	1.11	6.38	1.05
	WLC	6.10	1.12	6.06	1.33	6.10	1.23	--	--
Time Efficiency	Treat	5.39	1.53	5.31	1.39	5.65	1.43	5.77	1.35
	WLC	5.43	1.42	5.50	1.40	5.36	1.49	--	--
Active Involvement	Treat	6.20	1.23	6.30	1.15	6.59	1.05	6.43	1.17
	WLC	6.35	1.21	6.14	1.31	6.35	1.25	--	--
Cooperative Teamwork	Treat	5.97	1.40	6.00	1.38	6.16	1.17	6.21	1.24
	WLC	6.13	1.40	6.01	1.44	6.09	1.45	--	--
Open Thinking	Treat	6.17	1.08	6.14	1.03	6.29	0.97	6.32	1.09
	WLC	6.13	1.14	6.11	1.12	6.23	1.12	--	--
Quality Seeking	Treat	6.51	1.10	6.42	1.16	6.66	1.07	6.52	1.12
	WLC	6.57	1.10	6.45	1.19	6.48	1.12	--	--
Self Efficacy	Treat	5.49	1.41	5.51	1.30	5.82	1.28	5.99	1.28
	WLC	5.50	1.32	5.47	1.34	5.68	1.40	--	--
Social Effectiveness	Treat	5.80	1.36	5.76	1.33	6.02	1.34	5.99	1.25
	WLC	5.87	1.26	5.77	1.38	5.91	1.30	--	--
Stress Management	Treat	5.45	1.57	5.42	1.36	5.84	1.37	5.82	1.39
	WLC	5.49	1.38	5.51	1.31	5.65	1.39	--	--
Leadership Ability	Treat	5.90	1.41	5.91	1.37	6.03	1.43	6.02	1.45
	WLC	6.12	1.49	5.90	1.47	6.14	1.41	--	--
Coping with Change	Treat	5.66	1.38	5.59	1.27	6.03	1.33	6.05	1.30
	WLC	5.71	1.33	5.65	1.25	5.83	1.38	--	--
Overall	Treat	5.88		5.92		6.14		6.14	
	WLC	5.95		5.87		5.98		--	

Table 36 *(continued)*

Scale		T0 (*N* = 346)		T1 (*N* = 378)		T2 (*N* = 380)		T3 (*N* = 156)	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
External Locus of Control	Treat	5.00	1.68	5.38	1.37	5.67	1.38	4.91	1.62
	WLC	4.74	1.45	5.03	1.50	5.02	1.59	--	--
Internal Locus of Control	Treat	6.51	1.19	6.60	1.56	6.86	1.23	6.59	1.08
	WLC	6.46	1.19	6.52	1.27	6.59	1.18	--	--
Overall	Treat	5.76		5.99		6.27		5.75	
	WLC	5.60		5.78		5.81		--	
SDQ									
General Academic	Treat	4.51	0.88	4.38	0.86	4.40	0.92	4.57	0.88
	WLC	4.16	1.01	4.19	0.93	4.21	1.01	--	--
Global-Esteem	Treat	4.62	0.82	4.66	0.80	4.86	0.73	4.85	0.76
	WLC	4.54	0.86	4.57	0.87	4.59	0.82	--	--
Physical Ability	Treat	4.23	1.17	4.31	1.09	4.60	1.01	4.55	1.15
	WLC	4.46	1.13	4.32	1.15	4.35	1.15	--	--
Parent Relations	Treat	5.17	0.87	5.24	0.86	5.50	0.74	5.13	0.91
	WLC	4.99	1.08	5.05	1.05	4.98	1.08	--	--
Same-Sex Relations	Treat	4.66	0.94	4.73	0.82	4.87	0.84	4.58	0.96
	WLC	4.78	0.88	4.76	0.85	4.72	0.81	--	--
Opposite-Sex Relations	Treat	4.22	0.96	4.29	0.93	4.43	0.97	4.38	0.94
	WLC	4.26	0.98	4.40	0.97	4.31	0.94	--	--
Problem Solving	Treat	4.34	0.78	4.31	0.85	4.47	0.75	4.64	0.81
	WLC	4.33	0.83	4.28	0.89	4.42	0.82	--	--
Overall	Treat	4.54		4.56		4.73		4.67	
	WLC	4.50		4.51		4.51		--	

Table 36 *(continued)*

Scale		T0 (*N* = 346)		T1 (*N* = 378)		T2 (*N* = 380)		T3 (*N* = 156)	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
BPNSFS									
Autonomy	Treat	4.19	1.01	4.19	1.01	4.34	0.95	4.51	0.89
	WLC	4.30	0.89	4.40	0.94	4.34	0.90	--	--
Relatedness	Treat	4.42	0.92	4.42	0.92	4.68	0.81	4.60	0.79
	WLC	4.49	0.85	4.46	0.94	4.32	0.94	--	--
Competence	Treat	4.73	0.84	4.73	0.84	4.87	0.81	4.84	0.74
	WLC	4.67	0.84	4.65	0.90	4.65	0.85	--	--
Overall	Treat	4.45		4.45		4.63		4.65	
	WLC	4.48		4.50		4.44		--	
MES									
Thought	Treat	4.78	0.76	4.78	0.76	4.88	0.71	4.82	0.75
	WLC	4.82	0.79	4.85	0.75	4.77	0.85	--	--
Behaviour	Treat	4.31	0.97	4.31	0.97	4.36	1.01	4.48	0.91
	WLC	4.38	1.02	4.36	0.95	4.25	1.02	--	--
Hampering	Treat	3.59	0.99	3.59	0.99	3.19	0.87	3.46	1.05
	WLC	3.36	0.97	3.49	0.90	3.54	0.94	--	--
Overall	Treat	4.23		4.23		4.14		4.26	
	WLC	4.18		4.24		4.19		--	
GHQ									
Positive	Treat	2.08	0.42	2.13	0.32	2.35	0.38	2.20	0.90
	WLC	2.08	0.41	2.14	0.38	2.01	0.43	--	--
Negative	Treat	2.12	0.65	2.32	0.58	2.46	0.50	2.19	0.83
	WLC	2.13	0.63	2.28	0.57	2.13	0.66	--	--
Overall	Treat	2.10		2.23		2.41		2.20	
	WLC	2.11		2.21		2.07		--	

Table 36 *(continued)*

Scale		T0 (*N* = 346)		T1 (*N* = 378)		T2 (*N* = 380)		T3 (*N* = 156)	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Single Factor Scales									
Well Being	Treat	3.55	0.65	3.70	0.60	3.88	0.65	3.83	0.68
	WLC	3.63	0.62	3.62	0.66	3.56	0.75	--	--
Life Satisfaction	Treat	7.51	2.16	7.51	2.16	8.21	1.64	8.08	1.72
	WLC	7.54	1.93	7.65	1.87	7.40	2.12	--	--
Gratitude	Treat	4.90	0.75	4.87	0.77	5.20	0.75	4.78	0.70
	WLC	4.73	0.84	4.86	0.83	4.84	0.80	--	--
Resilience	Treat	4.30	0.95	4.30	0.95	4.52	0.92	4.62	0.86
	WLC	4.28	1.06	4.41	0.90	4.38	0.97	--	--

Hypothesis 2.1: Experimental effects of the OAE intervention.

Overview. Hypothesis 2.1 predicted that at the close of the OAE program (T2), students allocated to the intervention would report favourable gains in self-concept, life effectiveness, LoC, and secondary personal resource variables (see Chapter 4: General Method). Furthermore, the greatest benefits would be seen on the two GHQ-12 clinical scales. The randomized controlled trial design provided a strong WLC condition, against which treatment effects were compared.

Results. Descriptive statistics for T0, T1 and T2 are shown in Table 36 (see Appendix 6 for Skewness and Kurtosis statistics), and standardized ESs with 95% confidence intervals (CIs) are displayed in Table 37. To assess the main effects of the intervention, psychological outcome variables at the pre-test (T0 and T1) and post-test (T2) time-waves were regressed on a dummy variable that indicated the intervention and WLC conditions (WLC as a reference group).

Analysis of the seven SDQ factors revealed small effect sizes in favour of the treatment condition for Physical Ability (β = .222, *SE* = .067, $p < .01$) and Same-Sex Relations (β = .191, *SE* = .095, $p < .05$), while medium to strong effect sizes were observed for Global Esteem (β = .262, *SE* = .075, $p < .001$) and Parent Relations (β = .435, *SE* = .075, $p < .001$) domains. No statistically meaningful difference was observed between treatment and WLC conditions on the Opposite-Sex Relations control scale ($p > .05$). This was as expected, as no change in Opposite-Sex Relations self-concept would be expected to occur as a result from the OAE program.

Table 37 *Experimental ESs and Summary Descriptive for all Scales at T2 (Both Groups); T3 (Treatment) Versus T2 (WLC); and the Change in ESs Over the Post-Treatment Interval (T3 - T2)*

	T2 (*N* = 380)		T3 (*N* = 344)		T3 – T2	
Scale	*ES*	*SE*	*ES*	*SE*	*ES*	*SE*
SDQ						
Academic	.017	.075	.179*	.080	.161*	.062
Global Esteem	.262***	.075	.247**	.083	.014	.060
Physical Ability	.222**	.067	.235**	.072	.013	.054
Parent Relations	.435***	.075	.072	.085	-.363***	.064
Sm-Sex Rels	.191*	.095	-.290**	.104	-.482***	.095
Op-Sex Rels	.103	.071	.170*	.084	.067	.071
Problem Solving	-.015	.077	.240**	.084	.255***	.069
Mean	.185		.122		-.048	
ROPELOC						
Self Confidence	.334***	.075	.179*	.083	-.154**	.057
Time Efficiency	.186*	.083	.281**	.092	.095	.077
Active Involvement	.186**	.071	.037	.077	-.148**	.057
Coop Team	.117	.073	.125	.081	.008	.059
Open Thinking	.050	.089	.100	.094	.050	.073
Quality Seeking	.178*	.083	-.074	.094	-.252**	.074
Self-Efficacy	.087	.078	.191*	.083	.104	.071
Social Effectiveness	.092	.073	.068	.081	-.024	.064
Stress Management	.112	.077	.167	.086	.055	.064
Leadership Ability	.025	.067	.016	.075	-.008	.063
Coping with Change	.156	.082	.171	.095	.015	.076
Mean	.139		.115		-.024	
ELoC	.372***	.090	-.196	.103	-.568***	.090
ILoC	.198**	.072	-.078	.082	-.276***	.061
Mean	.285		-.137		-.422	
BPNSFS						
Autonomy	.103	.081	.266**	.088	.162*	.066
Relatedness	.384***	.088	.276**	.092	-.109	.057
Competence	.283***	.076	.207**	.079	-.076	.058
Mean	.257		.250		-.008	

Note. ES = standardised effect sizes. SDQ scales are: Academic = General Academic, Sm-Sex Rels = Same-Sex Relations, Op-Sex Rels = Opposite-Sex Relations, ROPELOC scales are: Coop Team = Cooperative Teamwork, ELoC = External Locus of Control, ILoC = Internal Locus of Control. GHQ-12 scales have been reverse-scored. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 37 *(continued)*

	T2 (*N* = 380)		T3 (*N*=344)		T2 - T3	
Scale	*ES*	*SE*	*ES*	*SE*	*ES*	*SE*
MES						
Thought	.263**	.076	.056	.082	-.206**	.060
Behaviour	.128	.082	.282**	.093	.154*	.074
Hampering	.317***	.083	-.187*	.094	-.504***	.078
Mean	.236		.050		-.185	
GHQ-12						
Positive	.705***	.090	.392***	.095	-.313***	.074
Negative	.533***	.085	.263**	.087	-.269***	.062
Mean	.619		.328		-.291	
One-Factor Scales						
Wellbeing	.439***	.081	.344***	.087	-.094	.067
Life Satisfaction	.323***	.080	.243**	.087	-.079	.060
Gratitude	.349***	.084	-.031	.095	-.380***	.073
Resilience	.150	.079	.293**	.087	.143	.061

Note. ES = standardised effect sizes. * $p < .05$, ** $p < .01$, *** $p < .001$.

From the 13 ROPELOC factors, small treatment effects were observed for Quality Seeking ($\beta = .178$, $SE = .083$, $p < .05$), Active Involvement ($\beta = .186$, $SE = .071$, $p < .01$), Time Efficiency ($\beta = .186$, $SE = .083$, $p < .05$), and Internal LoC ($\beta = .198$, $SE = .072$, $p < .01$). Medium effect sizes occurred in areas of External LoC ($\beta = .372$, $SE = .090$, $p < .001$) and Self Confidence ($\beta = .334$, $SE = .075$, $p < .001$), again favouring the treatment condition.

With respect to SDT psychological needs, the OAE program was associated with medium effects for Competence ($\beta = .283$, $SE = .076$, $p < .001$) and Relatedness ($\beta = .384$, $SE = .088$, $p < .001$).

Secondary research outcomes analysis revealed medium treatment gains in the MES Thought ($\beta = .263$, $SE = .076$, $p < .01$) and Hampering ($\beta = .317$, $SE = .083$, $p < .001$) scales. Each indicator of psychological distress measured by the GHQ-12 evidenced strong effect sizes (Positive, $\beta = .705$, $SE = .090$, $p < .001$; Negative, $\beta = .533$, $SE = .085$, p

< .001), while single-factor scales showed effect sizes ranging from medium to strong for Life Satisfaction (β = .323, *SE* = .080, $p < .001$), Wellbeing (β = .439, *SE* = .081, $p < .001$) and Gratitude (β = .349, *SE* = .084, $p < .001$).

Conclusion. Hypothesis 2.1 predicted that the OAE intervention would result in benefits to self-concept, life effectiveness and LoC, and that the greatest gains would be observed on the two GHQ-12 clinical scales. Results from the analysis were partially consistent with this prediction, whereby 19 from 32 outcomes evidenced statistically significant gains in favour of OAE participants, effect sizes ranging from small (.178) to strong (.705). Additionally, the GHQ-12 average effect size (β = .619) vastly exceeded that seen for self-concept (β = .185), Life Effectiveness (β = .139), and LoC (β = .285) scales.

Hypothesis 2.2: Lasting effects of the intervention (T2 to T3).

Overview. Hypothesis 2.2 posited that the post-test experimental effects of the OAE program would be maintained at six-month follow-up, and that further gains in effects would be observed over the post-treatment interval.

Results. Descriptive statistics are shown in Table 36, while standardized ESs with 95% confidence intervals (CIs) are displayed in Table 37. To examine follow-up effects from the intervention, psychological outcomes at follow-up (T3) were regressed on a dummy variable that indicated the intervention and WLC conditions, controlling for the corresponding variables at the pre-test waves (T0 and T1). However, the WLC was no longer viable to act as a control comparison at T3, so an innovative approach was taken to address this issue. The regression analysis treatment-group factor scores for each outcome at T3 were matched to the WLC post-test scores at T2. As such, the analysis used the WLC post-test scores at T2 as a reference to test follow-up effects. A series of post-hoc comparisons were run between T2 and T3 effect sizes for all outcomes, to detect the presence of diminishing, maintenance or sleeper effects occurring over the post-treatment interval.

Analysis of the seven SDQ factors revealed medium follow-up effect sizes for Global Esteem (β = .247, *SE* = .083, $p < .01$) and Physical Ability (β = .235, *SE* = .072, $p < .01$) self-concept, while post hoc analysis indicated that experimental effects were maintained on these factors. Non-significant post-treatment effects (T2) became significant at follow-up (T3) for General Academic (β = .179, *SE* = .080, $p < .05$), Problem Solving (β = .240, *SE* = .084, $p < .01$) and Opposite-Sex Relations (β = .170, *SE* = .084, $p < .05$). Post hoc analysis revealed statistically significant increases in the General Academic effect size (β = .161, *SE* = .062, $p < .05$) over the post-treatment interval, reflecting the emergence of a sleeper effect on this variable (see Chapter 2: Literature Review). A non-significant effect for Parent Relations ($p > .05$) occurred due to a significant diminution of post-treatment effects on this factor (β = -.363, *SE* = .064, $p < .001$), whereas the follow-up analysis observed a medium effect size in a negative direction for Same-Sex Relations (β = -290, *SE* = .104, $p < .01$) self-concept. This similarly indicates that the positive treatment effects seen on this factor at T2 had diminished. Figure 15 shows the overall mean change in self-concept for the treatment and WLC conditions over the four time-waves.

For the ROPELOC scales, small and medium follow-up treatment effects were observed for Self Confidence (β = .179, *SE* = .083, $p < .05$) and Time Efficiency (β = .281, *SE* = .092, $p < .01$), respectfully, and maintenance of post-treatment effects was determined. Non-significant effects were seen at T3 for Active Involvement, Quality Seeking and Internal LoC scales ($p > .05$), indicating a diminishing of post-treatment effects. A medium follow-up effect size emerged at T3 for Self-Efficacy (β = .197, *SE* = .083, $p < .05$), however this increase in effect size from post-test was non-significant ($p > .05$). Figure 16 and Figure 17 shows the overall mean change in life effectiveness and LoC, respectively, for the treatment and WLC conditions over the four time-waves.

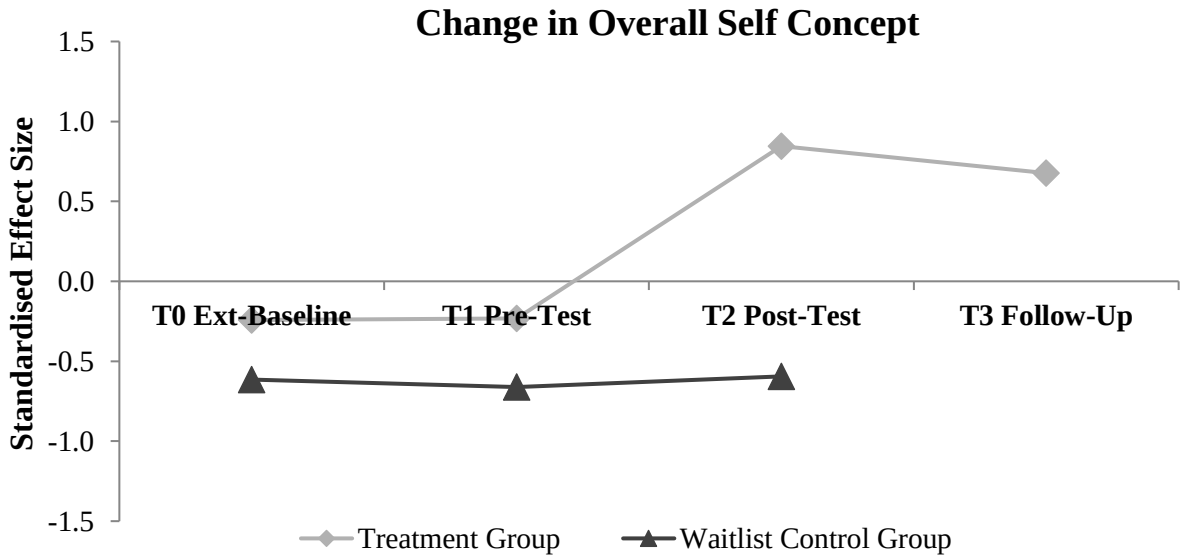


Figure 15. Change in overall mean self-concept for the treatment group and WLC over the four time-waves. *Note.* WLC data at T3 was derived from the WLC T2 scores. Scale Ms were calculated from SDQ factors: Global Esteem, Physical Ability, General Academic, Parent Relations, Same-Sex Relations, and Problem Solving.

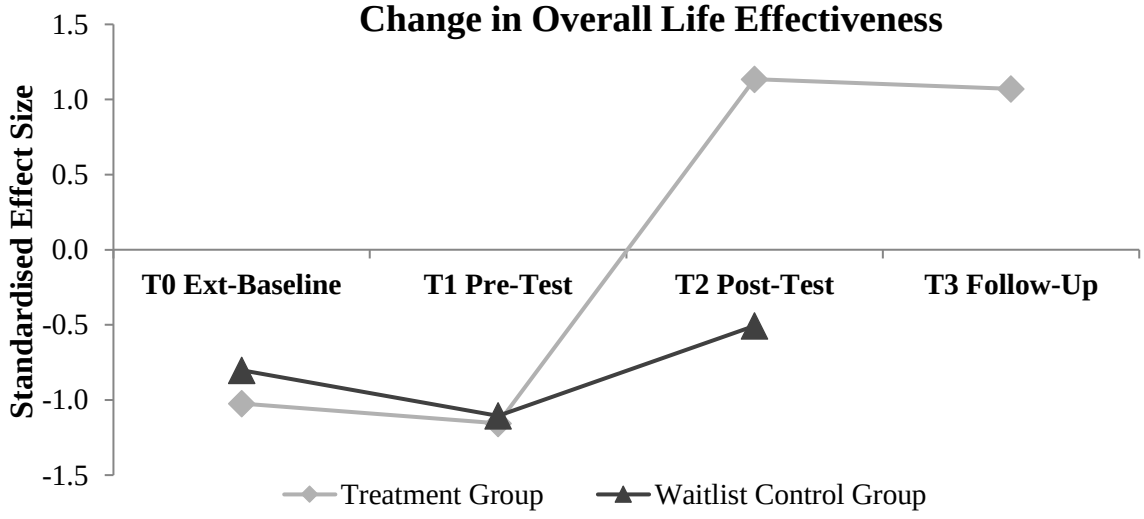


Figure 16. Change in overall life effectiveness for the treatment group and WLC over the four time-waves. *Note.* WLC data at T3 was derived from the WLC T2 scores. Scale *M*s were calculated from ROPELOC factors: Self Confidence, Time Efficiency, Active Involvement, Cooperative Teamwork, Open Thinking, Quality Seeking, Self Efficacy, Social Effectiveness, Stress Management, Leadership Ability, and Coping with Change.

For the BPNSFS scale, statistically significant medium effect sizes were seen at follow-up for Autonomy (β = .266, *SE* = .088, $p < .01$), Relatedness (β = .276, *SE* = .092, $p < .01$) and Competence (β = .207, *SE* = .079, $p < .01$).

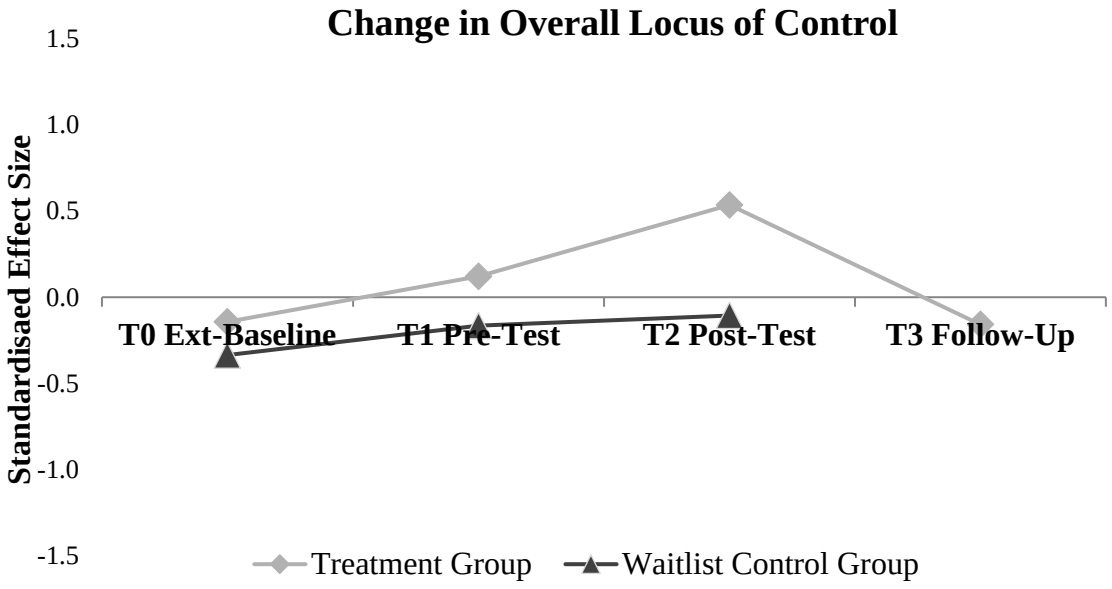


Figure 17. Change in overall locus of control for the treatment group and WLC over the four time-waves. *Note.* WLC data at T3 was derived from the WLC T2 scores. Scale *M*s were calculated from ROPELOC factors: External Locus of Control (reverse-scored) and Internal Locus of Control.

Post-hoc analysis revealed no statistically meaningful changes in effect sizes for Relatedness or Competence over the post-treatment interval ($p > .05$), and thus treatment gains were maintained on these factors. Contrasting to this was the statistically significant increase in effect size seen for Autonomy from T2 to T3 (β = .162, *SE* = .066, $p < .05$), indicating the emergence of a sleeper effect. The overall mean change in basic psychological needs satisfaction for both treatment and WLC conditions has been displayed in Appendix 7.

The longitudinal analysis for secondary outcomes revealed statistically significant medium effect sizes for Wellbeing (β = .344, *SE* = .087, $p < .001$), Life Satisfaction (β =

.243, *SE* = .087, $p < .01$) and Resilience (β = .293, *SE* = .087, $p < .01$). Post hoc analysis indicated no significant changes in effect sizes from T2 to T3. A non-significant finding for Gratitude at T3 was reflective of a significant diminishing of post-treatment gains on this factor (β = -.380, *SE* = .073, $p < .001$). Medium effect sizes were observed on the GHQ-12 Positive (β = .392, *SE* = .095, $p < .001$) and Negative (β = .263, *SE* = .087, $p <$ .01) factors at follow-up, however post hoc analysis revealed these to be significantly reduced compared to post-test (Positive, β = -.313, *SE* = .074, $p < .001$; Negative, β = -.269, *SE* = .062, $p < .001$). As such, although treatment participants reported significantly less psychological distress at T3 compared to the WLC reference group, the benefits were greatest at T2. Follow-up analysis of the MES Thought and Hampering scales found the T2 post-treatment effects to have diminished at T3 ($p > .05$). However, a statistically significant medium effect size was observed for MES Behaviour (β = .282, *SE* = .093, $p <$ .01) at follow-up. This was statistically greater than at post-test (β = .154, *SE* = .074, $p <$ .05), reflecting a favourable sleeper effect for behavioural motivational tendencies towards school work.

Conclusion. Hypothesis 2.2 predicted that post-treatment experimental effects would be maintained and that additional 'sleeper' effects would be observed. Maintenance of post-treatment effects were seen for Global Esteem and Physical Ability self-concept, Self Confidence and Time Efficiency ROPELOC scales, both Positive and Negative symptom GHQ-12 scales, Wellbeing, Life Satisfaction, and psychological needs satisfaction for Relatedness and Competence. Non-significant post-test effects became significant at follow-up for Opposite-Sex Relations self-concept, Resilience and Self Efficacy, while statistically significant increases in effect sizes over the post-treatment interval were reflective of sleeper effects for General Academic and Problem Solving self-concept, Autonomy needs satisfaction, and Motivation and Engagement Behaviour. The

procedures used to test hypothesis 2.2 demonstrate the potential for OAE programs to foster long-term psychosocial benefits during adolescence.

Hypothesis 2.3 - Aptitude-treatment interaction (T1 to T2).

Overview. Hypothesis 2.3 predicted that the observed gains in each outcome variable that were attributable to the intervention over the experimental interval would be equivalent for all participants, irrespective of baseline aptitude. That is, the benefits from the intervention would be equivalent for all participants and thus, no aptitude-treatment interactions would be observed.

Results. To assess whether the intervention effects varied by different levels of baseline aptitude, an additional interaction term was specified (e.g., intervention X respective baseline outcome) to the regression models. The results are presented in Table 38. Consistently with hypothesis 2.3 predictions, apart from two exceptions. There were no other statistically significant interactions for baseline aptitude ($p > .05$) over the experimental interval, the case for 30 of the 32 outcomes. The exceptions included ROPELOC scales, Time Efficiency ($\beta = -.227$, $SE = .107$, $p < .05$), where treatment gains favoured low baseline aptitude, and Quality Seeking ($\beta = .201$, $SE = .101$, $p < .05$) where, interestingly, the greatest gains were seen among high baseline aptitude students.

Conclusion. Hypothesis 2.3 predicted that students' baseline aptitude on each research outcome would not determine the gains experienced from the OAE program, and thus that the intervention would benefit all students comparably. The findings were largely in favour of this prediction, whereby 30 from 32 outcomes reflected no preference for baseline aptitude on the effectiveness of the OAE program. Hypothesis 2.3 was therefore accepted and the OAE program was determined to be providing a robust intervention which provides equal benefits to students presenting with differing levels of baseline functioning.

Table 38 *ESs for the Interaction Between Intervention and Baseline Aptitude for all Scales at T1 to T2*

T1-T2 Aptitude Interactions (N = 380)					
Scale	β	*SE*	Scale	β	*SE*
SDQ			ROPELOC		
General Academic	-.106	.090	SC	-.040	.093
Global Esteem	-.110	.096	TE	-.227*	.107
Physical Ability	-.154	.080	AI	-.028	.080
Parent Relations	-.051	.101	CT	-.115	.090
Sm-Sex Relations	.219	.130	OT	.117	.102
Op-Sex Relations	.128	.092	QS	.201*	.101
Problem Solving	-.018	.126	SF	-.124	.083
Mean	-.037		SE	.020	.087
MES			SM	.044	.098
Thought	.068	.111	LA	-.083	.077
Behaviour	-.113	.123	CH	.071	.106
Hampering	.017	.101	Mean	-.015	
Mean	-.009		ELoC	-.106	.103
GHQ-12			ILoC	.073	.098
Positive	.288	.172	Mean	-.017	
Negative	.213	.130	BPNSFS		
Mean	.251		Autonomy	.100	.106
1-Fac. Scales			Relatedness	-.027	.121
Wellbeing	-.152	.113	Competence	-.030	.097
Life Satisfaction	-.102	.110	Mean	.014	
Gratitude	.128	.120			
Resilience	.136	.128			

Note. SDQ factors are: Sm-Sex Relations = Same-Sex Relations, Op-Sex Relations = Opposite-Sex Relations. ROPELOC factors are: SC = Self Confidence, TE = Time Efficiency, AI = Active Involvement, CT = Cooperative Teamwork, OT = Open Thinking, QS = Quality Seeking, SF = Self Efficacy, SE = Social Effectiveness, SM = Stress Management, LA = Leadership Ability, CH = Coping with Change, ELoC = External Locus of Control, and ILoC = Internal Locus of Control. * $p < .05$, ** $p < .01$, *** $p < .001$.

Hypothesis 2.4 - Mediation of the experimental effects

Overview. Hypothesis 2.4 predicted that the treatment condition would show increases in basic psychological needs satisfaction, which would constitute the mediating mechanisms through which the OAE intervention benefits research outcomes. To test this, constructs were measured post-treatment (T2) so to evaluate intervention students'

experience of psychological needs satisfaction due to the OAE program when compared to WLC participants. Preliminary analysis examined the correlations between each outcome at T2 with students' post-test (T2) psychological needs satisfaction. Mediation analysis was conducted in accordance with procedures detailed in the Study 2 Method section, pertaining to those outcomes evidencing significant total effects following hypothesis 2.1 testing.

Results. Table 39 shows the intercorrelation matrix for the three SDT psychological needs and the 29 outcome variables at T2. As expected, satisfaction of each psychological need (Autonomy, Relatedness and Competence) correlated significantly and in the expected direction with all outcomes: a total of 87 comparisons.

The results from the mediation analysis are presented in Table 40. As shown in Figure 18, from four total effects observed on the seven-factor SDQ instrument at T2 (see Table 37), three were primarily mediated by basic psychological needs satisfaction; the other was partially mediated. Significant indirect effects were seen for treatment on Same-Sex Relations (β = .100, *SE* = .039, $p < .05$; 95% CI [.036, .189]) and Parent Relations (β = .059, *SE* = .027, $p < .05$; 95% CI [.013, .124]), as mediated by the psychological need for Relatedness. Indirect effects for treatment via Competence needs satisfaction were observed for Global Esteem (β = .182, *SE* = .052, $p < .001$; 95% CI [.086, .291]) and Physical Ability (β = .086, *SE* = .033, $p < .01$; 95% CI [.028, .162]) self-concept.

The direct relationships between treatment with Global Esteem (β = .089, p = .107), Same-Sex Relations (β = .037, p = .641) and Physical Ability self-concept (β = .122, p = .087) became non-significant when controlling for psychological needs satisfaction, while the direct path from treatment condition to Parent Relations self-concept (β = .347, *SE* = .067, $p < .001$) remained significant.

Table 39 *Intercorrelation Matrix for SDT Needs Satisfaction and Each Dependent Measure Subscale at T2*

	Correlation (*r*)				Correlation (*r*)		
Scale	Auton	Rel	Comp	Scale	Auton	Rel	Comp
SDQ				ROPELOC			
Academic	*.47*	*.39*	*.68*	SC	*.63*	*.59*	*.81*
Global Esteem	*.66*	*.62*	*.82*	TE	*.57*	*.45*	*.60*
Physical	*.33*	*.41*	*.44*	AI	*.53*	*.57*	*.68*
Parent Relations	*.50*	*.51*	*.55*	CT	*.50*	*.53*	*.51*
Sm Sex	*.19*	*.31*	*.25*	OT	*.48*	*.43*	*.56*
Op Sex	*.25*	*.35*	*.29*	QS	*.49*	*.40*	*.67*
Problem Solving	*.51*	*.44*	*.60*	SF	*.58*	*.48*	*.69*
				SE	*.52*	*.56*	*.54*
MES				SM	*.45*	*.42*	*.51*
Thought	*.73*	*.59*	*.86*	LA	*.46*	*.44*	*.57*
Behaviour	*.64*	*.49*	*.59*	CH	*.53*	*.42*	*.57*
Hampering	*.41*	*.36*	*.58*	ELoC	*.29*	*.32*	*.41*
				ILoC	*.62*	*.51*	*.73*
1-Factor Scales							
Wellbeing	*.66*	*.65*	*.74*	GHQ-12			
Life Satisfaction	*.52*	*.56*	*.52*	Positive	*.59*	*.61*	*.64*
Gratitude	*.25*	*.35*	*.33*	Negative	*.51*	*.57*	*.59*
Resilience	*.61*	*.53*	*.66*				

Note. N = 376. BPNSFS scales are: Auton = Autonomy, Rel = Relatedness, and Comp = Competence. SDQ scales are: Academic = General Academic, Physical = Physical Ability, Sm Sex = Same-Sex Relations, and Op Sex = Opposite-Sex Relations. ROPELOC factors are: SC = Self Confidence, TE = Time Efficiency, AI = Active Involvement, CT = Cooperative Teamwork, OT = Open Thinking, QS = Quality Seeking, SF = Self Efficacy, SE = Social Effectiveness, SM = Stress Management, LA = Leadership Ability, CH = Coping with Change, ELoC = External Locus of Control, and ILoC = Internal Locus of Control. Significant results have been italicised.

Table 40 *ESs, SEs and CIs for Each Component of the SDT Mediation Model and the SDQ, ROPELOC, MES, GHQ-12, and One-Factor Scales*

	Direct Effect		Indirect Autonomy			Indirect Relatedness			Indirect Competence			Total Effect		
Scale	*β*	*SE*	*β*	*SE*	95% CI [LL, UL]	*β*	*SE*	95% CI [LL, UL]	*β*	*SE*	95% CI [LL, UL]	*β*	*SE*	95% CI [LL, UL]
SDQ														
Prnt	.347***	.067	.009	.010	[-.011, .033]	.059*	.027	[.013, .124]	.043	.028	[-.003, .114]	.458***	.077	[.308, .611]
Esteem	.089	.055	.000	.004	[-.013, .017]	.050	.026	[.004, .112]	.182***	.052	[.086, .291]	.321***	.075	[.171, .468]
Phys	.122	.072	.001	.006	[-.019, .018]	.009	.033	[-.050, .078]	.086*	.033	[.028, .162]	.219**	.066	[.093, .348]
Acad	-.085	.064	-.004	.006	[-.024, .010]	-.024	.020	[-.070, .016]	.226***	.055	[.122, .337]	.112	.078	[-.044, .268]
Prob	-.137*	.066	.010	.011	[-.010, .040]	.031	.026	[-.016, .090]	.107**	.036	[.043, .185]	.011	.075	[-.140, .159]
Sm Sx	.037	.080	-.009	.012	[-.044, .012]	.100*	.039	[.036, .189]	.023	.028	[-.033, .083]	.151*	.076	[.001, .303]
Op Sx	-.029	.066	-.007	.008	[-.033, .007]	.101***	.026	[.053, .155]	.022	.021	[-.020, .068]	.087	.065	[-.043, .211]
1-Factor														
WB	.296***	.067	.012	.014	[-.001, .049]	.091**	.031	[.038, .164]	.106**	.039	[.039, .192]	.505***	.081	[.344, .662]
LSat	.409**	.130	.011	.016	[-.017, .062]	.240**	.076	[.102, .397]	.059	.051	[-.032, .177]	.718***	.152	[.423, 1.021]
Grat	.199*	.082	-.006	.010	[-.042, .012]	.072*	.030	[.019, .141]	.058	.038	[-.008, .147]	.323***	.080	[.161, .482]
Res	.016	.069	.013	.014	[-.013, .049]	.032	.029	[-.022, .099]	.115**	.042	[.042, .206]	.176*	.078	[.018, .328]
MES														
Thou	.085	.046	.013	.015	[-.012, .049]	-.009	.018	[-.042, .031]	.202***	.049	[.106, .302]	.290***	.075	[.138, .441]
Behav	-.004	.065	.031	.032	[-.028, .103]	.006	.021	[-.039, .052]	.082**	.029	[.027, .147]	.115	.083	[-.051, .277]
Hamp	.212**	.064	-.002	.005	[-.020, .013]	-.020	.025	[.071, .036]	.122**	.037	[.057, .203]	.312***	.070	[.174, .454]
GHQ-12														
Pos	.144***	.020	.004	.005	[-.004, .016]	.029**	.010	[.012, .052]	.024*	.010	[.006, .047]	.714***	.090	[.537, . 890]
Neg	.188***	.038	.003	.004	[-.006, .017]	.049**	.017	[.019, .088]	.048*	.019	[.014, .093]	.545***	.084	[.381, . 709]

Note. *β* = standardised effect sizes. SDQ scales are: Acad = General Academic, Phys = Physical Ability, Sm Sx = Same-Sex Relations, Op Sx = Opposite-Sex Relations, Prnt = Parent Relations, Esteem = Global Esteem, and Prob = Problem Solving. One-Factor scales are: WB = Wellbeing, LSat = Life Satisfaction, Grat = Gratitude, and Res = Resilience. MES Scales are: Thou = Thought, Behav = Behaviour, and Ham = Hampering. GHQ-12 scales are: Pos = Positive and Neg = Negative. Opposite-Sex Relations scale (shaded 'grey') = Control scale. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 40 *(continued)*

Scale	Direct Effect		Indirect Autonomy			Indirect Relatedness			Indirect Competence			Total Effect		
	β	*SE*	*β*	*SE*	95% CI [LL, UL]	*β*	*SE*	95% CI [LL, UL]	*β*	*SE*	95% CI [LL, UL]	*β*	*SE*	95% CI [LL, UL]
ROPELOC														
SC	.184***	.052	-.003	.005	[-.020, .011]	.036	.022	[-.004, .085]	.186***	.052	[.089, .295]	.403***	.072	[.258, .543]
TE	.081	.074	.019	.022	[-.018, .077]	.008	.027	[-.055, .058]	.094**	.034	[.031, .017]	.202*	.082	[.037, .362]
AI	.008	.058	.001	.005	[-.019, .013]	.073**	.026	[.028, .136]	.099**	.033	[.042, .174]	.180**	.065	[.048, .304]
CT	-.019	.068	.006	.008	[-.008, .03]	.088**	.033	[.032, .167]	.035	.026	[-.013,.092]	.110	.070	[-.033, .248]
OT	-.083	.070	.002	.007	[-.013, .027]	.055*	.026	[.007, .114]	.090*	.035	[.026, .165]	.065	.076	[-.082, .213]
QS	.009	.060	-.001	.005	[-.018, .017]	-.020	.023	[-.068, .028]	.167***	.046	[.082, .270]	.155*	.071	[.016, .296]
SF	-.025	.067	.007	.009	[-.009, .034]	.010	.024	[-.033, .064]	.125**	.040	[.055, .213]	.117	.074	[-.031, .260]
SE	-.066	.068	.004	.007	[-.010, .024]	.111**	.037	[.048, .193]	.037	.024	[-.010, .088]	.087	.069	[-.051, .217]
SM	-.005	.068	.003	.006	[-.011, .023]	.061*	.028	[.010, .127]	.058	.031	[.002, .127]	.118	.072	[-.030, .260]
LA	-.079	.065	.001	.005	[-.014, .020]	.044	.026	[-.001, .102]	.060*	.026	[.014, .116]	.026	.064	[-.102, .150]
CH	.085	.078	.017	.019	[-.015, .066]	.009	.028	[-.041, .071]	.070*	.036	[.009, .147]	.181*	.079	[.018, .337]
EL	.258**	.076	-.003	.006	[-.025, .015]	.027	.031	[-.025, .096]	.053	.027	[.004, .117]	.335***	.075	[.183, .483]
IL	.073	.058	.012	.014	[-.012, .045]	.016	.020	[-.021, .061]	.107**	.034	[.049, .183]	.209**	.069	[.068, .346]

Note. β = standardised effect sizes. ROPELOC factors are: SC = Self Confidence, TE = Time Efficiency, AI = Active Involvement, CT = Cooperative Teamwork, OT = Open Thinking, QS = Quality Seeking, SF = Self Efficacy, SE = Social Effectiveness, SM = Stress Management, LA = Leadership Ability, CH = Coping with Change, EL = External Locus of Control, and IL = Internal Locus of Control. * $p < .05$, ** $p < .01$, *** $p < .001$.

These results indicate that the effects of the OAE program on students' Global Esteem, Same-Sex Relations and Physical Ability self-concept were primarily mediated by students' basic psychological needs satisfaction, while the Parent Relations domain was only partially mediated (see Figure 18). Of interest is the alignment between the mediating role of Relatedness with socially grounded self-concept domains, while those domains that focus on self-perceptions of capability and efficacy were predominantly mediated by the psychological need for Competence.

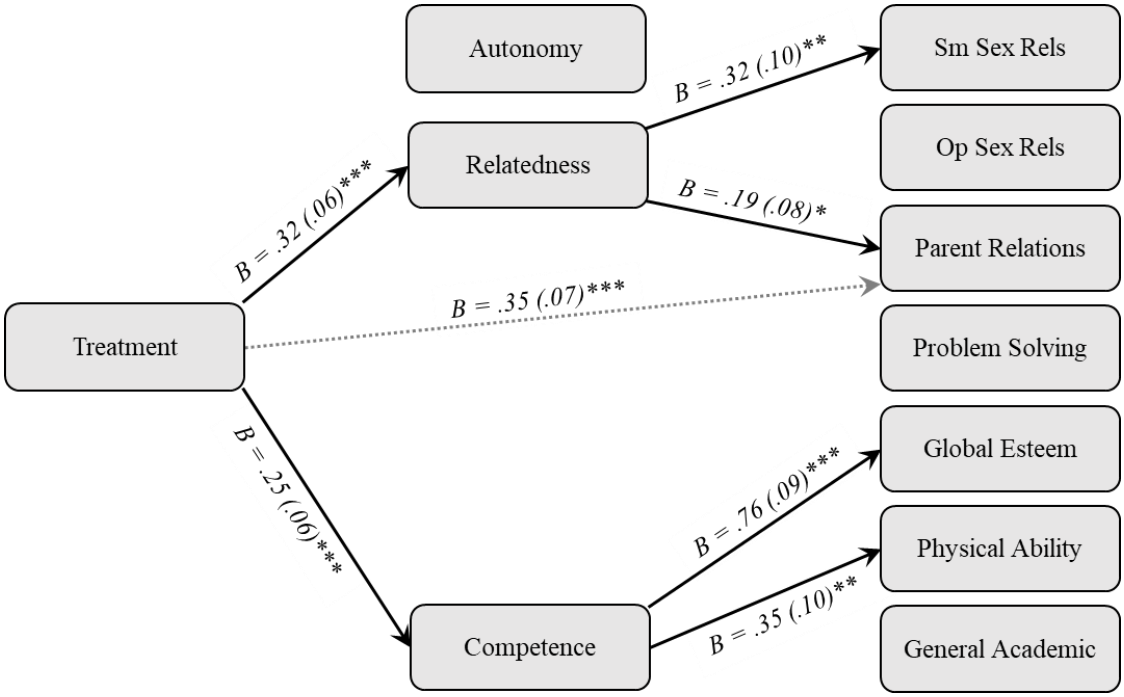


Figure 18. Direct and indirect effects of the treatment condition on SDQ outcomes when modelling SDT needs satisfaction measures as mediators.

Note. β = standardised effect sizes. SDT and SDQ factors were measured at T2 (post-treatment). SDQ scales are: Sm Sx Rels = Same-Sex Relations, Op Sx Rels = Opposite-Sex Relations. The analysis controlled for baseline (T1) needs satisfaction and dependent variable (T0-T1) group differences. * $p < .05$, ** $p < .01$, *** $p < .001$.

For the 13-factor ROPELOC, the inclusion of the three SDT mediators revealed significant indirect effects for Competence between treatment and Time Efficiency *(β* =

.094, *SE* = .034, $p < .01$; 95% CI [.031, .017]), Quality Seeking *(β* = .167, *SE* = .046, $p < .001$; 95% CI [.082, .270]), Coping with Change *(β* = .070, *SE* = .036, $p < .05$; 95% CI [.009, .147]), Self Confidence *(β* = .186, *SE* = .052, $p < .001$; 95% CI [.089, .295]) and Internal LoC *(β* = .107, *SE* = .034, $p < .01$; 95% CI [.049, .183]). The Active Involvement factor revealed significant indirect paths through both Competence *(β* = .099, *SE* = .033, $p < .01$; 95% CI [.042, .174]) and Relatedness *(β* = .073, *SE* = .026, $p < .01$; 95% CI [.028, .136]) needs satisfaction. No mediating effects were observed for treatment on External LoC via psychological needs satisfaction ($p > .05$).

When psychological needs were introduced to the mediation model, non-significant direct relationships were observed between treatment and Time Efficiency (*β* = .081, $p = .271$), Active Involvement (*β*= .008, $p = .887$), Quality Seeking (*β* = .009, $p = .886$), Coping with Change (*β* = .085, $p = .276$), and Internal LoC (*β* = .073, $p = .206$). The direct paths from treatment condition to Self Confidence (*β* = .184, *SE* = .052, $p < .001$) and External LoC (*β* = .258, *SE* = .076, $p < .01$) remained significant when modelling psychological needs satisfaction as covariates. These results indicate five of the 13 personal effectiveness factors, including Internal LoC, were primarily mediated, whereas one single factor was partially mediated, by satisfaction of students' basic psychological needs (see Figure 19).

Mediation analysis was next conducted for the MES and GHQ-12. Significant indirect relationships via Competence need satisfaction were observed between treatment and Thought (*β* = .202, *SE* = .049, $p < .001$; 95% CI [.106, .302]) and Hampering (*β* = .122, *SE* = .037, $p < .01$; 95% CI [.057, .203]) MES scales. Indirect mediation effects were again seen for the two GHQ-12 factors; this time however, via both Competence (Positive: *β* = .024, *SE* = .010, $p < .05$; 95% CI [.006, .047]; Negative: *β* = .048, *SE* = .019, $p < .05$; 95% CI [.014, .093]) and Relatedness (Positive: *β* = .029, *SE* = .010, $p < .01$; 95% CI

[.012, .052]; Negative: β = .049, *SE* = .017, p < .01; 95% CI [.019, .088]) needs satisfaction.

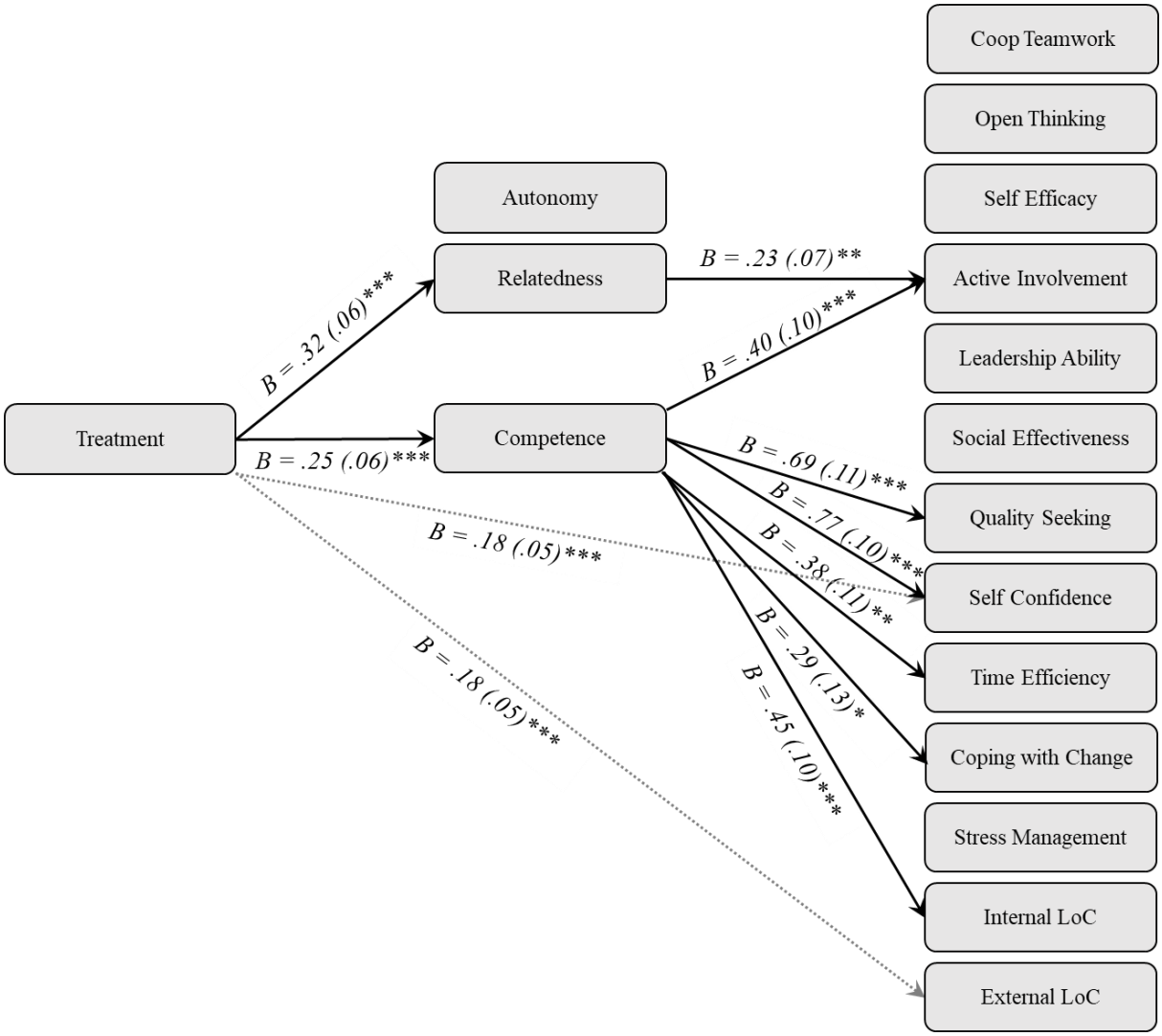


Figure 19. Direct and indirect effects of treatment condition on ROPELOC scale outcomes when modelling SDT needs satisfaction as mediators.

Note. β = standardised effect sizes. SDT and ROPELOC factors were measured at T2 (post-treatment). ROPELOC scales are: Coop Teamwork = Cooperative Teamwork, Internal LoC = Internal Locus of Control, External LoC = External Locus of Control. The analysis controlled for baseline (T1) needs satisfaction and dependent variable (T0-T1) group differences. * p < .05, ** p < .01, *** p < .001.

When controlling for basic psychological needs, the direct relationship between treatment and Thought became non-significant (β = .085, p = .069). The direct paths from treatment to MES Hampering (β = .212, SE = .064, p < .01) and GHQ-12 Positive (β = .144, SE = .020, p < .001) and Negative (β = .188, SE = .038, p < .001) scales remained significant when psychological needs were included as mediators. Consequently, adaptive motivational thought tendencies were primarily mediated by basic psychological needs, while maladaptive cognitive-affective motivational tendencies and indicators of psychological distress were partially mediated by SDT needs satisfaction (see Figure 20).

Mediation analysis lastly saw basic psychological needs modelled as covariates within each of the single-factor models. Results found a significant indirect effect of treatment on Resilience via Competence (β = .115, SE = .042, p < .01; 95% CI [.042, .206]). Indirect effects via Relatedness were further observed between treatment with Life Satisfaction (β = .240, SE = .076, p < .01; 95% CI [.102, .396]) and Gratitude (β = .072, SE = .030, p < .05; 95% CI [.019, .141]), while both Relatedness (β = .091, SE = .031, p < .01; 95% CI [.038, .164]) and Competence (β = .106, SE = .039, p < .01; 95% CI [.039, .192]) showed significant indirect effects for treatment on Wellbeing. When modelling psychological needs satisfaction as covariates, the direct path from treatment to Resilience became non-significant (β = .016, p = .816). In contrast, the paths from treatment to Wellbeing (β = .296, SE = .067, p < .001), Life Satisfaction (β = .409, SE = .130, p < .01), and Gratitude (β = .199, SE = .082, p < .05) were significant after controlling for psychological needs satisfaction. This analysis revealed psychological need satisfaction as primarily mediating the effects of treatment on Resilience, while the effects of treatment on Wellbeing, Life Satisfaction and Gratitude showed partial mediation effects (see Figure 20).

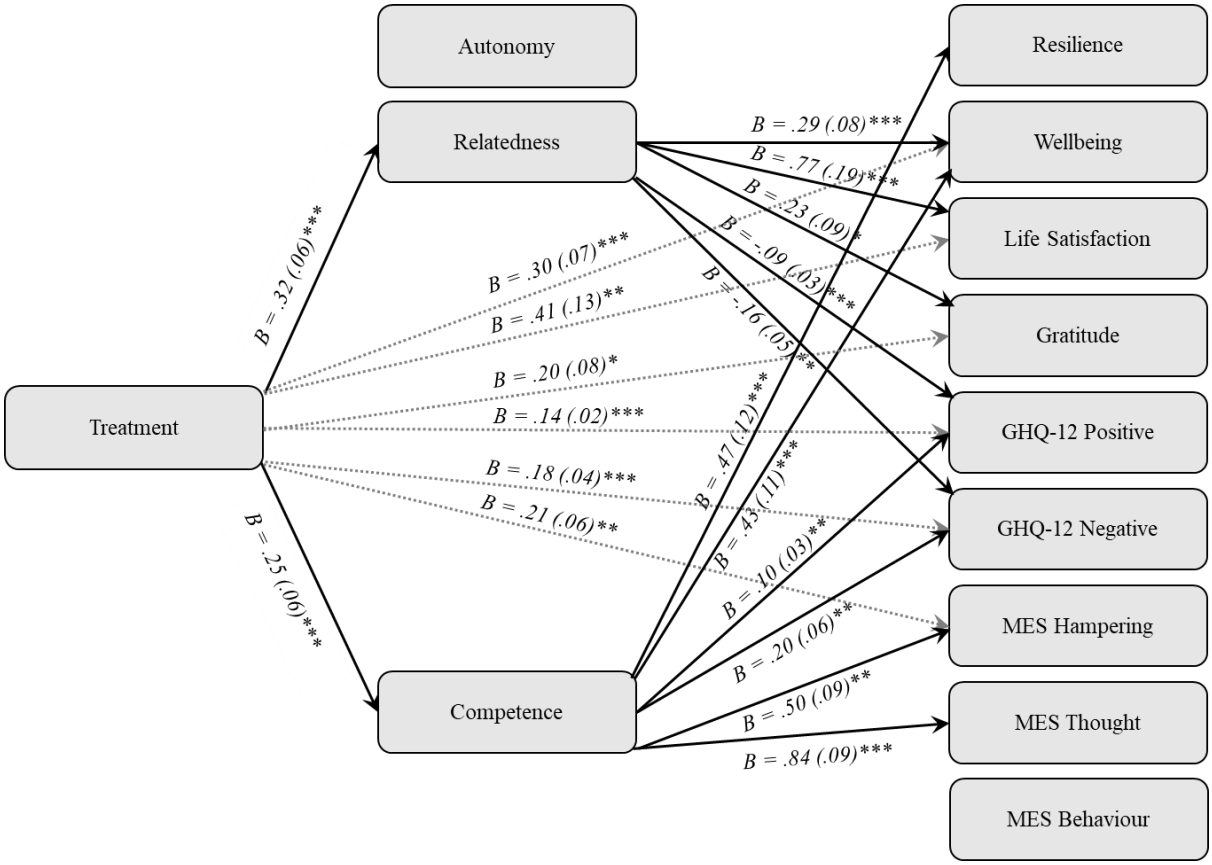


Figure 20. Direct and indirect effects of treatment condition on secondary outcomes when modelling SDT needs satisfaction as mediators.

Note. β = standardised effect sizes. SDT needs and dependent variables were measures at T2 (post-treatment). The analysis controlled for baseline (T1) needs satisfaction and dependent variable (T0-T1) group differences.* $p < .05$, ** $p < .01$, *** $p < .001$.

Conclusion. Hypothesis 2.4 predicted that the intervention would foster increases in SDT basic psychological needs satisfaction and that these gains would mediate the OAE intervention effects on research outcomes. This prediction was partially accepted, whereby the program led to greater Relatedness and Competence needs satisfaction compared to WLC participants. Mediation analysis revealed that 10 from 19 total effects were primarily mediated by one or both of these factors, while a further 8 from 19 total effects were partially mediated.

Summary of Results

Study 2 set out to empirically evaluate the effectiveness of a 5-month OAE program on male adolescents' psychological development. At the end of the OAE program, treatment participants showed benefits in multiple facets of self-concept and life effectiveness, LoC, and personal resources, including school motivation, life satisfaction, subjective wellbeing and general resilience. Longitudinal analysis saw maintenance of experimental effects on 11 from 18 factors, while a further seven domains of psychosocial and emotional functioning showed beneficial effects in favour of the treatment group at follow-up, which were not otherwise apparent at the conclusion of the OAE program. The immediate effects of the OAE program were largely mediated by SDT basic psychological needs satisfaction for Competence and Relatedness, although contrary to *a priori* predictions, Autonomy showed no mediation pathways. Examination of students' baseline aptitude on treatment effectiveness revealed no systematic trend and thus, the OAE program was determined to comprise a robust intervention that offers equivalent benefits irrespective of participants' baseline functioning.

These findings, arising from a strong research design, demonstrate the exciting potential for OAE programs to facilitate healthy psychological development in teenagers with benefits that last. The following chapter provides a detailed discussion of the Study 2 findings in the context of the broader literature area, along with the research strengths and limitations, contributions to the field and implication for future research.

CHAPTER 7: DISCUSSION AND CONCLUSION

Introduction

For decades, Outdoor Adventure Education (OAE) has provided a promising framework for enhancing adolescents' personal and social development. While previous research has shown that OAE can benefit some of the most widely studied outcomes in developmental literature, several pervasive issues have prevented progress in the area. These include: (a) a lack of well-defined outcome measures that reflect the specific goals of the OAE program and which are supported by strong measurement models; (b) a lack of longitudinal, controlled trial studies with adequate sample size and statistical power; and (c) insufficient investigation into independent factors that account for the positive outcomes from OAE. Thus, the aims of this research were to undertake rigorous psychometric evaluation of the research instrument (Study 1), and to then conduct a longitudinal randomized-controlled trial (RCT) investigation into the link between OAE and healthy psychological development in teenagers (Study 2). Firstly, this chapter summarises the key findings from the two empirical studies and discusses these within the broader OAE literature. Secondly, the strengths and limitations of this investigation, and directions for future research are addressed. Finally, the implications for educational policy and practice are discussed.

Summary of Findings

Study 1. Study 1 set out to establish the psychometric properties of the nine instruments selected to evaluate primary outcomes of self-concept, life effectiveness and Locus of Control (LoC), and secondary outcomes, including basic psychological needs satisfaction, psychological ill-being, school motivation and engagement, satisfaction with life, wellbeing, gratitude, and resilience (see Chapter 4: General Method). As the present investigation was dependent on quantitative procedures, a decision deemed to be a critical

prerequisite for exploring the important substantive issues that are the focus of Study 2. The results from Study 1, derived using Structural Equation Modelling (SEM) procedures, revealed that:

(1) all 32 scales showed acceptable internal consistency;

(2) the nine instruments' *a priori* hypothesised factor structures reflected good fit to the research sample (see Chapter 4: General Method);

(3) there was evidence for strong longitudinal invariance of all measurement models;

(4) the combined latent model (32-factors and 131-items) containing all component scales and items showed good to excellent structural validity;

(5) multitrait-multimethod (MTMM) procedures showed strong support for constructs' convergent and discriminant validity over time; and,

(6) a new three-factor, 11 item adaptation of the Motivation and Engagement Scale (MES; Martin, 2007, 2009) provided psychometrically robust measurement of cognitive, affect and behavioural changes in school motivation and engagement.

The findings from Study 1 show that the survey instrumentation selected on the basis of the OAE program's philosophy and objectives provided robust psychometric measurement of the outcomes under investigation. These results strengthen and extend existing validation into the instruments used in this research by demonstrating their suitable application in youth OAE settings.

Study 2. Study 2 set out to empirically examine the link between an OAE program and positive psychological development in male adolescents. Specifically, the aims were to:

(1) evaluate short-term program effects;

(2) evaluate six-month follow-up effects;

(3) examine the influence of students' baseline aptitude for each outcome on their sensitivity to benefiting from the OAE experience; and,

(4) examine the mediating role of Self Determination Theory's (STD) basic psychological needs on the relationship between the OAE program and the research outcomes.

Preliminary analysis. Study 2 conducted initial screening of the extended baseline period (T0 to T1). The results showed support for balanced groups at the extended baseline time-wave (T0) for 29 from 32 outcomes. The remaining three outcomes reflected significant group by time interactions between T0 and T1, resulting in a diminishing of group difference over the baseline interval. A further three outcomes showed beneficial main effects for time that were equivalent for both groups (see Table 35).

Short-term effects. To assess the main effects of the OAE intervention, the analysis compared the treatment and waitlist-control (WLC) participant scores on each outcome at post-test (T2), while controlling for group differences at the pre-test time-waves (T0 and T1). The response rate at T2 was 92% (N = 380), and all scales were scored so that higher positive ESs reflect more favourable treatment effects. The results found that students who attended the OAE program displayed statistically significant gains in a range of psychosocial and emotional domains, including the primary outcomes in this investigation: self-concept, life effectiveness and LoC. More specifically, short-term benefits were observed in:

(1) self-concept domains of Global Esteem (ES = .26), Physical Ability (ES = .22), Parent Relations (ES = .44) and Same-Sex Relations (ES = .19), an average small positive change across all six self-concept scales of ES = .19;

(2) life effectiveness areas, including Self Confidence (ES = .33), Time Efficiency (ES = .19), Active Involvement (ES = .19), and Quality Seeking (ES = .18): an average small positive change across all 11 life effectiveness scales of ES = .14;

(3) Internal LoC (ILoC; ES = .20) and External LoC (ELoC; .37)—an average moderately positive change in LoC of ES = .29;

(4) Relatedness (ES = .38) and Competence (ES = .28) basic psychological needs satisfaction;

(5) Positive (ES = .71) and Negative (ES = .53) symptoms of psychological distress: an average large positive change of ES = .62 and the highest of any outcome category; and

(6) positive psychological resources, including school motivation and engagement Thought (ES = .26) and Hampering (ES = .32), Wellbeing (ES = .44), Life Satisfaction (ES = .32) and Gratitude (ES = .35).

The findings from Study 2 support meta-analytic research observing OAE's positive short-term gains in adolescents' psychological development, including self-concept, life effectiveness and locus of control. School-embedded OAE programs are somewhat marginalised in OAE research as they are described to be less rigorous and as achieving fewer outcomes (e.g. Hattie et al., 1997). This suggestion is, however, somewhat in opposition to the identification of schools as powerful contexts to implement developmental interventions (e.g. Kaplan & Flum, 2012; O'Mara et al., 2006). As a result, school OAE programs have undergone less empirical investigation. It follows that psychological outcomes in direct reference to school experiences, such as academic motivation and engagement, and school autonomy, relatedness and competence, are less understood. This research narrows this gap in knowledge, demonstrating OAE to foster positive increases in the value students place on school, their interpersonal connectedness and their self-belief that they can be successful at school.

Long-term effects. Study 2 included a six-month follow-up (T3) to the end of program analysis (T2). Because the WLC had already embarked on the OAE program by the time the T3 data was collected, the analysis used the WLC post-test data (T2) as the

comparative reference group at T3. Post hoc analysis compared the short- and long-term regression models to estimate the changes in ES for each outcome over the post-treatment interval (T2 to T3; see Table 37). The response rate at T3 was 74%.

The analysis found that the students who attended the OAE program reported significant long-term gains in areas of self-concept, life effectiveness, basic psychological need satisfaction and other psychological resources. Some of these effects were observed at the end of the program, while others were not and thus these new gains emerged well after the OAE experience had ended. In accordance with SDT, post-program gains such as these may indicate that new values, beliefs and capacities associated with the OAE experience were successfully internalised by students. Among the greatest gains were seen in the personal resource domains of Wellbeing, Life Satisfaction and Resilience, despite these outcomes having received less attention in OAE literature (e.g. Bettmann et al., 2016; Cason & Gillis, 1994). Specifically, the longitudinal analysis observed:

1) maintenance of post-treatment effects for Global Esteem and Physical Ability self-concept, Self Confidence and Time Efficiency ROPELOC scales, Positive and Negative GHQ-12 scales, Wellbeing, Life Satisfaction, and psychological needs satisfaction for Relatedness and Competence;
2) the emergence of new, statistically significant effects for Opposite-Sex Relations self-concept, Resilience and Self Efficacy;
3) statistically significant additional increases in ESs from post-test (T2) to follow-up (T3), referred to as sleeper effects (see Chapter 2: Literature Review), for General Academic and Problem Solving self-concept, Autonomy needs satisfaction, and the motivation and engagement Behaviour scale; and,

4) diminishing of end of program gains and small negative ESs in favour of the WLC for Same-Sex Relations self-concept and the motivation and engagement Hampering scale.

This empirical investigation strengthens previous literature demonstrating long-term gains following OAE by employing rigorous methodological procedures that overcome prior criticisms of OAE research. .

Aptitude-treatment effects. In response to the call for greater investigation into the independent factors that influence OAE's effectiveness, Study 2 examined whether the effects of the OAE program on each outcome at post-test (T2) varied according to students' pre-test (T1) aptitude. An additional interaction term (i.e. intervention X respective baseline outcome at T1) was specified to assess whether target variable levels at pre-test predicted effect levels after the intervention. The results showed no systematic trends for students' pre-test aptitude on the gains that were attributable to the OAE intervention; this was the case for 30 from 32 outcomes.

Intuitively, one may anticipate that students who incur greater challenge during OAE will benefit differently to their peers. The results from this research indicate otherwise, and showed that the benefits from the outdoor experience were equally available for all students regardless of baseline aptitude for the psychological outcomes under investigation. For example, students who started the OAE program with low self-esteem, wellbeing or resilience did not benefit more or less in those categories when compared to peers who indicated more favourable pre-program scores. These findings have particular implications for the generalisability of OAE to youths of different backgrounds and levels of psychosocial functioning.

Mediation, SDT. Study 2 investigated the mediating role of SDT's basic psychological needs within the OAE treatment context. The analysis examined how

changes in Autonomy, Relatedness and Competence, due to the effects of the OAE program, accounted for changes in each outcome post-treatment (T2). The results revealed that of 19 significant treatment effects, basic psychological needs satisfaction primarily mediated 10 and partially mediated a further 8. Specifically, it was found that:

1) Relatedness primarily mediated the OAE program's effect on Same-Sex Relations self-concept and partially mediated the effects on Parent-Relations self-concept;
2) Competence primarily mediated the program's effects on Global Esteem and Physical Ability self-concept, Quality Seeking, Time Efficiency, Coping with Change life effectiveness scales, ILoC, Resilience and the school motivation and engagement Thought scale. Competence further partially mediated the program effects on Self Confidence and the Hampering motivation and engagement scale;
3) both Relatedness and Competence primarily mediated the program effects on Active Involvement; and,
4) the treatment did not predict short-term increases in Autonomy need satisfaction and thus, this psychological need showed no mediating pathways.

Few investigations have examined the role of basic psychological needs satisfaction in OAE settings, despite a growing body of empirical literature demonstrating the strong link between psychological needs satisfaction and healthy adolescent development (see Ryan & Deci, 2017). Study 2 statistically positioned SDT's basic psychological needs so to explicitly examine how these constructs influenced the effectiveness of the OAE program. The result indicated that without satisfying students basic psychological needs for Competence and Relatedness, many of the positive short-term effects from the OAE program would not have been observed. Additionally, the psychological need for Relatedness was particularly important to enable benefits in socially-oriented constructs, such as peer and parent relations self-concept. Consequently,

OAE programs aiming to enhance interpersonal functioning and social connectedness should endeavour to foster peer and staff relationships during the OAE experience.

Implications of General Findings

Baseline stability. Preliminary screening of the baseline interval (T0 to T1) revealed favourable changes on the ELoC, Hampering and Negative symptom scales, which were equivalent for both treatment and WLC groups. Interestingly, these scales represent three of the four scales in the 32 scale inventory that aim to measure negative indicators of psychological functioning. These findings may suggest firstly, that contextual or environmental factors external to the influence of the treatment setting are asserting a systematic and statistically meaningful impact on students' experience of psychosocial wellness, and secondly, that negative symptom variables are more sensitive to these effects. Such factors may relate to variations in students' response patterns on psychological measures when taken at the end, versus the start of a school term (i.e. due to fatigue, anticipatory excitement/anxiety etc.); the effects of age differences (see Chapter 2: Literature Review); and/or student perceptions of the social-educational experience when in Years 9 compared to Year 8 (e.g. beliefs around autonomy, interpersonal relatedness and self-belief). The greater malleability of clinical indicators of psychopathology in OAE settings when compared to positive psychological constructs is well documented (e.g. Bowen & Neill, 2013; Cason & Gillis, 1994; Hattie et al., 1997; Neill, 2003), although the contributing factors to the detected improvements in negative symptom constructs over the baseline period are not well understood and require further investigation.

Comparison of Effect Sizes to other major OAE reviews. The computation of standardised ESs and CIs in this investigation allows for the strength of the program's effects to be easily compared with findings from other OAE investigations and meta-analyses. Overall in this thesis, there were positive short-term changes of approximately

one third of a standard deviation unit (all 32 scale ESs ranging between .15 and .62; overall *M* ES = .30), which were similar to those reported in major OAE meta-analytic reviews (.31 to .34; Cason & Gillis, 1994; Hattie et al., 1997). Small overall positive long-term changes were approximately one sixth of a standard deviation unit (all 32 scale ESs ranging between -.14 and .33; overall *M* ES = .16) and indicated a diminishing of post-treatment effects (T2 to T3; overall *M* ES = -.14) that did not reflect the additional gains that were reported by Hattie et al. (1997; ES =.17) and Bowen and Neill (2013; g =.17) over the equivalent interval. More specifically, the results from this investigation observed:

(1) small positive increases in overall self-concept (short-term ES = .19; long-term ES = .12; see), that was less than that reported by Hattie et al. (1997; short-term ES =.28; long-term ES =.23) and Cason and Gillis (1994; short-term ES =.34);

(2) small gains in life effectiveness (short-term ES = .14; long-term ES = .12) that were substantially less than those reported by Neill (2008; short-term ES =.47; long-term ES =.31) and less than those reported by Hattie et al. (1997; short-term ES =.37; long-term ES =.18) on matching life effectiveness domains (Leadership, Teamwork, Self Efficacy, Confidence and Time Management);

(3) short-term gains in LoC (ES = .29) that were similar to those reported by Cason and Gillis (1994; ES =.30) and Hans (2000; unstandardised ES =.38), but exceeding those reported by Wilson and Lipsey (2000; ES =.10) within a delinquent youth sample;

(4) as predicted *a priori*, greatest short- and long-term gains were detected using the GHQ-12 clinical instrument (short-term ES = .62; long-term ES = .33; see Chapter 6: Study 2-Intervention, Statement of Hypotheses e.g. Bowen & Neill, 2013; Cason & Gillis, 1994);

(5) the small decreases in effect size over the post-treatment interval for self-concept (T2 to T3; ES = -.06) was similar to that reported by Hattie et al. (1997; T2 to T3; ES = -

.05). Similarly, the small decrease in overall mean life effectiveness over the post-treatment interval was only marginal (T2 to T3; ES = -.02) and indicative of good retention of short term gains that exceeded those reported by Neill (2008; T2 to T3; ES = -.10) and Hattie et al. (1997; T2 to T3; ES = -.16) on matching domains.

In summary, these results demonstrate overall end of program gains similar to those reported in meta-analytic reviews incorporating the breadth of outcomes seen in this research (e.g. Cason & Gillis, 1994; Hattie et al., 1997), larger than the effect size reported in delinquent youth focussed studies (Wilson & Lipsey, 2000), however smaller when compared to the effect sizes reported in reviews into specific outcomes of LoC (Hans, 2000) and life effectiveness (Neill, 2008). The overall long-term gains in self-concept and life effectiveness were smaller than those previously reported (see Hattie et al., 1997; Neill, 2008), although demonstrated greater retention of post-program effects in each category. Conversely, the significant end of program effects for LoC were not maintained and thus, did not support previous meta-analytic findings (e.g. Hans, 2000). This was somewhat surprising in light of the maintenance of gains seen for SDT's Competence needs satisfaction, as these constructs are closely related (see Chapter 2: Literature Review; e.g. Pelletier et al., 1999; Ryan & Deci, 2017). One possible explanation for the notable decreases in LoC over the post-treatment interval in lieu of changes in Competence surround the timing of the follow-up measure, as discussed under point (d) in the below section.

Possible explanations for smaller ESs. There are several explanations for the smaller ESs that were detected in this OAE investigation. These include: (a) the addition of adult participants in the Neill (2008) and Hattie et al. (1997) calculated ESs, as age is a widely acknowledged moderator of effect size in OAE research (e.g. Bowen & Neill, 2013; Hattie et al., 1997); (b) the 'high quality' research design (see Cason & Gillis, 1994),

which has been associated with smaller ESs compared to weaker design OAE investigations; (c) the age of participants having been confined to mid-adolescence (ages 13-15), which is typically associated with greater resistance to change; (d) the six-month follow-up not providing ample time for the long-term effects to be fully realised. Hattie et al. (1997) included studies with follow-up intervals up to 18 months, while Davis-Berman and Berman (1994) observed that post-treatment gains in self-efficacy and LoC had diminished at four-month assessment, however returned one and two years later; (e) the inclusion of an active control group who were attending a well-resourced Independent boys school; (f) the use of randomization in group-assignment, as randomized experimental designs have been shown to evidence smaller ESs when compared to alternative research designs with fewer controls over pre-existing differences (Cheung & Slavin, 2016). Such alternative designs are primarily used in OAE meta-analyses that form the basis of comparison in this thesis; and/or (g) the influence of Post Group Euphoria (PGE), program halo effects, lack of transfer learning or failure to adequately integrate the specific learnings from the OAE experience into subsequent class curriculum (e.g. see Scrutton, 2015).

An alternative approach to long-term analysis. OAE literature is largely characterised by repeated measures analysis, whereby participants' post-program scores are compared to their own pre-test scores (Neill, 2002). To provide an alternative and comparative approach to the between-groups follow-up analysis already presented (T3 treatment group versus T2 WLC), the long-term changes in the treatment group's scores from pre-test (T1) to follow-up (T3; n = 194) were examined (see Appendix 8). The results for overall self-concept (ES = .12) and overall LoC (ES = -.14) were identical when using within- and between-groups methods, however, reduced *SEs* for the within-group approach led to the detection of an additional effect for ELoC. For overall life effectiveness, the

within-groups method reflected substantially greater improvements at follow-up (T1 to T3; ES = .19) and seven significant effects, compared to the between-groups method (overall ES = .12), where three effects were detected. The mean *SE* for the 32 scales using the within-groups approach was .07—notably less than the .09 computed in between-groups analysis. Additionally, the within-group approach reflected consistently greater retention of positive short-term effects to follow-up (T2 to T3; self-concept ES = -.03 versus ES = -.05; life effectiveness ES = -.00 versus ES = -.02; LoC ES = -.35 versus ES = -.42). These results demonstrate the potential for OAE studies, without control comparisons, to produce stronger effect sizes, with lower *SE*s that are more prone to detecting statistically significant findings. This stands even after accounting for the influence of pre-test biases (e.g. see Hattie et al., 1997). These findings support and extend research by Cason and Gillis (1994), who observed that weaker designed OAE studies produced greater ESs. Such findings have implication for meta-analytic research, which should endeavour to include criteria of research quality and design when synthesising the findings from a body of OAE literature. Doing so would enhance the interpretability and generalisability of meta-analytic research against specific OAE investigations and other research disciplines, while generating much-needed insight into this important substantive area.

Self Determination Theory and The Glengarry Progam (TGP). The school-embedded OAE program in this thesis, TGP, was associated with short- and long-term increases in students' Relatedness and Competence psychological needs satisfaction (see Figure 21 and Figure 22, respectively), and long-term gains in Autonomy (see Figure 23) that were not otherwise observed at the time when the OAE program concluded.

TGP aimed to foster students' school-orientated Relatedness and Competence satisfaction by: (a) establishing meaningful relationships between students and their peers/staff; (b) appropriately challenging students relative to their individual capacities and

prior experience; (c) assisting students to pursue personal goals and master new skills; and (d) providing students with positive and encouraging feedback on their performance.

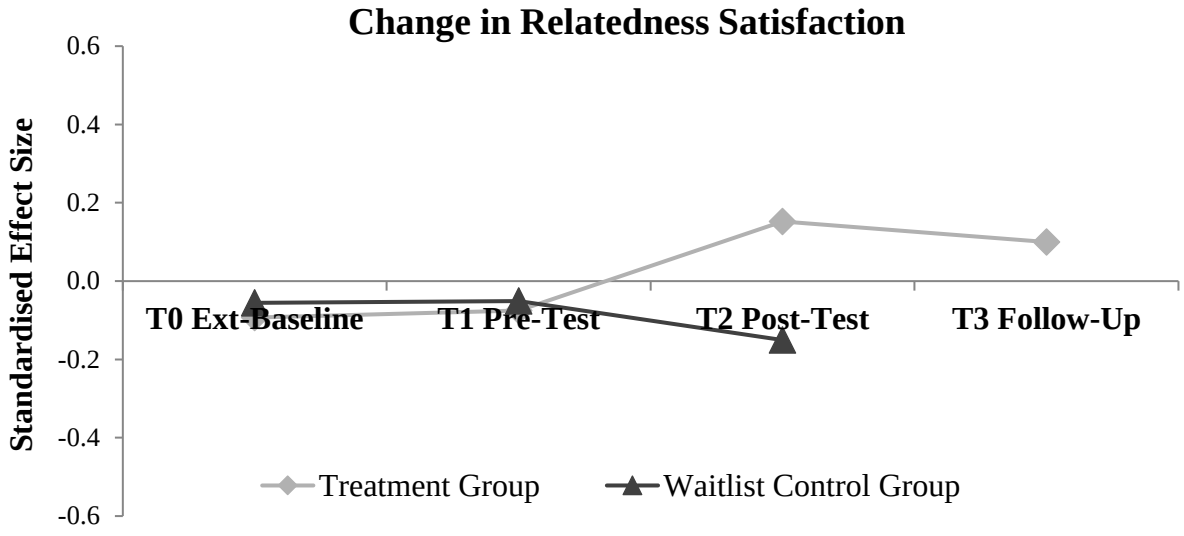


Figure 21. Overall mean change in Relatedness need satisfaction for the treatment group and WLC over the four time-waves.

Note. WLC data at T3 was derived from the WLC T2 scores.

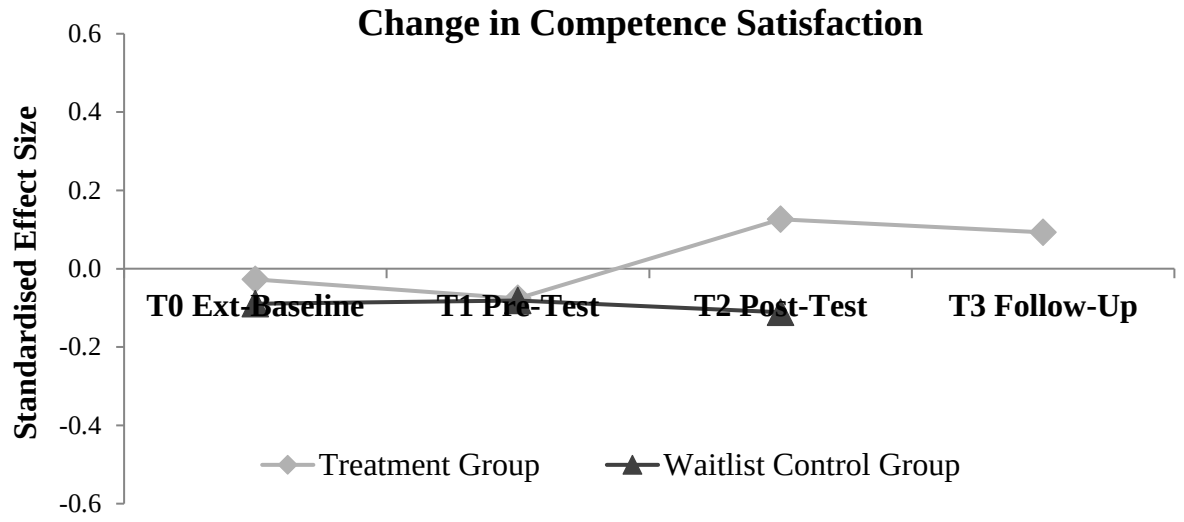


Figure 22. Overall mean change in Competence need satisfaction for the treatment group and WLC over the four time-waves.

Note. WLC data at T3 was derived from the WLC T2 scores.

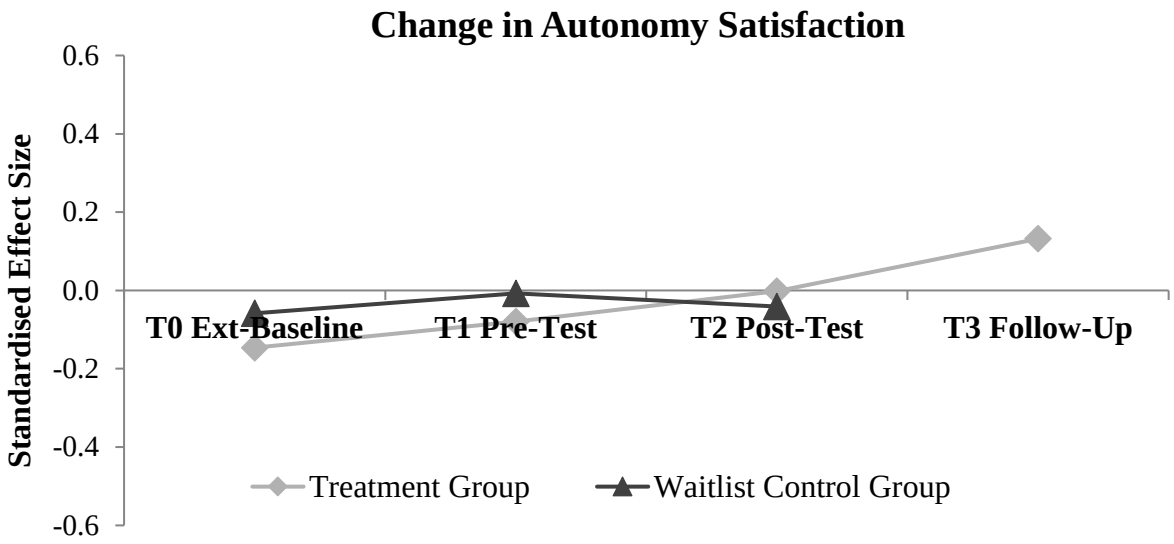


Figure 23. Overall mean change in Autonomy for the treatment group and WLC over the four time-waves.
Note. WLC data at T3 was derived from the WLC T2 scores.

Autonomy-supportive processes were: (a) exposing students to the OAE program philosophy and rationale from the time they enrol at the high-school, thus encouraging students to attribute greater meaning to the experience; (b) providing students with opportunities to make decisions and pursue interests; and (c) structuring the program so as to offer a positive and encouraging experience (see Wang et al., 2004). However, external regulations with potential to inhibit Autonomy satisfaction during the five-month OAE experience were the compulsory nature of the program and the dense, structured timetable. This latter feature is in part needed for safety, but otherwise is driven by the program's strong educational and developmental philosophy. In accordance with SDT, OAE programs need to consider the balance between autonomy supportive processes and external control expectancies. In doing so, students' intrinsically motivated engagement can be enhanced, thus allowing new values, attitudes and skills from the OAE experiences

to be successfully internalised so they may emerge in the form of long-term positive change.

Of interesting note were the higher correlations observed between basic psychological needs satisfaction and Opposite-Sex Relations self-concept when compared to Same-Sex-relations self-concept. While not an a priori prediction, one may speculate whether this was influenced by the single-sex school context. Students from non-coeducational environments receive greater exposure to same-gendered peers and thus have more opportunities to successfully establish friendships that inform domain related self-concept. Conversely, establishing opposite-gendered friendships outside of the school setting would be expected to require greater effort and intrinsically motivated behaviour. In such instances, satisfaction of basic psychological needs could provide the resources to enable a student to seek and achieve opposite-gendered friendships that subsequently inform the development of positive Opposite-Sex Relations self-concept. However, further research is required to test this idea.

Autonomy support and OAE. The OAE intervention in this research was not associated with short-term increases in school-orientated Autonomy satisfaction, although increases in Autonomy were observed for treatment participants at six-month follow-up. These findings may be explained by the fundamental aim of educational/developmental OAE experiences: to provide an optimal climate of growth where successful internalisation of new values and belief systems can later be applied to participants' everyday environments. It is this 'downstream effect' that is most important in SDT theory (see Chapter 2: Literature Review, Organismic integration theory). Thus, one may argue that the better indication of an OAE intervention's success is not the maximisation of real-time enjoyment, pursuit of personal interest and self-determined engagement, but more how these variables are operationalised in life after the experience. The linear increase in school

autonomy seen from T0 to T3 (see Figure 23), an approximate 12-month period, demonstrates this downstream effect. Although the OAE program was not associated with greater school-orientated autonomy at the end of the program, six months later students reported greater value and interest in what they were doing at school. Additionally, the alternative within-groups long-term analysis that is presented in Appendix 8 shows that the greatest improvement in students' psychological needs satisfaction over the OAE experience was in Autonomy, the most pertinent of the three psychological needs according to SDT theorists. It is not suggested that autonomy needs satisfaction is less important in OAE settings—in fact, the significant zero order *r*s between autonomy and all 32 outcomes at the end of the OAE program (T2) clearly suggest otherwise (see Table 39). It is also imperative to clarify that the non-significant post-test result for autonomy in this investigation does not indicate that the OAE experience was associated with low autonomy, but rather that students' autonomy was similar across the control and experiment groups, the usual school program versus OAE. In this research, the mean autonomy scale score for the WLC reference group was 4.34/6, with an average response to autonomy supportive statements falling between '*Agree a little*' and '*Agree*'. One possible inference is that when satisfaction of students' baseline autonomy is inferred, it is not necessary for OAE to enhance autonomy above and beyond that experienced in students' daily lives in order to be effective.

Strengths

Methodological contributions. The research design and statistical methods used in this OAE investigation were informed by best-practice recommendations in leading OAE reviews (e.g., Cason & Gillis, 1994; Dillon, 2013; Marsh et al., 1986c; Scrutton & Beames, 2015; Wilson & Lipsey, 2000). The strengths from the investigation include: (1) a longitudinal RCT research design supported by adequate sample size;

(2) the use of advanced statistical procedures, including SEM, which enabled imputation methods for missing data, minimisation of error in measurement models, and application of multi-level regression models and latently derived factor scores;

(3) results computed as standardised effect sizes that can be easily interpreted against, and synthesized into meta-analytic research;

(4) multiple pre/post-test measures and the inclusion of a control scale that helped identify and control for confounding influences (e.g. pre/post-test biases); and,

(5) psychometric validation of nine quantitative psychological instruments within a youth population.

Statistical procedures. SEM procedures formed the basis of the psychometric analysis in this thesis. This made it possible to examine complex multivariate relationships between indicator items and latent variables, and provided critical psychometric information that enabled model modifications, justified both theoretically and substantively. The effects of the intervention were inferred using factor score regression weights, to represent students' responses on each scale at the four time-waves. This method minimised measurement error by adopting a unit loading approach to the computation of derived scores that enhanced the reliability and validity of the analysis (Rowe & Rowe, 1999). Following the recommendations of Neill (2008), this research presents standardised effect sizes and CIs, as doing so enables greater convergence of literature and outcome generalizability (also see Cason & Gillis, 1994; Scrutton & Beames, 2015).

Empirical SEM literature recommends that both CFA and ESEM methods should be compared during structural validation procedures (Marsh et al., 2014; Marsh et al., 2013). CFA enables direct comparison of competing models' latent construct relationships, a critical step for theory testing, however ESEM offers greater flexibility (Marsh et al.,

2014) and thus has particular utility in applied social science research. The psychometric procedures used in this research demonstrate the complementary nature of both CFA and ESEM methods in identifying the strongest structural representation through which complex multivariate relationships and substantive research questions can be explored.

Many OAE investigations have small sample sizes. This, according to Scrutton and Beames (2015), begins to limit statistical parameters (see Cohen, 1988) in samples fewer than 80 to 100. To maximise sample size, this investigation applied stacked, or 'long' format data using the Mplus complex design program tool (Muthén & Muthén, 1998-2017). Doing so allowed for the combined structural model that integrated all scales and items, to be tested for structural validity—a step that has critical substantive implications in light of the administration procedures, which required participants to complete all nine instruments in a single session. Furthermore, post-hoc power analysis was conducted to confirm that the sample size was appropriate for the statistical procedures used in this research (See Chapter: General Method).

Survey design. A variety of self-report psychological instruments are available, however many have not been adequately tested in regard to their psychometric utility in OAE research. Scrutton and Beames (2015) recommend that researchers independently design and rigorously test the questionnaires they use, rather than accepting already-established measurement tools that are inappropriate to the specific OAE setting. This was the process in this thesis, to address criticisms with regard to inadequate selection and measurement of dependent variables in OAE research, and establish a psychometrically strong premise for the intervention analysis.

Research design. This investigation addressed several design protocol issues which have left OAE literature open to scrutiny. The multiple-cohort structure (using two consecutive Year 9 groups) increased the sample size and generalisability of results. The

randomly assigned WLC provided a strong representative comparison against which to infer treatment effects, while the inclusion of multiple pre/post-test measures allowed for investigation of group differences and the stability of constructs over the baseline period, extraneous variables to be identified and statistically controlled for (e.g. pre/post-test biases), and substantive longitudinal hypotheses to be explored.

Controlling for confounding variables and pre/post-test biases. OAE literature has been criticised for using design procedures that inadequately control for measurement error due to confounding variables. As recommended in major OAE reviews (e.g. Scrutton & Beames, 2015), the methods used in this research minimised the influence of extraneous variables by: (a) including a control group against which pre-treatment biases could be examined; (b) including statistical controls for baseline group differences in experimental analyses; (c) incorporating a control scale and follow-up assessment on all outcomes to test for PGE effects (see Marsh et al., 1986c); and (d) using randomized assignment methods to allocate students to the treatment or WLC conditions.

In accordance with the original aims of this research, rigorous psychometric procedures were applied to the survey instrumentation. This was a critical step to overcome criticisms regarding poor measurement and inappropriate survey design which have slowed forward progress of OAE literature. The RCT longitudinal design that was used in this research further enabled strong, empirical investigation into important substantive questions about OAEs effects. This acts to strengthen the body of OAE literature and is particularly important in the Australian context, as the government is in the midst of debate regarding the implementation of new national curriculum. The contribution of this research is thus timely to inform policy debate (e.g. see Australian Government Department of Education and Training, 2019).

Theoretical contributions. This investigation empirically evaluated an OAE program in relation to key psychological outcomes among adolescent males. The theoretical strengths from the investigation within the context of outdoor education include:

(1) evaluation of a powerful OAE intervention program conducted away from the school environment (i.e. five-month duration, residential, school embedded, clear philosophy and rationale; e.g. see Hans, 2000; Hattie et al., 1997; Hunter & Purcell, 1984; Neill, 2002; Richmond et al., 2018);

(2) exploration of short- and long-term effects;

(3) examination into the influence of 32 baseline psychological constructs on program effectiveness;

(4) the application of SDT as a mediating concept in an OAE setting;

(5) investigation into 32 psychosocial and emotional outcomes pertinent to healthy adolescent development;

(6) a minimum response rate of 74% (T3) and mean response rate of 86% across the four time-waves (T0, T1, T2 and T3); and,

(7) longitudinal analysis with comparison of alternative statistical approaches.

Breadth of psychological outcomes. Few independent OAE studies have examined the breadth of psychological constructs seen in this research. OAE programs vary substantially with regard to goals, content and philosophy (e.g. Priest & Gass, 1998), so selecting outcomes that are informed by these factors is an essential requirement to preserve research integrity. The OAE program in this research, TGP, aims to cater for all aspects of students' psychosocial and emotional development (see Chapter 3: The Intervention). Thus, there was a rationale to include outcomes spanning these areas. Outside of the present research, self-concept, life effectiveness and LoC have received

particular attention (e.g. Cason & Gillis, 1994; Gass et al., 2012; Hans, 2000; Hattie et al., 1997; Neill, 2008). The other personal psychological outcomes used in this investigation were selected on the basis of their theoretical and substantive alignment to the overarching philosophy of OAE (e.g. Cason & Gillis, 1994; Gass et al., 2012; Hans, 2000; Hattie et al., 1997; Neill, 2008), their predetermined psychometric suitability in adolescent populations, and how they matched the idiosyncratic goals of the OAE intervention.

Baseline aptitude as a mechanism for change. This investigation found that students' pre-test self-perceptions of self concept, personal effectiveness, LoC and other indicators of positive psychological functioning did not predict the benefits they subsequently reported on these same outcomes at the end of the OAE program. In a recent study by Scrutton (2015), however, school children showing the greatest benefits following a one-week OAE course perceived themselves as having poorer pre-program social skills. The important substantive question, 'do those with more room to grow show more growth through OAE?' warrants further investigation.

Criticisms of OAE literature regard the notable lack of longitudinal RCT studies and the insufficient investigation into the mechanisms by which OAE fosters positive growth in participants. The principles of SDT have been the subject of strong empirical investigation, with a growing body of literature demonstrating the link between basic psychological needs satisfaction and healthy outcomes among youths (e.g. Hansen & Jessop, 2017; Ryan & Deci, 2002, 2017). However, SDTs psychological needs have received surprisingly little attention in OAE settings. The OAE program in this research, while not specifically designed with the principles of SDT in mind, was shown to lead to increased basic psychological needs satisfaction compared to students attending their usual school timetable. The present investigation empirically examined these increases and

demonstrated that psychological needs satisfaction was a critical mechanism that explained the beneficial program effects on other outcome categories.

Limitations and Directions for Future Research

There were several important limitations to the present investigation which can be addressed in future OAE studies. These include:

(1) the lack of adequate evaluation of construct validity. While the analyses used in this research found strong support for each instrument's factor structure and measurement invariance, it is recommended that further psychometric validation be conducted across a variety of settings. This is especially relevant when component scales are to be administered and/or modelled in variation to the approach applied in this investigation;

(2) generalisability. The intervention studied in this research constitutes a best-practice OAE program in relation to what is typically and practically possible in applied school settings (see Chapter 3: The Intervention Program). Hence, a key issue is whether the positive results from this research are generalizable to other OAE settings which are less well-supported in terms of time, school commitment and resources. Furthermore, the research sample, while diverse in background characteristics (e.g. nationality, family composition, race, academic ability), all attend a single all-boys independent school. Hence there is need for further research in a range of different settings to test the generalisability of results;

(3) the breadth of the survey instrument. The quantitative instrument used in this research contained 32 scales and 131 items. Survey fatigue can impact on response accuracy and motivation, especially for students experiencing literacy or language difficulties. Furthermore, the number of observed variables and factors in the instrument relative to sample size meant that small clusters of theoretically related

scales needed to be psychometrically evaluated in isolation from others, to overcome limitations with software and computing power. Future studies could apply additional dependent variable ranking methods so as to include only those outcomes that are most relevant to program objectives. For example, program facilitators could be asked to rank outcomes in order of importance, and/or content analysis of the specific OAE program could be undertaken. In this way, the survey instrument could be refined to focus only on the most pertinent scales;

(4) use of iPads and paper-pencil survey methods. This research required hard-copy surveys to be completed for the treatment group at T1 (both cohorts), as well as for the first year of treatment participants' T2 data collection occurring off-site. (I.e. a New South Wales south coast beach house). Future studies could examine differences in responses as a function of survey delivery method and attempt to standardise procedures for all participants across all time-waves;

(5) exceptions to randomization procedures. As the research occurred in a high-school setting, instances occurred where students were required to attend the OAE program at a specified time. For example, due to student welfare or personal/family reasons;

(6) the absence of longitudinal true no-treatment control data. Given that the WLC condition was only temporarily available due to the school's requirements, the research design lacked data for the WLC at follow-up (T3). This was addressed by using a time-lapse control comparison (WLC T2 as the reference for Treatment T3) as a reasonable compromise, however future studies should strive to incorporate true control groups at all time-waves;

(7) the novel application of SDT. Many OAE programs, including TGP, were not designed with the principles of SDT in mind. Future studies should also explore

alternative theories which may enhance our knowledge of the processes of change following OAE experiences;

(8) reliance on empirical self-report data. While self- and other-report scores for self-concept show general agreement (Marsh, 1990b), it is recommended that future studies attempt to include student outcome indicators rated from multiple perspectives;

(9) the six-month follow-up. Longitudinal OAE research demonstrating program effects to manifest over time, frequently include follow-up measures of 12 months or longer (e.g. Davis-Berman & Berman, 1994; Hattie et al., 1997). It is recommended that future OAE studies endeavour to include multiple longitudinal measures extending 12 to 24 months after the OAE program concludes;

(10) the single measure of basic psychological needs satisfaction. In this research, basic psychological needs satisfaction comprised the measure on which SDT analysis and interpretations were inferred. Future OAE studies exploring the mechanisms of SDT may consider additional theory- informed instruments, such as measures of internalisation, intrinsic and extrinsic motivation, learning transference, perceived locus of causality, psychological needs frustration, and/or items targeting the 'meaning/value' dimension of autonomy;

(11) an extended OAE program. Meta-analyses and major OAE reviews consistently report longer programs are associated with more favourable outcomes (e.g. Bettmann et al., 2016; Cason & Gillis, 1994; Hattie et al., 1997; O'Brien et al., 2011; Wilson & Lipsey, 2000). However, to the author's best knowledge, there are no controlled studies that have tested an OAE program with the level of exposure as that in this research (i.e. 5-months duration with residential boarding). The field would benefit

from further research into whether comparable effect sizes can be gained from programs with less exposure/duration; and,

(12) control over extraneous variables. Because this research used true random assignment, many potential confounding factors were appropriately controlled for. However, the fact that a waitlist control was used meant that there could possibly be confounding influences for variations in time of year for treatment exposure. Future research should conduct further investigation into within-person and programmatic factors such social desirability (see Paulhus, 1991), post-course adjustment (see Allison, 2000) and response shift bias (see Ewert & Sibthorp, 2009; Sibthorp, Paisley, Gookin, & Ward, 2007). Furthermore, future research may include measures of achievement (e.g. physical performance/academic outcomes), which allow for control over their influence and investigation into potential interactions with psychological outcomes being explored.

Additionally, it is recommended that future OAE research investigates the influence of various aptitude variables that may have substantive relevance to the OAE setting, using different samples, while also considering how these interact with other independent factors (e.g. socioeconomic status; Scrutton, 2015). Advanced statistical methods and the reporting of inferential, standardised effect sizes and advanced SEM procedures should also be more widely applied in OAE research. Lastly, it is recommended that caution be taken when interpreting results from pre-/post-test OAE investigations, which may show an increased statistical sensitivity in respect of significance, that is independent of the inflated ESs associated with pre-/post-test response biases.

Implications for Educational Policy and Practice

Recent decades have seen educational policy makers place greater importance on the need for empirical research to inform the educational initiatives that are used in schools. Given the substantial human, environmental and financial capital, not to mention time, that OAE requires, the need for strong evidence with regard to design and anticipated outcomes appears justified. The findings from the present investigation have several important implications for the policy and practice of OAE.

First, there was clear evidence that the OAE program provided a valuable learning opportunity that were associated with long-term changes in multiple psychological and personal indicators of healthy development. Policy makers are placing increased attention on non-cognitive soft skills, such as those personal resource outcomes that are the focus of this research, and which appear to be more important than cognitive hard skills in determining academic outcomes (Gutman & Schoon, 2013). For example, this investigation found that the OAE experience led to greater Academic Ability self-concept six-months after the program had ended; this construct is linked closely with academic achievement (Marsh & Martin, 2011). However, teachers in Australia are often found to be evaluated, directly or indirectly, on the basis of objective indicators of their students' academic success. Such pressures would undermine the valuable developmental contribution that OAE appears to provide in light of the time involvement, and may thus inhibit educators' capacity to provide the most effective learning experiences for their students.

Second, the benefits from the OAE experience were available to all students, irrespective of their baseline levels on the psychosocial and emotional indicators used in this research. These results are consistent with OAE literature that shows comparable effect sizes following OAE for youths of varying behavioural, social and emotional

functioning (e.g. Bowen & Neill, 2013; Cason & Gillis, 1994; Hans, 2000; Hattie et al., 1997; Neill, 2003; Wilson & Lipsey, 2000). Williams (2013) found that while school-embedded OAE programs are on the rise, such opportunities discriminate against lower socioeconomic status (SES) catchment areas. Policy makers should explore strategies to provide equal access to OAE for all schools/students irrespective of SES or other considerations, as the benefits from OAE appear to be largely available to all teenagers. This suggestion is pertinent to current educational policy and financial reform that is the subject of debate in Australia (see Australian Government Department of Education and Training, 2019).

Third, the principles of SDT were largely applicable in the OAE setting, and support the well-documented relationship between basic psychological needs and positive health outcomes (e.g. Evans & Bonneville-Roussy, 2016; Jang et al., 2012; Ryan & Deci, 2017; Thomaes et al., 2017). This investigation observed that the OAE experiential framework was effective in providing students with a greater sense of social connectedness and perceived competence, which largely explained the benefits seen in other psychosocial domains. Align with recommendations by Wang et al. (2004), it is advised that OAE programs foster students' basic psychological needs by:

(1) providing a rationale for students' participation;

(2) fostering positive and meaningful relationships;

(3) including opportunities for decision-making; and

(4) providing positive feedback and opportunities to master new skills.

Furthermore, Scrutton (2015), suggests that integrating key learnings from an OAE experience into subsequent class curriculum is important for the longevity of program outcomes.

Conclusion

Outdoor Adventure Education (OAE) is used as a way to foster adolescents' technical, social and personal development in countries all over the world. This investigation furnished two empirical studies that strengthen previous research and establish the link between OAE and healthy psychological development in teenagers. The first important conclusion from this research is that OAE offers a robust educational framework to cultivate lasting benefits in students' self-concept, life effectiveness, and other key personal resources that are associated with desirable academic and vocational outcomes. Additionally, the benefits from OAE appear to be equally available to all students, irrespective of baseline psychosocial and emotional functioning.

The second important conclusion is that the principles of Self Determination Theory are well suited to OAE settings, in that basic psychological needs satisfaction offers a critical mechanism through which OAE may foster healthy development in teenagers. Until now, few studies have empirically tested SDT in outdoor settings. Together with the finding that all students benefitted equally from the OAE program, the close link between Competence needs satisfaction and positive outcomes in this research may suggest that the capability of OAE programs to provide a uniquely challenging experience for each student is an important design aspect. These results complement the broader body of SDT literature (e.g. Hansen & Jessop, 2017; Hughes et al., 2011; Ryan & Deci, 2002, 2017; Schunk & Pajares, 2005) and establish a strong case for the future application of SDT theory to OAE literature.

The results from this RCT investigation strengthen and extend previous OAE literature by enhancing the methodological rigor of the field of research and contributing much-needed empirical evidence into the design of best practice OAE programming.

These findings have the potential to inform national debate regarding educational policy and reform, while providing a strong platform for future investigations to build on.

REFERENCES

Abubakar, A., & Fischer, R. (2012). The factor structure of the 12-item General Health Questionnaire in a literate Kenyan population. *Stress Health, 28*(3), 248-254. doi:10.1002/smi.1420

Allison, P. (2000). Research from the ground up: Post-expedition adjustment. *Occasional Paper Number 1*, Ambleside, Cumbria, UK: Brathay Hall Trust.

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* (Sixth ed.): Washington, D.C. : American Psychological Association.

Andrich, D., & Van Schoubroeck, L. (1989). The General Health Questionnaire: A psychometric analysis using latent trait theory. *Psychological Medicine, 19,* 469-485.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*(3), 397-438. doi:10.1080/10705510903008204

Australian Bureau of Statistics. (2015). *National Health Survey: First Results, 2014-15*. Retrieved from http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012014-15

Australian Bureau of Statistics. (2016). Suicide in Australia. Retrieved from http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/3303.0~2015~Main%20Features~Intentional%20self-harm:%20key%20characteristics~8

Australian Government Department of Education and Training. (2019). The National School Reform Agreement. Retrieved from https://www.education.gov.au/national-school-reform-agreement-0

Bachman, J. G., O'Malley, P. M., & Johnston, J. (1978). *Youth in transition: Adolescence to adulthood- Change and stability in the lives of young men* (Vol. 6). Ann Arbor: Institute for Social Research.

Bandura, A. (1997). *Self-Efficacy: The exercise of control*. Freeman: New York.

Barrett, J., & Greenaway, R. (1995). *Why adventure? The role and value of outdoor adventure in young people's personal and social development*. Coventry, UK: Foundation for Outdoor Adventure.

Bartholomew, K. J., Ntoumanis, N., Ryan, R. M., Bosch, J. A., & Thogersen-Ntoumani, C. (2011). Self-determination theory and diminished functioning: the role of interpersonal control and psychological need thwarting. *Personality and Social Psychology Bulletin, 37*(11), 1459-1473. doi:10.1177/0146167211413125

Bartholomew, K. J., Ntoumanis, N., & Thøgersen-Ntoumani, C. (2010). The controlling interpersonal style in a coaching context: development and initial validation of a psychometric scale. *Journal of Sport & Exercise Psychology, 32*(2), 193. doi:10.1123/jsep.32.2.193

Baumeister, R. F. (1999). *The self in social psychology*: Philadelphia, Pa. : Taylor & Francis.

Bell, B. J., & Holmes, M. R. (2011). Important factors leading to outdoor orientation program outcomes: a qualitative exploration of survey results. *Journal of Outdoor Recreation, Education and Leadership, 3*, 26.

Bentler, P. M. (1990). Comparative Fit Indexes in Structural Models. *Psychological Bulletin, 107*(2), 238-246. doi:10.1037/0033-2909.107.2.238

Bettmann, J. E., Gillis, H. L., Speelman, E. A., Parry, K. J., & Case, J. M. (2016). A Meta-analysis of Wilderness Therapy Outcomes for Private Pay Clients. *Journal of Child and Family Studies, 25*(9), 2659-2673. doi:10.1007/s10826-016-0439-0

Blakemore, S. J., Burnett, S., & Dahl, R. E. (2010). The role of puberty in the developing adolescent brain. *Human Brain Mapping, 31*(6), 926-933. doi:10.1002/hbm.21052

Blaxter, M. (1990). *Health and lifestyle*. London: Routledge.

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., & 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

Briggs, N. (2006). *Estimation of the standard error and confidence interval of the indirect effect in multiple mediator models.* ProQuest Dissertations Publishing.

Brookes, A. R. (2006). *Situationist outdoor education in the country of lost children.*

Browne, M. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36*(1), 111-150. doi:10.1207/S15327906MBR3601_05

Browne, M., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research, 21*(2), 230-258.

Burton, K. D., Lydon, J. E., D'Alessandro, D. U., & Koestner, R. (2006). The Differential Effects of Intrinsic and Identified Motivation on Well-Being and Performance: Prospective, Experimental, and Implicit Approaches to Self-Determination Theory. *Journal of Personality and Social Psychology, 91*(4), 750-762. doi:10.1037/0022-3514.91.4.750

Bye, D., Pushkar, D., & Conway, M. (2007). Motivation, Interest, and Positive Affect in Traditional and Nontraditional Undergraduate Students. *Adult Education Quarterly: A Journal of Research and Theory, 57*(2), 141-158. doi:10.1177/0741713606294235

Byrne, B. M. (1996). *Measuring self-concept across the life span: issues and instrumentation* (1st ed.). Washington, DC: Washington, DC: American Psychological Association.

Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life: Perceptions, evaluations and satisfaction*. New York: Russell Sage.

Campbell, D. T., & 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

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105.

Casas, F., Bălţătescu, S., Bertran, I., González, M., & Hatos, A. (2012). School Satisfaction Among Adolescents: Testing Different Indicators for its Measurement and its Relationship with Overall Life Satisfaction and Subjective Well-Being in Romania and Spain. *Social Indicators Research, 111*(3), 665-681. doi:10.1007/s11205-012-0025-9

Casas, F., & Rees, G. (2015). Measures of Children's Subjective Well-Being: Analysis of the Potential for Cross-National Comparisons. *The official Journal of the International Society for Child Indicators, 8*(1), 49-69. doi:10.1007/s12187-014-9293-z

Casas, F., Sarriera, J. C., Abs, D., Coenders, G., Alfaro, J., Saforcada, E., & Tonon, G. (2011). Subjective Indicators of Personal Well-Being among Adolescents. Performance and Results for Different Scales in Latin-Language Speaking Countries: A Contribution to the International Debate. *Child Indicators Research, 5*(1), 1-28. doi:10.1007/s12187-011-9119-1

Cason, D., & Gillis, H. L. (1994). A Meta-Analysis of Outdoor Adventure Programming with Adolescents. *Journal of Experiential Education, 17*(1), 40-47. doi:10.1177/105382599401700109

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., . . . Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion, 39*(2), 216-236. doi:10.1007/s11031-014-9450-1

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.

Chubb, N. H., Fertman, C. I., & Ross, J. L. (1997). Adolescent self-esteem and locus of control: a longitudinal study of gender and age differences. *Adolescence, 32*, 113-129.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. 2nd: Hillsdale, NJ: Lawrence Erlbaum.

Cooley, C. H. (1902). Human nature and the social order. New York: Scribner's. *The Looking-Glass Self". http://wizard. ucr. edu/~ bkaplan/soc/lib/coollkgl. html*.

Coopersmith, S. A. (1967). *The Antecedents of Self Esteem*. San Francisco: W. H. Freeman.

Cronbach, L. J. (1951). Coefficent alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Cummins, R. A., & Lau, A. D. L. (2005). *Personal Wellbeing Index - School Children (PWI_SC) (3rd Ed)*.

Daniel, B., Bobilya, A. J., Kalisch, K. R., & McAvoy, L. H. (2014). Autonomous Student Experiences in Outdoor and Adventure Education. *Journal of Experiential Education, 37*(1), 4-17. doi:10.1177/1053825913518892

Davis-Berman, J., & Berman, D. S. (1994). Research Update: Two-year follow-up report for the Wilderness Therapy Program. *Journal of Experiential Education, 17*, 48-50. doi:10.1177/105382599401700110

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*, 105-115.

Deci, E. L. (1972a). Effects of contingent and non-contingent rewards and controls on intrinsic motivation. *Organizational Behavior and Human Performance, 8*, 217-229.

Deci, E. L. (1972b). Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology, 22*, 113-120.

Deci, E. L. (1975). *Intrinsic Motivation*. New York: Plenum.

Deci, E. L., & Ryan, M. R. (1985). Intrinsic motivation and self-determination in human behavior. *New York : Plenum*.

Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality. *Perspectives on motivation, 38*, 237-288.

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry, 11*(4), 227-268. doi:10.1207/S15327965PLI1104_01

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and Education: The Self Determination Perspective. *Educational Psychologist, 26*, 325-346.

Department of Health (Producer). (2016, October 17). Australia's Physical Activity and Sedentary Behaviour Guidelines for Young People (13-17 years). Retrieved from http://www.health.gov.au/internet/main/publishing.nsf/content/health-pubhlth-strateg-phys-act-guidelines#apa1317

Dewey, J. (1938). *Experience and Education*. New York, NY: MacMillian.

Dillon, J. (2013). Barriers and Benefits to Learning in Natural Environments: Towards a Reconceptualisation of the Possibilities for Change. *Cosmos, 8*(2), 153-166. doi:10.1142/s0219607712300056

Dolgin, R. (2014). Into the Wild: A Group Wilderness Intervention to Build Coping Strategies in High School Youth Through Collaboration and Shared Experience.

Journal of Creativity in Mental Health, 9(1), 83-98. doi:10.1080/15401383.2013.864963

Dunn, T. J., Baguley, T., & Brunsden, 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

Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward: Reality or myth? *American Psychologist, 51*, 1153-1166.

Ekstrom, R. B., Goertz, M. E., Pollack, J. M., & Rock, D. A. (1986). Who drops out of high school and why? Findings from a national study. *Teachers College Record, 87*, 356-373.

Ellis, L. A., Marsh, H. W., & Craven, R. G. (2009). Addressing the challenges faced by early adolescents: a mixed-method evaluation of the benefits of peer support. *American Journal of Community Psychology, 44*(1-2), 54-75. doi:10.1007/s10464-009-9251-y

Enders, C. K. (2008). A Note on the Use of Missing Auxiliary Variables in Full Information Maximum Likelihood-Based Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal, 15*(3), 434-448. doi:10.1080/10705510802154307

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Erikson, E. (1968). Identity: Youth and crisis. *New York: Norton.*

Evans, P., & Bonneville-Roussy, A. (2016). Self-determined motivation for practice in university music students. *Psychology of Music, 44*, 1095-1110. doi:10.1177/0305735615610926

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

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi:10.3758/BF03193146

Froh, J., Bono, G., & Emmons, R. (2010). Being grateful is beyond good manners: Gratitude and motivation to contribute to society among early adolescents. *Motivation and Emotion, 34*(2), 144-157. doi:10.1007/s11031-010-9163-z

Fry, S. K., & Heubeck, B. G. (1998). The effects of personality and situational variables on mood states during outward bound wilderness courses: An exploration. *Personality and Individual Differences, 24*(5), 649-659. doi:10.1016/S0191-8869(97)00230-4

Fuentes, M. C., García, J. F., Gracia, E., & Lila, M. (2011). Self-concept and psychosocial adjustment in adolescence. *Psicothema, 23*(1), 7-12.

Gass, M. A., Garvey, D. E., & Sugerman, D. A. (2003). The Long-Term Effects of a First-Year Student Wilderness Orientation Program. *The Journal of Experiential Education, 26*, 34-40.

Gass, M. A., Gillis, H. L., & Russell, K. C. (2012). *Adventure therapy: theory, research, and practice*. New York: Routledge.

Geelong Grammar School Exceptional Education. (2016). TIMBERTOP. Retrieved from https://www.ggs.vic.edu.au/School/Campuses/timbertop

Gillet, N., Vallerand, R., & Lafrenière, M.-A. (2012). Intrinsic and extrinsic school motivation as a function of age: the mediating role of autonomy support. *Social Psychology of Education, 15*(1), 77-95. doi:10.1007/s11218-011-9170-2

Gillis, H. L., & Speelman, E. (2008). Are Challenge (Ropes) Courses an Effective Tool? A Meta-Analysis. *Journal of Experiential Education, 31*(2), 111-135. doi:10.1177/105382590803100202

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

Gnambs, T., & Hanfstingl, B. (2015). The decline of academic motivation during adolescence: an accelerated longitudinal cohort analysis on the effect of psychological need satisfaction. *Educational Psychology, 36*(9), 1691-1705. doi:10.1080/01443410.2015.1113236

Goldberg, D. P. (1972). *The detection of psychiatric illness by questionnaire: A technique for the identification and assessment of non-psychotic psychiatric illness*. London, New York: Oxford University Press.

Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Journal of Psychological Medicine, 27*, 191-197.

Gottfried, A. E., Gottfried, A. W., Morris, P. E., & Cook, C. R. (2008). *Low Academic Intrinsic Motivation as a Risk Factor for Adverse Educational Outcomes: A*

Longitudinal Study from Early Childhood Through Early Adulthood: Oxford University Press.

Graham, J. W. (2003). Adding Missing-Data-Relevant Variables to FIML-Based Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal, 10*(1), 80-100. doi:10.1207/s15328007sem1001_4

Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in Children's Learning: An Experimental and Individual Difference Investigation. *Journal of Personality and Social Psychology, 52*(5), 890-898. doi:10.1037/0022-3514.52.5.890

Gutman, L. M., & Schoon, I. (2013). *The impact of non-cognitive skills on outcomes for young people*. Education Endowment Foundation.

Hahn, K. (1957). *Outward Bound*. New York: World Books.

Hans, T. A. (2000). A Meta-Analysis of the Effects of Adventure Programming on Locus of Control. *Journal of Contemporary Psychotherapy, 30*, 33-60.

Hansen, D. M., & Jessop, N. (2017). A Context for Self-Determination and Agency: Adolescent Developmental Theories *Development of self-determination through the life-course*: Dordrecht, Netherlands: Springer.

Harter, S. (2006). The Self. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (6th ed., Vol. 3, pp. 505-570). New York: John Wiley & Sons.

Hattie, J. (2009). Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement. *Routledge*.

Hattie, J., Marsh, H. W., Neill, J. T., & Richards, G. E. (1997). Outward Bound and adventure education: Out-of-class experiences that make a lasting difference. *Review of Educational Research, 67*, 43-87.

Hobby, L. (2015). *Promoting potential: a mixed methods study evaluating the impact of differing school settings on high achieving students academic and psychosocial outcomes*. Australian Catholic University. Institute for Positive, Psychology and Education, degree granting institution.

Holman, D., Pavlica, K., & Thorpe, R. (1997). Rethinking Kolb's theory of experiential learning in management education: The contribution of social constructivism and activity theory. *Management Learning, 28*, 135-148.

Howitt, D., & Cramer, D. (2005). *Introduction to SPSS in Psychology For SPSS 10, 11, 12 and 13* (3rd ed.). Essex, England: Pearson Education Limited.

Hu, L., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Hughes, A., Galbraith, D., & White, D. (2011). Perceived Competence: A Common Core for Self-Efficacy and Self- Concept? *Journal of Personality Assessment, 93*(3), 278-289. doi:10.1080/00223891.2011.559390

Hunter, I. R., & Purcell, K. D. (1984). Program characteristics and success in a resocialization program for adjudicated delinquents. *Corrective & Social Psychiatry*, 25–35.

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

IBM Corp. (Released 2013). IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

Jang, H., Kim, E. J., ., & Reeve, J. (2012). Longitudinal test of self-determination theory's motivation mediation model in a naturally occurring classroom context. *Journal of Educational Psychology, 104*(4), 1175-1188. doi:10.1037/a0028089

Jang, H., Reeve, J., Ryan, R. M., & Kim, A. (2009). Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically oriented Korean students? *Journal of Educational Psychology, 101*(3), 644-661. doi:10.1037/a0014241

Janić, S. R., Jurak, G., Milanović, I., Lazarević, D., Kovač, M., & Novak, D. (2014). Physical Self-Concept of Adolescents in Western Balkan Countries: A Pilot Study. *Perceptual and Motor Skills, 119*(2), 629-649. doi:10.2466/08.PMS.119c23z7

Janssen, I., & Leblanc, A. G. (2010). Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioural Nutrition and Physical Activity, 7*, 40. doi:10.1186/1479-5868-7-40

Jöreskog, K. G., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models*. Cambridge, MA: Abt Books.

Kaplan, A., & Flum, H. (2012). Identity formation in educational settings: A critical focus for education in the 21st century. *Contemporary Educational Psychology, 37*(3), 171-175. doi:10.1016/j.cedpsych.2012.01.005

Karoly, L. A., & Panis, C. (2004). *21st Century at Work* Rand.

Kashdan, T. B., & Ciarrochi, J. (2013). *Mindfulness, acceptance and positive psychology: the seven foundations of well-being*: Oakland, CA: Context Press.

Kasser, T., Ryan, R. M., Zax, M., & Sameroff, A. J. (1995). The relations of maternal and social environments to late adolescents' materialistic and prosocial values. *Developmental Psychology, 31*, 907-914.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 593-602.

Kline, T. (2005). *Psychological testing a practical approach to design and evaluation*. Thousand Oaks, California: Thousand Oaks, California: Sage Publications.

Knoop, R. (1981). Age and correlates of locus of control. *Journal of psychology, 108*, 103-106.

Kolb, D. A. (1984). *Experiential learning: experience as the source of learning and development*. Englewood Cliffs, NJ; Prentice Hall.

La Guardia, J. G. (2009). Developing Who I Am: A Self-Determination Theory Approach to the Establishment of Healthy Identities. *Educational Psychologist, 44*(2), 90-104.

Lam, S., Cheng, R. W., & Ma, W. Y. K. (2009). Teacher and Student Intrinsic Motivation in Project-Based Learning. *Instructional Science: An International Journal of the Learning Sciences, 37*(6), 565-578. doi:10.1007/s11251-008-9070-9

Langer, A. I., Ulloa, V. G., Aguilar-Parra, J. M., Araya-Veliz, C., & Brito, G. (2016). Validation of a Spanish translation of the Gratitude Questionnaire (GQ-6) with a Chilean sample of adults and high schoolers. *Health and Quality of Life Outcomes, 14*, 53. doi:10.1186/s12955-016-0450-6

Larson, B. A. (2007). Adventure Camp Programs, Self-Concept, and Their Effects on Behavioral Problem Adolescents. *The Journal of Experiential Education, 29*(3), 313-330.

Lee, A., & Lam, E. (2014). EXPLORATORY STRUCTURAL EQUATION MODELING: A NEW TREND OF FACTOR ANALYSIS. *International Journal of Mathematics, Game Theory, and Algebra, 23*(3), 211-216.

Lefcourt, H. M. (1976). *Locus of Control*. Hillsdale, NJ: Erlbaum.

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and Extrinsic Motivational Orientations in the Classroom: Age Differences and Academic Correlates. *Journal of Educational Psychology, 97*(2), 184-196. doi:10.1037/0022-0663.97.2.184

Lerner, R. M., Kier, C. A., & Brown, J. (2002). *Adolescence: Development, diversity, context, and application*.

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. *The Journal of Experimental Education, 83*(1), 1-23. doi:10.1080/00220973.2013.876229

Little, T. D., Card, N. A., Slegers, D. W., & Ledford, E. C. (2007). Representing contextual effects in multiple-group MACS models. *Modeling contextual effects in longitudinal studies*, 121-147.

Little, T. D., Hawley, P. H., Henrich, C. C., & Marsland, K. (2002). Three views of the agentic self: A developmental synthesis. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 389-404): Rochester, NY: University of Rochester Press.

Louw, P. J., Meyer, C. D., Strydom, G. L., Kotze, H. N., & Ellis, S. (2012). The impact of an adventure based experiential learning programme on the life effectiveness of black high school learners. *African Journal for Physical, Health Education, Recreation & Dance, 18*(1), 55-64.

Lucas, R. E., & Donnellan, M. B. (2012). Estimating the Reliability of Single-Item Life Satisfaction Measures: Results from Four National Panel Studies. *Social Indicators Research, 105*(3), 323-331. doi:10.1007/s11205-011-9783-z

Makikangas, A., Feldt, T., Kinnunen, U., Tolvanen, A., Kinnunen, M.-L., & Pulkkinen, L. (2006). The Factor Structure and Factorial Invariance of the 12-Item General Health Questionnaire (GHQ-12) across Time: Evidence from Two Community-Based Samples. *Psychological Assessment, 18*(4), 444-451.

Marcoulides, K. M., & Yuan, K.-H. (2016). New Ways to Evaluate Goodness of Fit: A Note on Using Equivalence Testing to Assess Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(1), 148-153. doi:10.1080/10705511.2016.1225260

Marsh, H. W. (1988a). *Self-Description Questionnaire: A Theoretical and Empirical Basis for the Measurement of Multiple Dimensions of Preadolescent Self-Concept: A Test*

Manual and a Research Monograph. The Psychological Corporation: San Antonio, Texas.

Marsh, H. W. (1988b). *The Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of the late adolescent self-concept: An interim test manual and a research monograph.* San Antonio, TX: The Psychological Corporation.

Marsh, H. W. (1989). Age and Sex Effects in Multiple Dimensions of Self-Concept: Preadolescence to Early Adulthood. *Journal of Educational Psychology, 81*(3), 417-430. doi:10.1037/0022-0663.81.3.417

Marsh, H. W. (1990a). 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. (1990b). *Self-Description Questionnaire - II manual*. Sydney: University of Western Sydney.

Marsh, H. W. (1994). Confirmatory factor analysis models of factorial invariance: A multifaceted approach. *Structural Equation Modeling: A Multidisciplinary Journal, 1*(1), 5-34. doi:10.1080/10705519409539960

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical processes. *Advanced structural equation modeling techniques* 315-353.

Marsh, H. W., Byrne, B. M., & Shavelson, R. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology, 80*, 366-380.

Marsh, H. W., Chanal, J. P., & Sarrazin, P. G. (2006). Self-belief does make a difference: A reciprocal effects model of the causal ordering of physical self-concept and gymnastics performance. *Journal of Sports Sciences, 24*(1), 101-111. doi:10.1080/02640410500130920

Marsh, H. W., Craven, R. G., & Debus, R. (1998). Structure, Stability, and Development of Young Children's Self-Concepts: A Multicohort–Multioccasion Study. *Child Development, 69*(4), 1030-1053. doi:10.1111/j.1467-8624.1998.tb06159.x

Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G. E., & 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., 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., & Hau, K. T. (1996). Assessing goodness of fit: Is parsimony always desirable? *The Journal of Experimental Education, 64*(4), 364-390.

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* (pp. 225-340). Mahwah, NJ: Lawrence Erlbaum Associates.

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., & Martin, A. J. (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., & Hau, K. (2006). *A Multimethod Perspective on Self-Concept Research in Educational Psychology: A Construct Validity Approach*: American Psychological Association.

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., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). 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., 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., Parada, R. H., & Ayotte, V. (2004). A Multidimensional Perspective of Relations Between Self-Concept (Self Description Questionnaire II) and

Adolescent Mental Health (Youth Self-Report). *Psychological Assessment, 16*, 27-41.

Marsh, H. W., & Richards, G. E. (1988). The Outward Bound Bridging Course for low achieving high-school males: Effect on academic achievement and multidimensional self-concepts. *Australian Journal of Psychiatry, 40*, 281-298.

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*, 475-492. doi:10.1177/0146167286124011

Marsh, H. W., Richards, G. E., & Barnes, J. (1986b). Multidimensional self-concepts: A long term follow-up of the effect of participation in an Outward Bound program. *Personality and Social Psychology Bulletin, 12*, 475-492.

Marsh, H. W., Richards, G. E., & Barnes, J. (1986c). Multidimensional self-concepts: The effect of participation in an Outward Bound program. *Personality and Social Psychology Bulletin, 45*, 173-187.

Marsh, H. W., Smith, I. D., Barnes, J., & Butler, S. (1983). Self-concept: Reliability, dimensionality, validity, and the measurement of change. *Journal of Educational Psychology, 75*, 772-790.

Martin, A. J. (2007). Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology, 77*, 413-440.

Martin, A. J. (2009). Motivation and Engagement Across the Academic Life Span (Vol. 69, pp. 794-824).

Martin, A. J., & Marsh, H. W. (2006). Academic Resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools, 43*, 267-281. doi:10.1002/pits

McCullough, M. E., Emmons, R. A., & Tsang, J.-A. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology, 82*(1), 112-127. doi:10.1037//0022-3514.82.1.112

McDonald, R. P. (1999). *Test Theory: A Unified Treatment* (Vol. 95). Mahwah, NJ: Lawrence Erlbaum.

McLeod, B., & Allen-Craig, S. (2004). *An evaluation of an Experiential Learning and Outdoor Education school program on the Life Effectiveness skills of middle*

school boys. Paper presented at the International Outdoor Education Research Conference, Latrobe University Bendigo, Victoria, July 6-9.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525-543. doi:10.1007/BF02294825

Mission Australia. (2018). 2018 Youth Survey Report. Retrieved from https://www.missionaustralia.com.au/what-we-do/research-impact-policy-advocacy/youth-survey

Molina, J. G., Rodrigo, M. F., Losilla, J.-M., & Vives, J. (2014). Wording effects and the factor structure of the 12-item General Health Questionnaire (GHQ-12). *Psychological Assessment, 26*(3), 1031-1037.

Molloy, L. E., Ram, N., & Gest, S. D. (2011). The storm and stress (or calm) of early adolescent self-concepts: within- and between-subjects variability. *Developmental Psychology, 47*(6), 1589-1607. doi:10.1037/a0025413

Mora, P. A., Berkowitz, A., Contrada, R. J., Wisnivesky, J., Horne, R., Leventhal, H., & Halm, E. A. (2011). Factor structure and longitudinal invariance of the Medical Adherence Report Scale-Asthma. *Psychology & health, 26*(6), 713-727. doi:10.1080/08870446.2010.490585

Mouratidis, A., Vansteenkiste, M., Lens, W., Michou, A., & Soenens, B. (2013). Within-Person Configurations and Temporal Relations of Personal and Perceived Parent-Promoted Aspirations to School Correlates Among Adolescents. *Journal of Educational Psychology, 105*(3), 895-910. doi:10.1037/a0032838

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide. Eighth Edition*. Los Angeles CA: Muthén & Muthén.

Nader, P. R., Bradley, R. H., Houts, R. M., McRitchie, S. L., & O'Brien, M. (2008). Moderate-to-vigorous physical activity from ages 9 to 15 years. *JAMA, 300*, 295-305.

National Health and Medical Research Council. (2013). Australian Dietary Guidelines. Retrieved from https://www.eatforhealth.gov.au/sites/default/files/files/the_guidelines/n55_australian_dietary_guidelines.pdf

Neff, K. D., & McGehee, P. (2010). Self-compassion and Psychological Resilience Among Adolescents and Young Adults. *Self and Identity, 9*(3), 225-240. doi:10.1080/15298860902979307

Neill, J. T. (2002). *Meta-analytic research on the outcomes of outdoor education*. Paper presented at the 6th Biennial Coalition for Education in the Outdoors Research Symposium, Bradford Woods, IN.

Neill, J. T. (2003). Reviewing and Benchmarking Adventure Therapy Outcomes: Applications of Meta-Analysis. *Journal of Experiential Education, 25*(3), 316-321. doi:10.1177/105382590302500305

Neill, J. T. (2008). Enhancing Life Effectiveness: The Impacts of Outdoor Educ Programs. *PhD Thesis. Sydney: Univ of Western Sydney*.

Neill, J. T., & Heubeck, B. G. (1998). *Adolescent Coping Styles and Outdoor Education: Searching for the Mechanisms of Change*. Paper presented to the 1st International Adventure Therapy Conference, University of Western Australia, Perth, W.A., Australia, July 1-5, 1997.

Neill, J. T., Marsh, H. W., & Richards, G. E. (1997a). *Development and Psychometrics of the Life Effectiveness Questionnaire*. Sydney, Australia: University of Western Sydney.

Neill, J. T., Marsh, H. W., & Richards, G. E. (1997b). The Life Effectiveness Questionnaire: Development and Psychometrics. Sydney, NSW, Australia: University of Western Sydney.

Nix, G., Ryan, R. M., Manley, J. B., & Deci, E. L. (1999). Revitalization through self-regulation: The effects of autonomous and controlled motivation on happiness and vitality. *Journal of Experimental Social Psychology, 35*, 266-284.

Nunn, G. D. (1987). Concurrent validity between children's locus of control and attitudes toward home, school, and peers. *Educational and Psychological Measurement, 47*, 1087-1089.

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, 4*, 524-533.

O'Brien, L., Burls, A., Bentsen, P., Hilmo, I., Holter, K., Haberling, D., . . . McLoughlin, J. (2011). Outdoor Education, Life Long Learning and Skills Development in Woodlands and Green Spaces: The Potential Links to Health and Well-Being *Forests, Trees and Human Health* (pp. 343-372): Springer Netherlands.

Outdoor Education Australia. (2018). Scope and Sequence of Outdoor Education in Australian Schools. Retrieved from https://outdooreducationaustralia.org.au/education/sequencing/

Outward Bound: Inspiring Australians. (2016). Classic School Programs. Retrieved from http://www.outwardbound.org.au/course-category/school/

Özmen, O., & Sümer, Z. H. (2011). Predictors of risk-taking behaviors among Turkish adolescents. *Personality and Individual Differences*, *50*(1), 4-9.

Pajares, F., & Schunk, D. H. (2005). Self-efficacy and self-concept beliefs: Jointly contributing to the quality of human life. In M. H. Marsh, R. Craven, & D. McInerney (Eds.), *International advances in self research* (Vol. 2, pp. 95–121). Greenwich, CT: Information Age Publishing.

Paulhus, D. L. (1991). CHAPTER 2 - Measurement and Control of Response Bias. In J. Robinson., P. Shaver., & L. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (pp. 17-59): New York: Academic Press.

Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nature Reviews Neuroscience, 9*, 947-957. doi:10.1038/nrn2513

Pelletier, L. G., Dion, S., Tuson, K., & Green-Demers, I. (1999). Why do people fail to adopt environmental protective behaviors? Toward a taxonomy of environmental amotivation. *Journal of Applied Social Psychology, 29*(12), 2481-2504. doi:10.1111/j.1559-1816.1999.tb00122.x

Petersen, C., Maier, S. F., & Seligman, M. E. P. (1995). *Learned helplessness: A theory for the age of personal control*. New York, NY: Oxford University Press.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879-891. doi:10.3758/brm.40.3.879

Priest, S. (1997). *Effective leadership in adventure programming*. Champaign, IL: Champaign, IL : Human Kinetics.

Priest, S., & Gass, M. A. (1998). *Effective leadership in adventure programming*. Human Kinetics: Champaign, IL: Human Kinetics: Champaign, IL.

Rahman, H. A. (2009). *The Effect of Outdoor Education on Students' Social Emotional Competencies Development and Engagement in Classroom*: La Trobe University, Beechworth, Victoria, Australia.

Rankin, J. L., Lane, D. J., Gibbons, F. X., & Gerrard, M. (2004). Adolescent Self-Consciousness: Longitudinal Age Changes and Gender Differences in Two Cohorts. *Journal of Research on Adolescence, 14*(1), 1-21. doi:10.1111/j.1532-7795.2004.01401001.x

Reschly, A. L., Betts, J., & Appleton, J. J. (2014). An examination of the validity of two measures of student engagement. *International Journal of School & Educational Psychology, 2*(2), 106-114. doi:10.1080/21683603.2013.876950

Revelle, W., & Zinbarg, R. E. (2009). Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika, 74*(1), 145-154. doi:10.1007/s11336-008-9102-z

Richards, G. E. (1977). *Some Educational Implications & Contributions of Outward Bound*: G. E. Richards.

Richards, G. E., Ellis, L. A., & Neill, J. T. (2002, 6-8 August). *The ROPELOC: Review of Personal Effectiveness and Locus of Control: A comprehensive instrument for reviewing life effectiveness.* Paper presented at the Self-Concept Research: Driving International Research Agendas, Sydney.

Richards, G. E., & Neill, J. T. (2000a). *Review of personal effectiveness scale*. Canberra, Australia: National Outdoor Education and Leadership Services.

Richards, G. E., & Neill, J. T. (2000b). *Review of personal effectiveness scale*. Canberra, Australia: National Outdoor Education and Leadership Services.

Richards, K., Carpenter, C., & Harper, N. (2011). Looking at the landscape of adventure therapy: making links to theory and practice. *Journal of Adventure Education & Outdoor Learning, 11*(2), 83-90. doi:10.1080/14729679.2011.632877

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

Rickinson, M., Dillon, J., Teamey, K., Morris, M., Choi, M. Y., Sanders, D., & Benefield, P. (2004). *A review of research on outdoor learning*. Preston Montford, England: Field Studies Council.

Ringdal, R., Bradley Eilertsen, M.-E., BjГёrnsen, H. N., Espnes, G. A., & Moksnes, U. K. (2018). Validation of two versions of the Warwick-Edinburgh Mental Well-Being Scale among Norwegian adolescents. *Scandinavian journal of public health, 46*, 718-725. doi:10.1177/1403494817735391

Ritchie, S. D., Wabano, M. J., Russell, K., Enosse, L., & Young, N. L. (2014). Promoting resilience and wellbeing through an outdoor intervention designed for Aboriginal adolescents. *Rural and remote health, 14*, 2523.

Rodríguez-Meirinhos, A., Antolín-Suárez, L., Brenning, K., Vansteenkiste, M., & Oliva, A. (2019). A Bright and a Dark Path to Adolescents' Functioning: The Role of Need Satisfaction and Need Frustration Across Gender, Age, and Socioeconomic Status. *Journal of Happiness Studies*, 1-22. doi:10.1007/s10902-018-00072-9

Rotter, J. B. (1954). *Social learning and clinical psychology*: Prentice-Hall, Inc.

Rowe, K. J., & Rowe, K. S. (1999). Investigating the relationship between students' attentive– inattentive behaviors in the classroom and their literacy progress: Chapter 2 Purpose, data and analytic methods for the investigation. *International Journal of Educational Research, 31*(1), 17-30. doi:10.1016/S0883-0355(98)00069-X

Ryan, R. M., & Deci, E. L. (2000a). The darker and brighter sides of human existence: Basic psychological needs as a unifying concept. *Psychological Enquiry, 11*, 319-338. doi:10.1207/S15327965PLI1104_03

Ryan, R. M., & Deci, E. L. (2000b). Intrinsic and Extrinsic Motivation: Classic Definitions and New Directions. *Contemporary Educational Psychology, 25*, 54–67. doi:doi:10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2000c). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist, 55(1)*, 68-78. doi:00000487-200001000-00007

Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of Self-Determination Research* (pp. 3 – 33): Rochester, NY: University of Rochester Press.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: basic psychological needs in motivation, development, and wellness*: New York, New York London, England : The Guilford Press.

Ryan, R. M., Deci, E. L., & Grolnick, W. S. (1995). Autonomy, relatedness, and the self: Their relation to development and psychopathology. *Developmental psychopathology, 1*, 618-655.

Santrock, J. W. (2003). *Adolescence (9th ed.)*: Boston: McGraw-Hill.

Savahl, S., Tiliouine, H., Casas, F., Adams, S., Mekonen, Y., Dejene, N., . . . Witten, H. (2017). Children's subjective well-being in Africa: A comparative analysis across three countries. *Children and Youth Services Review, 80*, 31-40. doi:10.1016/j.childyouth.2017.06.063

Scarf, D., Hayhurst, J. G., Riordan, B. C., Boyes, M., Ruffman, T., & Hunter, J. A. (2017). Increasing resilience in adolescents: the importance of social connectedness in adventure education programmes. *Australasian Psychiatry, 25*(2), 154-156. doi:10.1177/1039856216671668

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 Self-Esteem Following an Adventure Education Programme. *Journal of Adventure Education and Outdoor Learning, 18*(2), 165-175. doi:10.1080/14729679.2017.1385496

Schmidt, M., Valkanover, S., & Conzelmann, A. (2013). Veridicality of Self-Concept of Strength in Male Adolescents. *Perceptual and Motor Skills, 116*(3), 1029-1042. doi:10.2466/06.10.pms.116.3.1029-1042

Schumacker, R. E. (2010). *A Beginner's guide to structural equation modeling* (3rd ed.). New York: New York : Routledge.

Schunk, D. H., & Pajares, F. (2005). Competence Perceptions and Academic Functioning. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of Competence and Motivation* (pp. 85-104). New York: London: Guilford Publications.

Scrutton, R. A. (2014). Outdoor adventure education for children in Scotland: quantifying the benefits. *Journal of Adventure Education & Outdoor Learning, 15*(2), 1-15. doi:10.1080/14729679.2013.867813

Scrutton, R. A. (2015). Outdoor Adventure Education for Children in Scotland: Quantifying the Benefits. *Journal of Adventure Education and Outdoor Learning, 15*(2), 123-137. doi:10.1080/14729679.2013.867813

Scrutton, R. A., & 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

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Validation of construct interpretations. *Review of Educational Research, 46*, 407-441.

Shehu, J., & Mokgwathi, M. M. (2008). Health Locus of Control and Internal Resilience Factors among Adolescents in Botswana: A Case-control Study with Implications for Physical Education. *South African Journal for Research in Sport, Physical Education & Recreation (SAJR SPER), 30*(2), 95-105.

Sheldon, K. M. (1995). Creativity and self-determination in personality. *Creativity Research Journal, 8*, 25-36.

Shevlin, M., & Adamson, G. (2005). Alternative factor models and factorial invariance of the GHQ-12: a large sample analysis using confirmatory factor analysis. *Psychological Assessment, 17*(2), 231-236. doi:10.1037/1040-3590.17.2.231

Shih, H.-P. (2006). Assessing the effects of self-efficacy and competence on individual satisfaction with computer use: an IT student perspective. *Computers in Human Behavior, 22*, 1012-1026. doi:10.1016/j.chb.2004.03.025

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., 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., & Morgan, C. (2011). Adventure-Based Programming: Exemplary Youth Development Practice. *New Directions for Youth Development, 2011*, 105-119. doi:10.1002/yd.400

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., Paisley, K., Gookin, J., & Ward, P. (2007). Addressing response-shift bias: Retrospective pretests in recreation research and evaluation. *Journal of Leisure Research, 39*, 295-315.

Simon, P. (2016). An exploratory study about the buffering effect of gratitude on the relationship between neuroticism and anhedonic depression. *Psychological Studies, 61*(4), 340-347. doi:10.1007/s12646-016-0377-3

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience & Biobehavioral Reviews, 24*(4), 417-463.

Steiger, J. H. (1989). *EzPATH: Causal modeling*. Evanston, IL: SYSTAT.

Stewart-Brown, S., Platt, S., Tennant, A., Maheswaran, H., Parkinson, J., Weich, S., . . . Clarke, A. (2011). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a valid and reliable tool for measuring mental well-being in diverse

populations and projects. *Journal of Epidemiology and Community Health, 65*, A38-A39. doi:10.1136/jech.2011.143586.86

Stewart-Brown, S., Tennant, A., Tennant, R., 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*, 15.

Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology, 5*, 1-25. doi:10.1146/annurev.clinpsy.032408.153639

Strickland, B. R. (1989). Internal-external control expectancies: From contingency to creativity. *American Psychologist, 44*, 1-12.

Sundjoto, S. (2017). The Role of Internal Locus of Control on Intrinsic Motivation and Employee Performance of Ceramic Company in East Java. *Journal of Business and Management, 19*(7), 29-35. doi:10.9790/487X-1907072935

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 Qual Life Outcomes, 5*, 63. doi:10.1186/1477-7525-5-63

The Scots College. (2016). *Glengarry Parent Handbook: General Information*. Retrieved from the Scots College Glengarry website: https://glengarry.tsc.nsw.edu.au/

Thomaes, S., Sedikides, C., van den Bos, N., Hutteman, R., & Reijntjes, A. (2017). Happy To Be "Me?" Authenticity, Psychological Need Satisfaction, and Subjective Well-Being in Adolescence. *Child Development, 88*(4), 1045-1056. doi:10.1111/cdev.12867

Thomson, C. J., & Burr, J. F. (2015). Take A Hike! The Impact of Adventure-based Learning on Psychological Wellbeing in Troubled Youth. *Medicine & Science in Sports & Exercise, 47*, 512. doi:10.1249/01.mss.0000477840.11719.1f

Tucker, L., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1-10. doi:10.1007/BF02291170

Usher, E. L., & Pajares, F. (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary Educational Psychology, 31*(2), 125-141. doi:10.1016/j.cedpsych.2005.03.002

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relations between self-beliefs and academic achievement: A systematic review. *Educational Psychologist, 39*, 111-133.

Vlamis, E., Bell, B. J., & Gass, M. A. (2011). Effects of a College Adventure Orientation Program on Student Development Behaviors. *Journal of Experiential Education, 34*(2), 127-148. doi:10.5193/JEE34.2.127

Vörösvári, E. (2016). Effects of an Outdoor Adventure Camp Experience on Adults with Brain Injury. *Palaestra, 30*(2), n/a.

Wang, C. K. J., Ang, R. P., Teo-Koh, S. M., & Kahlid, A. (2004). Motivational predictors of young adolescents' participation in an outdoor adventure course: A self-determination theory approach. *Journal of Adventure Education and Outdoor Learning, 4*(1), 57-65. doi:10.1080/14729670485200421

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th-through 11th-grade Australian students. *Child Development, 75*, 1556-1574. doi:10.1111/j.1467-8624.2004.00757.x

Werneke, U., Goldberg, D. P., Yalcin, I., & Ustun, B. T. (2000). The stability of the factor structure of the General Health Questionnaire. *Psychological Medicine, 30*, 823-829.

Williams, J., & Mackinnon, D. P. (2008). Resampling and Distribution of the Product Methods for Testing Indirect Effects in Complex Models. *Structural Equation Modeling: A Multidisciplinary Journal, 15*(1), 23-51. doi:10.1080/10705510701758166

Williams, R. (2013). Woven into the Fabric of Experience: Residential Adventure Education and Complexity. *Journal of Adventure Education and Outdoor Learning, 13*(2), 107-124. doi:10.1080/14729679.2012.731725

Wilson, C., & Sibthorp, J. (2017). The Developmental Importance of Success and Failure Experiences at Summer Camp. *Journal of Youth Development, 12*(3), 18-36. doi:10.5195/jyd.2017.502

Wilson, S. L., & Lipsey, M. W. (2000). Wilderness challenge programs for delinquent youth: a meta-analysis of outcome evaluations. *Evaluation and Program Planning, 23*(1), 1-12.

Wolf, F. M. (1986). *Meta-analysis quantitative methods for research synthesis*. Beverly Hills, London: SAGE.

Wood, A. M., Maltby, J., Gillett, R., Linley, P. A., & Joseph, S. (2008). The role of gratitude in the development of social support, stress, and depression: Two longitudinal studies. *Journal of Research in Personality, 42*(4), 854-871. doi:10.1016/j.jrp.2007.11.003

Wood, P. (2008). Confirmatory Factor Analysis for Applied Research. *62*, 91-92. doi:10.1198/tas.2008.s98

World Health Organisation. (2016). Global Strategy on Diet, Physical Activity and Health. Retrieved from http://www.who.int/dietphysicalactivity/factsheet_inactivity/en/

Ye, S. (2009). Factor structure of the General Health Questionnaire (GHQ-12): The role of wording effects. *Personality and Individual Differences, 46*(2), 197-201. doi:10.1016/j.paid.2008.09.027

Yeung, A. S. (2011). Student self-concept and effort: gender and grade differences. *Educational Psychology, 31*(6), 749-772. doi:10.1080/01443410.2011.608487

Yeung, A. S., & McInerney, D. (2005). Students' School Motivation and Aspiration Over High School Years. *Educational Psychology, 25*(5), 537-554. doi:10.1080/01443410500046804

Yin, P., & Fan, X. (2003). Assessing The Factor Structure Invariance Of Self-Concept Measurement Across Ethnic And Gender Groups: Findings From A National Sample. *Educational and Psychological Measurement, 63*(2), 296-318. doi:10.1177/0013164403251328

Zinbarg, R. E., Revelle, W., & Yovel, I. (2007). Estimating ωh for Structures Containing Two Group Factors: Perils and Prospects. *Applied Psychological Measurement, 31*(2), 135-157. doi:10.1177/0146621606291558

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating Generalizability to a Latent Variable Common to All of a Scale's Indicators: A Comparison of Estimators for ωh. *Applied Psychological Measurement, 30*(2), 121-144. doi:10.1177/0146621605278814

Zink, R. (2010). Coming to know oneself through experiential education. *Discourse: Studies in the Cultural Politics of Education, 31*(2), 209-219. doi:10.1080/01596301003679727

APPENDICES

Appendix 1: Evidence of Human Research Ethics Committee Approval

Dear Applicant,

Principal Investigator: Prof Herbert Marsh, Dr Keong Yap
Student Researcher: Mathew Pfeiffer
Ethics Register Number: 2015-245H
Project Title: Enhancing Psychological Outcomes in Adolescence: The Effects of Outdoor Adventure Education
Risk Level: Low Risk Date
Approved: 10/12/2015
Ethics Clearance End Date: 1/11/2021

This email is to advise that your application has been reviewed by the Australian Catholic University's Human Research Ethics Committee and confirmed as meeting the requirements of the National Statement on Ethical Conduct in Human Research.

The data collection of your project has received ethical clearance but the decision and authority to commence may be dependent on factors beyond the remit of the ethics review process and approval is subject to ratification at the next available Committee meeting. The Chief Investigator is responsible for ensuring that outstanding permission letters are obtained, interview/survey questions, if relevant, and a copy forwarded to ACU HREC before any data collection can occur. Failure to provide outstanding documents to the ACU HREC before data collection commences is in breach of the National Statement on Ethical Conduct in Human Research and the Australian Code for the Responsible Conduct of Research. Further, this approval is only valid as long as approved procedures are followed.

If your project is a Clinical Trial, you are required to register it in a publicly accessible trials registry prior to enrolment of the first participant (e.g. Australian New Zealand Clinical Trials Registry http://www.anzctr.org.au/) as a condition of ethics approval.

If you require a formal approval certificate, please respond via reply email and one will be issued.

Researchers who fail to submit a progress report may have their ethical clearance revoked and/or the ethical clearances of other projects suspended. When your project has been completed a progress/final report form must be submitted. The information researchers provide on the security of records, compliance with approval consent procedures and documentation and responses to special conditions is reported to the NHMRC on an annual basis. In accordance with NHMRC the ACU HREC may undertake annual audits of any projects considered to be of more than low risk.

It is the Principal Investigators / Supervisors responsibility to ensure that:
1. All serious and unexpected adverse events should be reported to the HREC with 72 hours.
2. Any changes to the protocol must be reviewed by the HREC by submitting a Modification/Change to Protocol Form prior to the research commencing or continuing. http://research.acu.edu.au/researcher-support/integrity-and-ethics/
3. Progress reports are to be submitted on an annual basis. http://research.acu.edu.au/researcher-support/integrity-and-ethics/
4. All research participants are to be provided with a Participant Information Letter and consent form, unless otherwise agreed by the Committee.
5. Protocols can be extended for a maximum of five (5) years after which a new application must be submitted. (The five year limit on renewal of approvals allows the Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

Researchers must immediately report to HREC any matter that might affect the ethical acceptability of the protocol eg: changes to protocols or unforeseen circumstances or adverse effects on participants.

Please do not hesitate to contact the office if you have any queries.

Kind regards,

Kylie Pashley
on behalf of ACU HREC Chair, Dr Nadia Crittenden
Ethics Officer | Research Services
Office of the Deputy Vice Chancellor (Research) Australian Catholic University

Appendix 2: The Research Survey Instrument

Reference: Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS; Stewart-Brown et al., 2009)

Below are some statements about feelings and thoughts. Please circle the number that best describes your experience of each **over the last 2 weeks.**

		None of the time 1	**Rarely 2**	**Some of the time 3**	**Often 4**	**All of the time 5**
1	I've been feeling optimistic about the future	1	2	3	4	5
2	I've been feeling useful	1	2	3	4	5
3	I've been feeling relaxed	1	2	3	4	5
4	I've been dealing with problems well	1	2	3	4	5
5	I've been thinking clearly	1	2	3	4	5
6	I've been feeling close to other people	1	2	3	4	5
7	I've been able to make up my own mind about things	1	2	3	4	5

Reference: Overall Satisfaction with Life (OSL; PWI-SC, Cummins & Lau, 2005)

Please indicate your answer to the following statement on the scale below.
How happy are you with your life as a whole?

Very Sad				**Not Happy or Sad**					**Very Happy**
1	2	3	4	5	6	7	8	9	10

Reference: GHQ-12 (Goldberg, 1972)

Please indicate your answer to the following statements by marking a response

1. Have you recently been able to concentrate on whatever you're doing?

   O Better than usual
   O Same as usual
   O Less than usual
   O Much less than usual

2. Have you recently lost much sleep over worry?

   O Not at all
   O No more than usual
   O Rather more than usual
   O Much more than usual

3. Have you recently felt that you are playing a useful part in things?

   O More so than usual
   O Same as usual
   O Less useful than usual
   O Much less than usual

4. Have you recently felt capable of making decisions about things?

   O More so than usual
   O Same as usual
   O Less so than usual
   O Much less than usual

5. Have you recently felt constantly under strain?

O Not at all
O No more than usual
O Rather more than usual
O Much more than usual

6. Have you recently felt you couldn't overcome your difficulties?

O Not at all
O No more than usual
O Rather more than usual
O Much more than usual

7. Have you recently been able to enjoy your normal day-to-day activities?

O More so than usual
O Same as usual
O Less so than usual
O Much less than usual

8. Have you recently been able to face up to your problems?

O More so than usual
O Same as usual
O Less so than usual
O Much less than usual

9. Have you recently been feeling unhappy and depressed?

O Not at all
O No more than usual
O Rather more than usual
O Much more than usual

10. Have you recently been losing confidence in yourself?

- O Not at all
- O No more than usual
- O Rather more than usual
- O Much more than usual

11. Have you recently been thinking of yourself as a worthless person?

- O Not at all
- O No more than usual
- O Rather more than usual
- O Much more than usual

12. Have you recently been feeling reasonably happy, all things considered?

- O More so than usual
- O Same as usual
- O Less so than usual
- O Much less than usual

Reference: Review of Personal Effectiveness with Locus of Control (ROPELOC) scale (Richards et al., 2002)

PLEASE READ THESE INSTRUCTIONS FIRST

This is a chance for you to look at how you think and feel about yourself. It is important that you report **how you feel NOW** (not how you felt at another time in your life, or how you might feel tomorrow).

Use the eight point scale to indicate how true (like you) or how false (unlike you), each statement is as a description of you.

FALSE NOT LIKE ME							**TRUE LIKE ME**
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
This statement doesn't describe me at all; it isn't like me at all		More false than true			More true than false		This statement describes me very well; it is very much like me

STATEMENT		**FALSE**							**TRUE**
1	I like cooperating in a team	1	2	3	4	5	6	7	8
2	No matter what the situation is I can handle it	1	2	3	4	5	6	7	8
3	I can be a good leader	1	2	3	4	5	6	7	8
4	My own efforts and actions are what will determine my future	1	2	3	4	5	6	7	8
5	I prefer to be actively involved in things	1	2	3	4	5	6	7	8
6	I am open to different thinking if there is a better idea	1	2	3	4	5	6	7	8
7	In everything I do I try my best to get the details right	1	2	3	4	5	6	7	8
8	Luck, other people and events control most of my life	1	2	3	4	5	6	7	8
9	I am confident that I have the ability to succeed in anything I want to do	1	2	3	4	5	6	7	8
10	I am effective in social situations	1	2	3	4	5	6	7	8
11	I am calm in stressful situations	1	2	3	4	5	6	7	8
12	My overall effectiveness in life is very high	1	2	3	4	5	6	7	8
13	I plan and use my time efficiently	1	2	3	4	5	6	7	8
14	I cope well with changing situations	1	2	3	4	5	6	7	8
15	I cooperate well when working in a team	1	2	3	4	5	6	7	8

16	No matter what happens I can handle it	1	2	3	4	5	6	7	8
17	I am capable of being a good leader	1	2	3	4	5	6	7	8
18	I like being active and energetic	1	2	3	4	5	6	7	8
19	What I do and how I do it will determine my successes in life	1	2	3	4	5	6	7	8
20	I am open to new thoughts and ideas	1	2	3	4	5	6	7	8

STATEMENT		**FALSE**							**TRUE**
21	I try to get the best possible results when I do things	1	2	3	4	5	6	7	8
22	When I apply myself to something I am confident I will succeed	1	2	3	4	5	6	7	8
23	My future is mostly in the hands of other people	1	2	3	4	5	6	7	8
24	I am competent and effective in social situations	1	2	3	4	5	6	7	8
25	I can stay calm and overcome anxiety in almost all situations	1	2	3	4	5	6	7	8
26	I am efficient and do not waste time	1	2	3	4	5	6	7	8
27	Overall, in all things in life, I am effective	1	2	3	4	5	6	7	8
28	When things around me change I cope well	1	2	3	4	5	6	7	8
29	I am good at cooperating with team members	1	2	3	4	5	6	7	8
30	I can handle things no matter what happens	1	2	3	4	5	6	7	8
31	I am seen as a capable leader	1	2	3	4	5	6	7	8
32	I like to get into things and make action	1	2	3	4	5	6	7	8
33	I can adapt my thinking and ideas	1	2	3	4	5	6	7	8
34	If I succeed in life it will be because of my efforts	1	2	3	4	5	6	7	8
35	I try to get the very best results in everything I do	1	2	3	4	5	6	7	8
36	I am confident in my ability to be successful	1	2	3	4	5	6	7	8
37	I communicate effectively in social situations	1	2	3	4	5	6	7	8
38	My life is mostly controlled by external things	1	2	3	4	5	6	7	8
39	I am calm when things go wrong	1	2	3	4	5	6	7	8
40	I am efficient in the way I use my time	1	2	3	4	5	6	7	8
41	I cope well when things change	1	2	3	4	5	6	7	8
42	Overall, in my life I am a very effective person	1	2	3	4	5	6	7	8

Reference: Motivation and Engagement Scale (MES; Martin, 2007, 2009)

Please read each sentence and choose an answer. There are six possible answers for each question - "Disagree a lot", "Agree a lot", and four answers in between. DO NOT say your answer out loud or talk about them with anyone else.

		Disagree a lot 1	**Disagree 2**	**Disagree a little 3**	**Agree a little 4**	**Agree 5**	**Agree a lot 6**
1	I believe I can do a good job in my schoolwork	1	2	3	4	5	6
2	What I learn in my schoolwork is important and useful	1	2	3	4	5	6
3	In my schoolwork, I am focused on learning and improving more than competing and being the best	1	2	3	4	5	6
4	I plan out how I will do my schoolwork and study	1	2	3	4	5	6
5	I use my study/homework time well and try to study and do homework under conditions that bring out my best	1	2	3	4	5	6
6	I persist at schoolwork even when it is challenging or difficult	1	2	3	4	5	6
7	I get quite anxious about schoolwork and tests	1	2	3	4	5	6
8	I mainly do my schoolwork to avoid failing or disapproval from parents or the teacher/s	1	2	3	4	5	6
9	I don't think I have much control over how well I do in my schoolwork	1	2	3	4	5	6
10	In my schoolwork I sometimes reduce my chances of doing well (e.g. waste time, not study, disrupt others, procrastinate)	1	2	3	4	5	6
11	I often feel like giving up in my schoolwork	1	2	3	4	5	6

Reference: Basic Psychological Needs Satisfaction and Frustration Scale (BPNSFS; Chen et al., 2015)

The following questions concern your feelings about **your last term at school.**

Please indicate how much you agree with each of the following statements given your experiences at school. Remember that your teachers will never know how you responded to the questions. Use the following scale in responding to the items.

		Disagree a lot 1	**Disagree 2**	**Disagree a little 3**	**Agree a little 4**	**Agree 5**	**Agree a lot 6**
1	At school, I feel a sense of choice and freedom in the things I undertake	1	2	3	4	5	6
2	I feel that my decisions at school reflect what I really want	1	2	3	4	5	6
3	I feel my choices at school express who I really am	1	2	3	4	5	6
4	I feel I have been doing what really interests me at school	1	2	3	4	5	6
5	I feel that the people I care at school about also care about me	1	2	3	4	5	6
6	I feel connected with people who care for me at school, and for whom I care at school	1	2	3	4	5	6
7	At school, I feel close and connected with other people who are important to me	1	2	3	4	5	6
8	I experience a warm feeling with the people I spend time with at school	1	2	3	4	5	6
9	I feel confident that I can do things well at school	1	2	3	4	5	6
10	At school, I feel capable at what I do	1	2	3	4	5	6
11	When I am at school, I feel competent to achieve my goals	1	2	3	4	5	6
12	At school, I feel I can successfully complete difficult tasks	1	2	3	4	5	6

Reference: Self-Description Questionnaire II/III (SDQ-II/III; Marsh, 1988b, 1990b)

Please select a number beside each statement to indicate how much you agree with it.

		Disagree a lot 1	**Disagree 2**	**Disagree a little 3**	**Agree a little 4**	**Agree 5**	**Agree a lot 6**
1	I learn things quickly in most school subjects	1	2	3	4	5	6
2	I have always done well in most school subjects	1	2	3	4	5	6
3	Compared to others my age I am good at most school subjects	1	2	3	4	5	6
4	Work in most school subjects is easy for me	1	2	3	4	5	6
5	I get good marks in most school subjects	1	2	3	4	5	6
1	I can run fast	1	2	3	4	5	6
2	I enjoy things like sports, gym, and dance	1	2	3	4	5	6
3	I am good at things like sport, gym, and dance	1	2	3	4	5	6
4	I am better than most of my friends at things like sports, gym, and dance	1	2	3	4	5	6
5	I can run a long way without stopping	1	2	3	4	5	6
1	Overall I have a lot to be proud of	1	2	3	4	5	6
2	I can do things as well as most people	1	2	3	4	5	6
3	Most things I do, I do well	1	2	3	4	5	6
4	Overall, most things I do turn out well	1	2	3	4	5	6
5	I do lots of important things	1	2	3	4	5	6
6	A lot of things about me are good	1	2	3	4	5	6
7	When I do something, I do it well	1	2	3	4	5	6
1	I get along well with my parents	1	2	3	4	5	6
2	My parents treat me fairly	1	2	3	4	5	6
3	My parents understand me	1	2	3	4	5	6
4	I do not like my parents very much	1	2	3	4	5	6
1	I can figure out unusual answers to new problems easily	1	2	3	4	5	6
2	I come up with new ideas all the time	1	2	3	4	5	6
3	I enjoy working out new ways of solving problems	1	2	3	4	5	6
4	I can often see better ways of doing things	1	2	3	4	5	6
5	I can often see better ways of working out answers to problems	1	2	3	4	5	6
6	I like to invent new things or ideas	1	2	3	4	5	6

7	I often think of new and unusual ways of doing things	1	2	3	4	5	6
8	I can always think of new ways of looking at things	1	2	3	4	5	6

Please select a number beside each statement to indicate how much you agree with it.

		Disagree a lot 1	**Disagree 2**	**Disagree a little 3**	**Agree a little 4**	**Agree 5**	**Agree a lot 6**
1	I am not very popular with members of the opposite sex	1	2	3	4	5	6
2	It is difficult to make friends with members of my own sex	1	2	3	4	5	6
3	I make friends easily with boys	1	2	3	4	5	6
4	I make friends easily with girls	1	2	3	4	5	6
5	I have lots of friends of the opposite sex	1	2	3	4	5	6
6	Not many people from my own sex like me	1	2	3	4	5	6
7	I do not get along very well with boys	1	2	3	4	5	6
8	I do not get along very well with girls	1	2	3	4	5	6
9	I make friends easily with members of my own sex	1	2	3	4	5	6

Reference: The Gratitude Questionnaire (GQ-6; McCullough et al., 2002)

1	I have so much in life to be thankful for	1	2	3	4	5	6
2	If I had to list everything that I felt grateful for, it would be a very long list	1	2	3	4	5	6
3	When I look at the world, I don't see much to be grateful for	1	2	3	4	5	6
4	I am grateful to a wide variety of people	1	2	3	4	5	6
5	As I get older I find myself more able to appreciate the people, events, and situations that have been part of my life history	1	2	3	4	5	6

Reference: Academic Resilience Scale (adaptation; Martin & Marsh, 2006)

Please select a number beside each statement to indicate how much you agree with it.

		Disagree a lot 1	**Disagree 2**	**Disagree a little 3**	**Agree a little 4**	**Agree 5**	**Agree a lot 6**
1	I believe I am mentally tough when it comes to overcoming life challenges	1	2	3	4	5	6
2	I don't usually let life stresses get on top of me	1	2	3	4	5	6
3	I'm good at bouncing back from disappointments in my life	1	2	3	4	5	6
4	I think I'm good at dealing with sources of pressure in my life	1	2	3	4	5	6
5	I don't let difficulties and disappointments in life affect my confidence	1	2	3	4	5	6
6	I'm good at dealing with setbacks (e.g. negative feedback on what I do, disappointing outcomes)	1	2	3	4	5	6

Appendix 3: Factor Loadings for the Combined Research Instrument

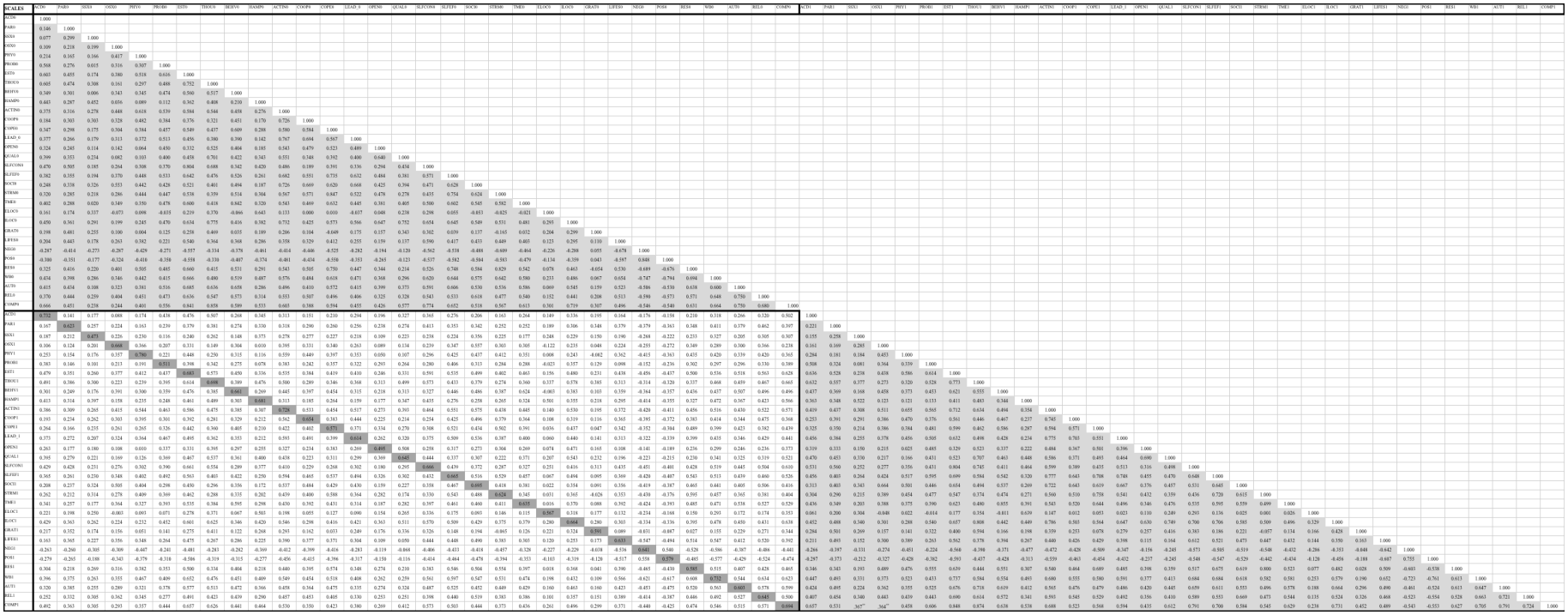
Table 41 *Summary Statistics for Scale Loadings Based on the Combined Research Instrument Model*

Scale	# of Items	Min	Max	Mean	Median
CT	3	.76	.86	.82	.85
SF	3	.55	1.10	.78	.70
LA	3	.58	.99	.80	.84
AI	3	.37	.74	.60	.70
OT	3	.34	.77	.55	.55
QS	3	.37	.74	.60	.68
SC	3	.34	.47	.41	.44
SE	3	.82	.91	.86	.84
SM	3	.66	.78	.73	.75
OE	3	.11	.18	.14	.14
TE	3	.68	.78	.73	.74
CH	3	.70	.81	.76	.77
IL	3	.77	.84	.80	.79
EL	3	.72	.77	.74	.74
POS	6	.46	.66	.58	.59
NEG	6	.61	.76	.68	.67
AUT	4	.73	.87	.79	.78
COMP	4	.82	.86	.84	.83
REL	4	.65	.86	.78	.81
THOU	3	.35	.73	.52	.49
BEHA	3	.48	.84	.70	.76
HAMP	5	.40	.69	.58	.58
ACAD	5	.70	.89	.80	.81
PHYS	5	.65	.95	.79	.81
EST	7	.74	.85	.78	.76
PROB	8	.64	.90	.78	.79
PRNT	4	.45	.91	.77	.86
SSX	5	.23	.84	.55	.66
OSX	4	.32	.87	.63	.66
WB	7	.57	.77	.68	.67
LSAT	1	1.00	1.00	1.00	1.00
RES	5	.76	.87	.82	.94
GRAT	6	.26	.81	.62	.63

Note: All parameters were estimated using stacked data and are reported in standardised format. ROPELOC factors are: CT = Cooperative Teamwork, SF = Self Efficacy, LA = Leadership Ability, AI = Active Involvement, OT = Open Thinking, QS = Quality Seeking, SC = Self Confidence, SE = Social Effectiveness, SM = Stress Management, OE = Overall Effectiveness, TE = Time Efficiency, CH = Coping with Change, IL = Internal Locus of Control, and EL = External Locus of Control. GHQ-12 factors are: POS = Positive and NEG = Negative. BPNSFS factors are: AUT = Autonomy, COMP = Competence, and REL = Relatedness. MES factors are: THOU = Thought, BEHA = Behavioural, and Hamp = Hampering. SDQ factors are: ACAD = General Academic, PHYS = Physical Abilities, EST = General Esteem, PROB = Problem Solving, PRNT = Parent Relations, SSX = Same-Sex Relations, and OSX = Opposite-Sex Relations. WB = Wellbeing, LSAT = Life Satisfaction, RES = Resilience, and GRAT = Gratitude.

Appendix 4: Multitrait-Multimethod Matix

Table 42 *Multitrait-Multimethod Matrix*



Note: Multitrait-multimethod matrix of correlations between matching factors at different time-waves. Convergent validities (highlighted in the diagonal of bolded box) are all statistically significant and consistently higher than correlations of non-matching factors; heterotrait-heteromethod (different trait, different method) correlations between non-matching traits at different time-waves (off-diagonal values in the bolded box) and heterotrait-monomethod (different trait, same method) correlations (off-diagonal values within each of the triangular submatrices highlighted in light grey).

Table 43 *MTMM Heterotrait-Monomethod Correlation Matrix for Time-Wave 0 (T0)*

SCALES	ACD0	PAR0	SSX0	OSX0	PHY0	PROB0	SEst0	THOU0	BEHV0	HAMP0	OVEF0	ACIN0	COOP0	COPE0	LEAD0	OPEN0	QUAL0	CON0	EFF0	SOCI0	STRM0	TME0	ELOC0	ILOC0	GRAT0	LIFES0	NEG0	POS0	RES0	WB0	AUT0	REL0	CMP0
ACD0	1.000																																
PAR0	0.146	1.000																															
SSX0	0.077	0.299	1.000																														
OSX0	0.109	0.218	0.199	1.000																													
PHY0	0.214	0.165	0.166	0.417	1.000																												
PROB0	0.568	0.276	0.015	0.316	0.307	1.000																											
SEst0	0.603	0.455	0.174	0.380	0.518	0.616	1.000																										
THOU0	0.605	0.474	0.308	0.161	0.297	0.488	0.752	1.000																									
BEHV0	0.349	0.301	0.006	0.343	0.345	0.474	0.560	0.517	1.000																								
HAMP0	0.443	0.287	0.452	0.036	0.089	0.112	0.362	0.408	0.210	1.000																							
OVEF0	0.093	0.206	0.059	0.086	0.064	0.151	0.169	0.128	0.066	0.039	1.000																						
ACIN0	0.375	0.316	0.278	0.448	0.618	0.539	0.584	0.544	0.458	0.276	0.120	1.000																					
COOP0	0.184	0.303	0.303	0.328	0.482	0.384	0.376	0.321	0.451	0.170	0.067	0.726	1.000																				
COPE0	0.347	0.298	0.175	0.304	0.384	0.457	0.549	0.437	0.609	0.288	0.195	0.580	0.584	1.000																			
LEAD0	0.377	0.266	0.179	0.313	0.372	0.513	0.456	0.380	0.390	0.142	0.154	0.767	0.694	0.567	1.000																		
OPEN0	0.324	0.245	0.114	0.142	0.064	0.450	0.332	0.525	0.404	0.185	-0.041	0.543	0.479	0.523	0.489	1.000																	
QUAL0	0.399	0.353	0.234	0.082	0.103	0.400	0.458	0.701	0.422	0.343	0.124	0.551	0.348	0.392	0.400	0.640	1.000																
CON0	0.470	0.505	0.185	0.264	0.308	0.370	0.804	0.688	0.342	0.420	0.137	0.486	0.189	0.391	0.336	0.294	0.434	1.000															
EFF0	0.382	0.355	0.194	0.370	0.448	0.533	0.642	0.476	0.526	0.261	0.064	0.682	0.551	0.735	0.632	0.484	0.381	0.571	1.000														
SOCI0	0.248	0.338	0.326	0.553	0.442	0.428	0.521	0.401	0.494	0.187	0.140	0.726	0.669	0.620	0.668	0.425	0.394	0.471	0.628	1.000													
STRM0	0.320	0.285	0.218	0.286	0.444	0.447	0.538	0.359	0.514	0.304	0.079	0.567	0.571	0.847	0.522	0.478	0.278	0.435	0.754	0.624	1.000												
TME0	0.402	0.288	0.020	0.349	0.350	0.478	0.600	0.418	0.842	0.320	0.095	0.543	0.469	0.632	0.445	0.381	0.405	0.500	0.602	0.545	0.582	1.000											
ELOC0	0.161	0.174	0.337	-0.073	0.098	-0.035	0.219	0.370	-0.066	0.643	-0.068	0.133	0.000	0.010	-0.037	0.048	0.238	0.298	0.055	-0.053	-0.025	-0.021	1.000										
ILOC0	0.450	0.361	0.291	0.199	0.245	0.470	0.634	0.775	0.416	0.382	0.108	0.732	0.425	0.573	0.566	0.647	0.752	0.654	0.645	0.549	0.531	0.481	0.293	1.000									
GRAT0	0.198	0.481	0.255	0.100	0.004	0.125	0.258	0.469	0.035	0.189	-0.006	0.206	0.104	-0.049	0.175	0.157	0.343	0.302	0.039	0.137	-0.165	0.032	0.204	0.299	1.000								
LIFES0	0.204	0.443	0.178	0.263	0.382	0.221	0.540	0.364	0.368	0.286	0.067	0.358	0.329	0.412	0.255	0.159	0.137	0.590	0.417	0.433	0.449	0.403	0.123	0.295	0.110	1.000							
NEG0	-0.287	-0.414	-0.273	-0.287	-0.429	-0.271	-0.557	-0.334	-0.378	-0.461	-0.028	-0.414	-0.446	-0.525	-0.282	-0.194	-0.120	-0.562	-0.538	-0.488	-0.609	-0.464	-0.226	-0.288	0.055	-0.678	1.000						
POS0	-0.300	-0.351	-0.177	-0.324	-0.410	-0.350	-0.558	-0.330	-0.407	-0.374	-0.101	-0.481	-0.434	-0.550	-0.353	-0.265	-0.123	-0.537	-0.582	-0.504	-0.583	-0.479	-0.134	-0.359	0.043	-0.597	0.848	1.000					
RES0	0.325	0.416	0.220	0.401	0.505	0.485	0.660	0.415	0.531	0.291	0.063	0.543	0.505	0.750	0.447	0.344	0.214	0.526	0.748	0.584	0.829	0.542	0.078	0.463	-0.054	0.530	-0.689	-0.676	1.000				
WB0	0.434	0.398	0.286	0.346	0.442	0.415	0.666	0.480	0.519	0.487	0.036	0.576	0.484	0.618	0.471	0.368	0.296	0.620	0.644	0.575	0.642	0.580	0.233	0.486	0.067	0.654	-0.747	-0.794	0.694	1.000			
AUT0	0.415	0.434	0.108	0.323	0.381	0.516	0.685	0.636	0.658	0.286	0.136	0.496	0.410	0.572	0.415	0.399	0.373	0.591	0.606	0.530	0.536	0.586	0.069	0.545	0.159	0.523	-0.506	-0.530	0.638	0.600	1.000		
REL0	0.370	0.444	0.259	0.404	0.451	0.473	0.636	0.547	0.573	0.314	0.125	0.553	0.507	0.496	0.406	0.325	0.328	0.543	0.533	0.618	0.477	0.540	0.152	0.441	0.208	0.513	-0.590	-0.573	0.571	0.648	0.750	1.000	
CMP0	0.666	0.451	0.238	0.244	0.401	0.556	0.841	0.858	0.589	0.533	0.132	0.603	0.388	0.594	0.455	0.426	0.577	0.774	0.652	0.518	0.567	0.613	0.301	0.719	0.307	0.496	-0.546	-0.540	0.631	0.664	0.750	0.680	1.000

Note. Time-wave 0 heterotrait-monomethod (different trait, same method) correlations (off-diagonal values within each of the triangular submatrices highlighted in light grey.

Table 44 *MTMM Heterotrait-Monomethod Correlation Matrix for Time-Wave 1 (T1)*

SCALES	ACD1	PAR1	SSX1	OSX1	PHY1	PROB1	SEst1	THOU1	BEHV1	HAMP1	OVEF1	ACIN1	COOP1	COPE1	LEAD1	OPEN1	QUAL1	CON1	EFF1	SOCII	STRM1	TME1	ELOC1	ILOC1	GRAT1	LIFES1	NEG1	POS1	RES1	WB1	AUT1	REL1	CMP1
ACD1	1.000																																
PAR1	0.221	1.000																															
SSX1	0.155	0.258	1.000																														
OSX1	0.161	0.169	0.285	1.000																													
PHY1	0.284	0.181	0.184	0.453	1.000																												
PROB1	0.508	0.324	0.081	0.364	0.339	1.000																											
SEst1	0.636	0.528	0.238	0.438	0.586	0.614	1.000																										
THOU1	0.632	0.557	0.377	0.273	0.320	0.528	0.773	1.000																									
BEHV1	0.437	0.369	0.168	0.458	0.373	0.453	0.621	0.555	1.000																								
HAMP1	0.363	0.348	0.522	0.123	0.121	0.133	0.411	0.483	0.344	1.000																							
OVEF1	0.087	0.043	0.101	0.103	0.053	0.048	0.183	0.093	0.129	0.005	1.000																						
ACIN1	0.419	0.437	0.308	0.511	0.655	0.565	0.712	0.634	0.494	0.354	0.117	1.000																					
COOP1	0.253	0.391	0.291	0.386	0.470	0.376	0.561	0.446	0.467	0.237	0.073	0.745	1.000																				
COPE1	0.325	0.350	0.214	0.386	0.384	0.481	0.599	0.462	0.586	0.287	0.257	0.594	0.571	1.000																			
LEAD1	0.456	0.384	0.255	0.378	0.456	0.505	0.632	0.498	0.428	0.234	0.152	0.775	0.703	0.551	1.000																		
OPEN1	0.319	0.333	0.150	0.215	0.025	0.485	0.329	0.523	0.337	0.222	-0.064	0.484	0.367	0.501	0.396	1.000																	
QUAL1	0.470	0.453	0.330	0.217	0.166	0.431	0.523	0.707	0.463	0.448	0.119	0.586	0.371	0.493	0.464	0.690	1.000																
CON1	0.531	0.560	0.252	0.277	0.356	0.431	0.804	0.745	0.411	0.464	0.158	0.599	0.389	0.435	0.513	0.316	0.498	1.000															
EFF1	0.456	0.403	0.264	0.424	0.517	0.595	0.699	0.584	0.542	0.320	0.097	0.777	0.643	0.708	0.748	0.455	0.470	0.648	1.000														
SOCII	0.313	0.403	0.343	0.664	0.501	0.446	0.654	0.494	0.537	0.269	0.197	0.722	0.643	0.619	0.667	0.376	0.457	0.531	0.645	1.000													
STRM1	0.304	0.290	0.215	0.389	0.454	0.477	0.547	0.374	0.474	0.271	0.102	0.560	0.510	0.758	0.541	0.432	0.359	0.436	0.720	0.615	1.000												
TME1	0.436	0.349	0.203	0.388	0.375	0.390	0.623	0.480	0.855	0.391	0.198	0.543	0.520	0.644	0.496	0.346	0.476	0.535	0.595	0.559	0.499	1.000											
ELOC1	0.061	0.200	0.304	-0.048	0.022	-0.014	0.177	0.354	-0.011	0.639	-0.088	0.147	0.012	0.053	0.023	0.110	0.249	0.293	0.136	0.025	0.001	0.026	1.000										
ILOC1	0.452	0.488	0.340	0.301	0.288	0.540	0.657	0.808	0.442	0.449	0.142	0.786	0.503	0.564	0.647	0.630	0.749	0.700	0.706	0.585	0.509	0.496	0.329	1.000									
GRAT1	0.284	0.501	0.269	0.157	0.141	0.322	0.400	0.594	0.166	0.198	-0.034	0.359	0.253	0.078	0.279	0.257	0.416	0.383	0.186	0.221	-0.057	0.134	0.166	0.428	1.000								
LIFES1	0.211	0.493	0.152	0.300	0.389	0.263	0.562	0.378	0.394	0.267	0.047	0.440	0.426	0.429	0.398	0.115	0.164	0.612	0.521	0.473	0.447	0.432	0.144	0.350	0.163	1.000							
NEG1	-0.266	-0.397	-0.331	-0.274	-0.451	-0.224	-0.560	-0.390	-0.371	-0.477	-0.010	-0.472	-0.428	-0.509	-0.347	-0.156	-0.245	-0.573	-0.505	-0.519	-0.548	-0.432	-0.286	-0.353	-0.048	-0.642	1.000						
POS1	-0.297	-0.373	-0.212	-0.327	-0.428	-0.382	-0.593	-0.437	-0.428	-0.313	-0.084	-0.559	-0.463	-0.454	-0.432	-0.237	-0.245	-0.548	-0.547	-0.529	-0.442	-0.434	-0.120	-0.456	-0.188	-0.607	0.755	1.000					
RES1	0.346	0.343	0.193	0.489	0.476	0.555	0.639	0.444	0.551	0.307	0.086	0.540	0.464	0.689	0.485	0.398	0.359	0.517	0.675	0.619	0.800	0.523	0.077	0.482	0.028	0.509	-0.603	-0.538	1.000				
WB1	0.447	0.493	0.331	0.373	0.523	0.433	0.737	0.584	0.554	0.493	0.013	0.680	0.555	0.580	0.591	0.377	0.413	0.684	0.684	0.618	0.582	0.581	0.253	0.579	0.190	0.652	-0.723	-0.761	0.613	1.000			
AUT1	0.424	0.495	0.224	0.362	0.355	0.525	0.676	0.718	0.619	0.412	0.003	0.565	0.476	0.479	0.486	0.420	0.445	0.659	0.611	0.553	0.496	0.578	0.188	0.664	0.296	0.490	-0.461	-0.524	0.613	0.647	1.000		
REL1	0.407	0.454	0.340	0.443	0.439	0.443	0.690	0.614	0.572	0.341	0.066	0.593	0.545	0.529	0.492	0.356	0.410	0.589	0.553	0.669	0.473	0.544	0.135	0.524	0.326	0.468	-0.523	-0.554	0.528	0.663	0.721	1.000	
CMP1	0.657	0.531	.367**	.364**	0.458	0.606	0.848	0.874	0.638	0.538	0.073	0.688	0.523	0.568	0.594	0.435	0.612	0.791	0.700	0.584	0.545	0.629	0.238	0.731	0.452	0.489	-0.543	-0.553	0.627	0.705	0.791	0.724	1.000

Note. Time-wave 1 heterotrait-monomethod (different trait, same method) correlations (off-diagonal values within each of the triangular submatrices highlighted in light grey.

Table 45 *MTMM Monotrait-Heteromethod Correlation Matrix*

SCALES	ACD0	PAR0	SSX0	OSX0	PHY0	PROB0	SEst0	THOU0	BEHV0	HAMP0	OVEF0	ACIN0	COOP0	COPE0	LEAD0	OPEN0	QUAL0	CON0	EFF0	SOCI0	STRM0	TME0	ELOC0	ILOC0	GRAT0	LIFES0	NEG0	POS0	RES0	WB0	AUT0	REL0	CMP0
ACD1	0.732	0.141	0.177	0.088	0.174	0.438	0.476	0.507	0.268	0.345	0.083	0.313	0.151	0.210	0.294	0.196	0.327	0.365	0.276	0.206	0.163	0.264	0.149	0.336	0.195	0.164	-0.176	-0.158	0.210	0.318	0.266	0.320	0.502
PAR1	0.167	0.623	0.257	0.224	0.163	0.239	0.379	0.381	0.274	0.330	0.055	0.318	0.290	0.260	0.256	0.238	0.274	0.413	0.353	0.342	0.252	0.252	0.189	0.306	0.348	0.379	-0.379	-0.363	0.348	0.411	0.379	0.462	0.397
SSX1	0.187	0.212	0.473	0.226	0.230	0.116	0.240	0.262	0.148	0.373	0.069	0.278	0.277	0.227	0.218	0.109	0.223	0.238	0.224	0.356	0.225	0.177	0.248	0.229	0.150	0.190	-0.268	-0.222	0.233	0.327	0.205	0.305	0.307
OSX1	0.106	0.124	0.201	0.668	0.366	0.207	0.331	0.149	0.304	0.010	0.045	0.395	0.331	0.340	0.263	0.089	0.134	0.239	0.347	0.557	0.303	0.305	-0.122	0.235	0.048	0.224	-0.255	-0.272	0.349	0.289	0.300	0.366	0.238
PHY1	0.253	0.154	0.176	0.357	0.780	0.221	0.448	0.250	0.315	0.116	0.079	0.559	0.449	0.397	0.353	0.050	0.107	0.296	0.425	0.437	0.412	0.351	0.008	0.243	-0.002	0.362	-0.415	-0.363	0.435	0.420	0.339	0.420	0.365
PROB1	0.383	0.146	0.101	0.213	0.191	0.511	0.398	0.342	0.275	0.078	-0.018	0.383	0.242	0.357	0.322	0.293	0.264	0.280	0.406	0.313	0.284	0.288	-0.023	0.357	0.129	0.098	-0.152	-0.236	0.302	0.297	0.296	0.330	0.389
SEst1	0.479	0.351	0.260	0.377	0.412	0.437	0.683	0.573	0.450	0.336	0.124	0.535	0.384	0.419	0.410	0.246	0.331	0.591	0.535	0.499	0.402	0.463	0.156	0.480	0.231	0.438	-0.456	-0.437	0.500	0.536	0.518	0.563	0.628
THOU1	0.491	0.386	0.300	0.223	0.239	0.395	0.614	0.698	0.389	0.476	0.107	0.500	0.289	0.346	0.368	0.313	0.499	0.573	0.433	0.379	0.274	0.360	0.337	0.578	0.385	0.313	-0.314	-0.320	0.337	0.468	0.459	0.467	0.665
BEHV1	0.301	0.249	0.176	0.391	0.300	0.359	0.476	0.385	0.661	0.269	0.090	0.445	0.397	0.454	0.315	0.218	0.313	0.327	0.446	0.486	0.387	0.624	-0.003	0.383	0.103	0.359	-0.364	-0.357	0.436	0.437	0.507	0.496	0.496
HAMP1	0.413	0.314	0.397	0.158	0.235	0.248	0.461	0.489	0.303	0.681	0.094	0.313	0.185	0.264	0.159	0.177	0.347	0.435	0.276	0.258	0.265	0.324	0.501	0.355	0.218	0.295	-0.414	-0.355	0.327	0.472	0.367	0.423	0.566
OVEF1	0.032	0.111	0.135	0.066	-0.034	0.064	0.114	0.092	0.007	-0.046	0.342	0.033	0.056	0.066	0.145	-0.047	0.024	0.064	0.051	0.117	-0.019	0.016	-0.060	0.090	0.097	-0.012	0.016	-0.024	0.038	-0.048	0.028	-0.011	0.053
ACIN1	0.386	0.309	0.265	0.415	0.544	0.463	0.586	0.475	0.385	0.307	0.143	0.728	0.533	0.454	0.517	0.273	0.393	0.464	0.551	0.575	0.438	0.445	0.140	0.530	0.195	0.372	-0.420	-0.411	0.456	0.516	0.430	0.522	0.571
COOP1	0.193	0.234	0.262	0.303	0.395	0.301	0.392	0.281	0.329	0.212	0.062	0.562	0.654	0.383	0.444	0.225	0.214	0.254	0.425	0.496	0.379	0.364	0.108	0.319	0.116	0.365	-0.395	-0.372	0.383	0.414	0.344	0.475	0.368
COPE1	0.264	0.166	0.235	0.261	0.265	0.326	0.442	0.360	0.405	0.210	0.092	0.422	0.402	0.571	0.371	0.334	0.270	0.308	0.521	0.434	0.502	0.391	0.036	0.437	0.047	0.342	-0.352	-0.304	0.489	0.399	0.423	0.382	0.439
LEAD1	0.373	0.272	0.207	0.324	0.364	0.467	0.495	0.362	0.353	0.212	0.114	0.593	0.491	0.399	0.614	0.262	0.320	0.375	0.509	0.536	0.387	0.400	0.060	0.440	0.141	0.313	-0.322	-0.339	0.399	0.435	0.346	0.429	0.441
OPEN1	0.263	0.177	0.180	0.108	0.010	0.337	0.331	0.395	0.297	0.255	0.064	0.327	0.234	0.383	0.269	0.495	0.508	0.258	0.317	0.273	0.304	0.269	0.074	0.471	0.165	0.108	-0.141	-0.189	0.236	0.299	0.246	0.236	0.373
QUAL1	0.395	0.279	0.221	0.169	0.126	0.369	0.467	0.537	0.361	0.400	0.145	0.438	0.223	0.311	0.299	0.369	0.645	0.444	0.337	0.307	0.222	0.371	0.207	0.543	0.232	0.196	-0.223	-0.215	0.230	0.341	0.325	0.319	0.521
CON1	0.429	0.428	0.231	0.276	0.302	0.390	0.661	0.554	0.289	0.377	0.098	0.410	0.229	0.268	0.302	0.180	0.295	0.666	0.439	0.372	0.287	0.327	0.251	0.416	0.313	0.435	-0.451	-0.401	0.428	0.519	0.445	0.504	0.610
EFF1	0.365	0.261	0.230	0.348	0.402	0.492	0.563	0.403	0.422	0.250	0.092	0.594	0.465	0.537	0.494	0.326	0.302	0.432	0.665	0.516	0.529	0.457	0.067	0.494	0.095	0.369	-0.420	-0.407	0.543	0.513	0.439	0.460	0.526
SOCI1	0.208	0.237	0.324	0.505	0.404	0.298	0.450	0.296	0.356	0.172	0.121	0.537	0.484	0.429	0.430	0.159	0.227	0.358	0.467	0.695	0.418	0.381	0.022	0.354	0.091	0.356	-0.419	-0.387	0.465	0.441	0.405	0.506	0.416
STRM1	0.262	0.212	0.314	0.278	0.409	0.369	0.462	0.288	0.335	0.202	0.093	0.439	0.400	0.588	0.364	0.282	0.174	0.330	0.543	0.488	0.624	0.345	0.031	0.365	-0.026	0.353	-0.430	-0.376	0.595	0.457	0.365	0.381	0.404
TME1	0.341	0.257	0.177	0.364	0.327	0.393	0.535	0.384	0.595	0.298	0.100	0.430	0.392	0.431	0.314	0.187	0.282	0.397	0.461	0.460	0.411	0.635	0.016	0.370	0.088	0.392	-0.424	-0.393	0.485	0.471	0.538	0.527	0.529
ELOC1	0.221	0.198	0.250	-0.003	0.093	0.071	0.278	0.371	0.067	0.503	0.038	0.198	0.055	0.127	0.090	0.154	0.265	0.336	0.175	0.093	0.146	0.115	0.567	0.318	0.177	0.132	-0.234	-0.168	0.150	0.293	0.172	0.174	0.353
ILOC1	0.429	0.363	0.262	0.224	0.232	0.452	0.601	0.625	0.346	0.420	0.149	0.546	0.298	0.416	0.421	0.363	0.511	0.570	0.509	0.429	0.375	0.379	0.280	0.664	0.280	0.303	-0.334	-0.336	0.395	0.478	0.450	0.431	0.638
GRAT1	0.217	0.352	0.174	0.156	0.051	0.141	0.275	0.411	0.122	0.268	-0.090	0.293	0.162	0.033	0.249	0.176	0.336	0.326	0.148	0.194	-0.065	0.126	0.221	0.324	0.591	0.089	-0.031	-0.087	0.027	0.155	0.229	0.271	0.344
LIFES1	0.163	0.365	0.227	0.356	0.348	0.264	0.475	0.267	0.286	0.225	0.038	0.390	0.377	0.371	0.304	0.109	0.050	0.444	0.448	0.490	0.383	0.303	0.120	0.253	0.173	0.633	-0.547	-0.494	0.514	0.547	0.412	0.520	0.392
NEG1	-0.263	-0.260	-0.305	-0.309	-0.447	-0.241	-0.481	-0.283	-0.282	-0.369	-0.033	-0.412	-0.399	-0.416	-0.283	-0.119	-0.068	-0.406	-0.433	-0.418	-0.457	-0.328	-0.227	-0.229	-0.038	-0.536	0.641	0.540	-0.528	-0.586	-0.387	-0.486	-0.441
POS1	-0.279	-0.265	-0.188	-0.343	-0.379	-0.310	-0.506	-0.319	-0.315	-0.277	-0.042	-0.456	-0.415	-0.396	-0.317	-0.150	-0.116	-0.414	-0.464	-0.478	-0.394	-0.353	-0.103	-0.319	-0.120	-0.517	0.558	0.579	-0.485	-0.577	-0.429	-0.524	-0.474
RES1	0.304	0.218	0.269	0.316	0.382	0.353	0.500	0.334	0.404	0.218	-0.014	0.440	0.395	0.574	0.348	0.274	0.210	0.383	0.546	0.504	0.554	0.397	0.018	0.368	0.041	0.390	-0.465	-0.430	0.585	0.515	0.407	0.428	0.465
WB1	0.396	0.375	0.263	0.355	0.467	0.409	0.652	0.476	0.451	0.409	0.128	0.549	0.454	0.518	0.408	0.262	0.259	0.561	0.597	0.547	0.531	0.474	0.198	0.432	0.109	0.566	-0.621	-0.617	0.608	0.732	0.544	0.634	0.623
AUT1	0.320	0.385	0.255	0.289	0.321	0.378	0.577	0.513	0.472	0.366	0.018	0.458	0.364	0.475	0.335	0.274	0.324	0.487	0.525	0.452	0.449	0.429	0.160	0.463	0.160	0.423	-0.453	-0.475	0.520	0.563	0.603	0.578	0.599
REL1	0.252	0.332	0.305	0.362	0.345	0.277	0.491	0.423	0.439	0.290	0.017	0.457	0.453	0.405	0.330	0.253	0.251	0.398	0.440	0.519	0.383	0.386	0.101	0.357	0.151	0.389	-0.414	-0.387	0.446	0.492	0.527	0.645	0.500
CMP1	0.492	0.363	0.305	0.293	0.357	0.444	0.657	0.626	0.441	0.464	0.047	0.530	0.350	0.423	0.380	0.269	0.412	0.573	0.503	0.444	0.373	0.436	0.261	0.496	0.299	0.371	-0.440	-0.425	0.474	0.546	0.515	0.571	0.694

Note. Convergent validities (highlighted in the diagonal of bolded box) are all statistically significant and consistently higher than correlations of non-matching factors. Heterotrait-heteromethod (different trait, different method) correlations between non-matching traits at different time-waves (off-diagonal values in the bolded box).

Appendix 5: Input Syntax, Fit Statistics and Item-Loadings for the Combined Instrument Model

Mplus Syntax (Input, Fit Indices and Loadings) for the Combined Latent Model

```
TITLE: Full Stacked Model F-Scores GO
DATA: FILE = "T0-T4_Mplus.FSfile.csv";
VARIABLE:
NAMES = Stu_ID Semester Time IDTime WEM_1 WEM_2 WEM_3 WEM_4 WEM_5 WEM_6 WEM_7
        LSat GHQ_1 GHQ_2 GHQ_3 GHQ_4 GHQ_5 GHQ_6 GHQ_7 GHQ_8 GHQ_9 GHQ_10
        GHQ_11 GHQ_12
        LEf_1 LEf_2 LEf_3 LEf_4 LEf_5 LEf_6 LEf_7 LEf_8 LEf_9 LEf_10 LEf_11 LEf_12
        LEf_13 LEf_14 LEf_15 LEf_16 LEf_17 LEf_18 LEf_19 LEf_20 LEf_21 LEf_22 LEf_23r
        LEf_24 LEf_25 LEf_26 LEf_27 LEf_28 LEf_29 LEf_30 LEf_31 LEf_32 LEf_33 LEf_34
        LEf_35 LEf_36 LEf_37 LEf_38r LEf_39 LEf_40 LEf_41 LEf_42 MvE_1 MvE_2 MvE_3
        MvE_4 MvE_5 MvE_6 MvE_7r MvE_8r MvE_9r MvE_10r MvE_11r PsNA1 PsNA2 PsNA3
        PsNA4
        PsNR5 PsNR6 PsNR7 PsNR8 PsNC9 PsNC10 PsNC11 PsNC12 SDQAc1 SDQAc2 SDQAc3
        SDQAc4
        SDQAc5 SDQPh1 SDQPh2 SDQPh3 SDQPh4 SDQPh5 SDQSE1 SDQSE2 SDQSE3 SDQSE4
        SDQSE5
        SDQSE6 SDQSE7 SDQPa1 SDQPa2 SDQPa3 SDQPa4r SDQPS1 SDQPS2 SDQPS3 SDQPS4
        SDQPS5
        SDQPS6 SDQPS7 SDQPS8 SDQR1r SDQR2r SDQR3 SDQR4 SDQR5 SDQR6r SDQR7r
        SDQR8r SDQR9
        Grat1 Grat2 Grat3r Grat4 Grat5 Res1 Res2 Res3 Res4 Res5 Res6;
         MISSING=.;
         CLUSTER = Stu_ID
         AUXILIARY = Time;
         AUXILIARY = IDTime;
         AUXILIARY = Semester;
         USEVARIABLES ARE
            WEM_1 WEM_2 WEM_3 WEM_4 WEM_5 WEM_6 WEM_7
            Res1 Res2 Res3 Res4 Res5 Res6
            Grat1 Grat2 Grat3r Grat4 Grat5
            MvE_1 MvE_2 MvE_3
            MvE_4 MvE_5 MvE_6
            MvE_7r MvE_8r MvE_9r MvE_10r MvE_11r
            LSat

            GHQ_1 GHQ_3  GHQ_4 GHQ_7 GHQ_8 GHQ_12
            GHQ_2 GHQ_5 GHQ_6 GHQ_9 GHQ_10 GHQ_11

            LEf_12 LEf_27 LEf_42
            LEf_1 LEf_15 LEf_29
            LEf_2 LEf_16 LEf_30
            LEf_3 LEf_17 LEf_31
            LEf_5 LEf_18 LEf_32
            LEf_6 LEf_20 LEf_33
            LEf_7 LEf_21 LEf_35
            LEf_9 LEf_22 LEf_36
            LEf_10 LEf_24 LEf_37
            LEf_11 LEf_25 LEf_39
            LEf_13 LEf_26 LEf_40
            LEf_14 LEf_28 LEf_41
            LEf_4 LEf_19 LEf_34
            LEf_8 LEf_23r LEf_38r
```

```
  PsNA1 PsNA2 PsNA3 PsNA4
  PsNR5 PsNR6 PsNR7 PsNR8
  PsNC9 PsNC10 PsNC11 PsNC12

  SDQAc1 SDQAc2 SDQAc3 SDQAc4 SDQAc5
  SDQPh1 SDQPh2 SDQPh3 SDQPh4 SDQPh5
  SDQSE1 SDQSE2 SDQSE3 SDQSE4 SDQSE5 SDQSE6 SDQSE7
  SDQPa1 SDQPa2 SDQPa3 SDQPa4r
  SDQPS1 SDQPS2 SDQPS3 SDQPS4 SDQPS5 SDQPS6
  SDQPS7 SDQPS8
  SDQR2r SDQR3 SDQR6r SDQR7r SDQR9
  SDQR1r SDQR4 SDQR5 SDQR8r;

ANALYSIS:
      TYPE = complex;
      Estimator=mlr;
      ROTATION = target(OBLIQUE);
      process=4;
MODEL:
  WellBei   by WEM_1-WEM_7 Res1-Grat5~0 (*a);
  Resil     by Res1-Res6 WEM_1-WEM_7~0 Grat1-Grat5 (*a);
  Grat      by Grat1-Grat5 WEM_1-Res6~0 (*a);

  Thought    by MvE_1-MvE_3 MvE_4-MvE_11r~0 (*t);
  Behav      by MvE_4-MvE_6 MvE_1-MvE_3~0 MvE_7r-MvE_11r~0 (*t);
  MufGuz     by MvE_7r-MvE_11r MvE_1-MvE_6~0 (*t);

  LifeSat by LSat@1; LSat@0;

  Positiv   by GHQ_1 GHQ_3  GHQ_4 GHQ_7 GHQ_8 GHQ_12;
  Negativ   by GHQ_2 GHQ_5 GHQ_6 GHQ_9 GHQ_10 GHQ_11;

  OvEff      by LEf_12-LEf_42 LEf_1-LEf_41~0 (*c);
  CoopTeam   by LEf_1-LEf_29 LEf_12-LEf_42~0 LEf_2-LEf_41~0 (*c);
  SelfEffic  by LEf_2-LEf_30 LEf_12-LEf_29~0 LEf_3-LEf_41~0 (*c);
  Lead       by LEf_3-LEf_31 LEf_12-LEf_30~0 LEf_5-LEf_41~0 (*c);
  ActInv     by LEf_5-LEf_32 LEf_12-LEf_31~0 LEf_6-LEf_41~0 (*c);
  OpenTh     by LEf_6-LEf_33 LEf_12-LEf_32~0 LEf_7-LEf_41~0 (*c);
  QualSeek   by LEf_7-LEf_35 LEf_12-LEf_33~0 LEf_9-LEf_41~0 (*c);
  SelfCon    by LEf_9-LEf_36 LEf_12-LEf_35~0 LEf_10-LEf_41~0 (*c);
  SocialEff  by LEf_10-LEf_37 LEf_12-LEf_36~0 LEf_11-LEf_41~0 (*c);
  StressMgnt by LEf_11-LEf_39 LEf_12-LEf_37~0 LEf_13-LEf_41~0 (*c);
  TimeEffic  by LEf_13-LEf_40 LEf_12-LEf_39~0 LEf_14-LEf_41~0 (*c);
  CopingCha  by LEf_14-LEf_41 LEf_12-LEf_40~0 (*c);

  ILoC by LEf_4-LEf_34 LEf_8-LEf_38r~0 (*d);
  ELoC by LEf_8-LEf_38r LEf_4-LEf_34~0 (*d);

  Autonom by PsNA1 PsNA2 PsNA3 PsNA4;
  Related by PsNR5 PsNR6 PsNR7 PsNR8;
  Compet  by PsNC9 PsNC10 PsNC11 PsNC12;

  Academ      by SDQAc1-SDQAc5 SDQPh1-SDQR8r~0 (*f);
  Physic      by SDQPh1-SDQPh5 SDQAc1-SDQAc5~0 SDQSE1-SDQR8r~0 (*f);
  SelfEste    by SDQSE1-SDQSE7 SDQAc1-SDQPh5~0 SDQPa1-SDQR8r~0 (*f);
  Parents     by SDQPa1-SDQPa4r SDQAc1-SDQSE7~0 SDQPS1-SDQR8r~0 (*f);
  Prob        by SDQPS1-SDQPS8 SDQAc1-SDQPa4r~0 SDQR2r-SDQR8r~0 (*f);
  SSexRels    by SDQR2r-SDQR9 SDQAc1-SDQPS8~0 SDQR1r-SDQR8r~0 (*f);
  OSexRels    by SDQR1r-SDQR8r SDQAc1-SDQR9~0 (*f);
```

```
    SAVEDATA:
    FILE IS ALL_T0T4_Long_FS.Comeon.csv;
    SAVE IS fscores;
    FORMAT IS free;
OUTPUT: sampstat STDYX modindices (4) residual;

THE MODEL ESTIMATION TERMINATED NORMALLY
MODEL FIT INFORMATION
Number of Free Parameters                     1429
Loglikelihood

        H0 Value                          -281675.545
        H0 Scaling Correction Factor            2.0974
          for MLR
        H1 Value                          -271898.641
        H1 Scaling Correction Factor            1.3962
          for MLR

Information Criteria

        Akaike (AIC)                       566209.091
        Bayesian (BIC)                     574031.729
        Sample-Size Adjusted BIC           569491.911
          (n* = (n + 2) / 24)

Chi-Square Test of Model Fit

        Value                              15434.953*
        Degrees of Freedom                 7750
        P-Value                            0.0000
        Scaling Correction Factor          1.2669
          for MLR

RMSEA (Root Mean Square Error Of Approximation)

        Estimate                           0.024
        90 Percent C.I.                    0.023  0.024
        Probability RMSEA <= .05           1.000

CFI/TLI

        CFI                                0.935
        TLI                                0.925

Chi-Square Test of Model Fit for the Baseline Model

        Value                         126462.776
        Degrees of Freedom                 8911
        P-Value                       0.0000

SRMR (Standardized Root Mean Square Residual)

        Value                              0.029
```

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Two-Tailed Est./S.E.	P-Value
WELLBEI BY				
WEM_1	0.569	0.036	15.730	0.000
WEM_2	0.757	0.028	27.414	0.000
WEM_3	0.672	0.031	21.518	0.000
WEM_4	0.765	0.025	30.374	0.000
WEM_5	0.742	0.029	25.201	0.000
WEM_6	0.630	0.033	19.316	0.000
WEM_7	0.632	0.033	19.044	0.000
RES1	0.018	0.027	0.647	0.518
RES2	0.027	0.024	1.123	0.261
RES3	-0.022	0.020	-1.111	0.266
RES4	0.006	0.020	0.307	0.759
RES5	0.003	0.020	0.176	0.860
RES6	-0.015	0.020	-0.750	0.453
GRAT1	-0.043	0.019	-2.299	0.021
GRAT2	-0.024	0.015	-1.565	0.118
GRAT3R	0.087	0.042	2.089	0.037
GRAT4	0.021	0.028	0.750	0.453
GRAT5	0.043	0.023	1.844	0.065
RESIL BY				
RES1	0.755	0.028	27.311	0.000
RES2	0.763	0.027	28.469	0.000
RES3	0.874	0.018	49.547	0.000
RES4	0.863	0.019	44.852	0.000
RES5	0.830	0.020	40.914	0.000
RES6	0.849	0.018	47.799	0.000
WEM_1	-0.004	0.025	-0.165	0.869
WEM_2	-0.023	0.027	-0.857	0.392
WEM_3	0.000	0.026	0.003	0.998
WEM_4	0.035	0.024	1.466	0.143
WEM_5	0.010	0.026	0.379	0.705
WEM_6	-0.009	0.030	-0.317	0.752
WEM_7	0.000	0.033	-0.009	0.993
GRAT1	0.476	0.059	8.017	0.000
GRAT2	0.460	0.057	8.052	0.000
GRAT3R	-0.051	0.043	-1.178	0.239
GRAT4	0.463	0.052	8.987	0.000
GRAT5	0.405	0.053	7.617	0.000
GRAT BY				
GRAT1	0.768	0.024	31.772	0.000
GRAT2	0.807	0.022	37.102	0.000
GRAT3R	0.257	0.029	8.815	0.000
GRAT4	0.618	0.032	19.212	0.000
GRAT5	0.633	0.032	19.489	0.000
WEM_1	0.120	0.030	4.000	0.000
WEM_2	0.016	0.021	0.789	0.430
WEM_3	-0.098	0.023	-4.179	0.000
WEM_4	-0.070	0.018	-3.892	0.000
WEM_5	-0.032	0.020	-1.597	0.110
WEM_6	0.073	0.030	2.469	0.014
WEM_7	0.029	0.029	1.004	0.315
RES1	0.162	0.026	6.206	0.000
RES2	-0.080	0.018	-4.416	0.000
RES3	0.011	0.017	0.644	0.520

	Estimate	S.E.	Est./S.E.	P-Value
RES4	0.001	0.015	0.088	0.930
RES5	-0.052	0.018	-2.969	0.003
RES6	-0.028	0.018	-1.596	0.110
THOUGHT BY				
MVE_1	0.730	0.036	20.551	0.000
MVE_2	0.487	0.057	8.533	0.000
MVE_3	0.346	0.052	6.588	0.000
MVE_4	0.084	0.025	3.354	0.001
MVE_5	0.026	0.024	1.077	0.282
MVE_6	0.363	0.037	9.689	0.000
MVE_7R	-0.346	0.049	-7.064	0.000
MVE_8R	-0.060	0.041	-1.475	0.140
MVE_9R	0.269	0.032	8.334	0.000
MVE_10R	-0.149	0.047	-3.184	0.001
MVE_11R	0.076	0.027	2.787	0.005
BEHAV BY				
MVE_4	0.762	0.025	30.923	0.000
MVE_5	0.844	0.024	34.834	0.000
MVE_6	0.480	0.034	14.119	0.000
MVE_1	0.052	0.036	1.450	0.147
MVE_2	0.228	0.049	4.687	0.000
MVE_3	0.371	0.042	8.825	0.000
MVE_7R	0.027	0.042	0.634	0.526
MVE_8R	0.027	0.040	0.682	0.495
MVE_9R	-0.290	0.028	-10.444	0.000
MVE_10R	0.211	0.040	5.257	0.000
MVE_11R	0.059	0.024	2.504	0.012
MUFGUZ BY				
MVE_7R	0.403	0.041	9.903	0.000
MVE_8R	0.583	0.031	18.531	0.000
MVE_9R	0.641	0.031	20.899	0.000
MVE_10R	0.580	0.036	16.164	0.000
MVE_11R	0.694	0.026	26.292	0.000
MVE_1	0.090	0.021	4.274	0.000
MVE_2	0.038	0.030	1.285	0.199
MVE_3	-0.031	0.031	-1.004	0.315
MVE_4	-0.033	0.018	-1.829	0.067
MVE_5	0.033	0.014	2.359	0.018
MVE_6	0.132	0.025	5.194	0.000
LIFESAT BY				
LSAT	1.000	0.000	999.000	999.000
POSITIV BY				
GHQ_1	0.457	0.034	13.617	0.000
GHQ_3	0.603	0.027	22.014	0.000
GHQ_4	0.548	0.027	19.991	0.000
GHQ_7	0.593	0.028	21.152	0.000
GHQ_8	0.594	0.029	20.304	0.000
GHQ_12	0.655	0.024	27.443	0.000
NEGATIV BY				
GHQ_2	0.606	0.025	24.172	0.000
GHQ_5	0.616	0.022	27.487	0.000
GHQ_6	0.606	0.023	26.415	0.000

	Estimate	S.E.	Est./S.E.	P-Value
GHQ_9	0.753	0.017	45.461	0.000
GHQ_10	0.762	0.017	43.718	0.000
GHQ_11	0.729	0.019	37.416	0.000
OVEFF BY				
LEF_12	-0.179	0.314	-0.571	0.568
LEF_27	0.110	0.343	0.321	0.748
LEF_42	0.143	0.295	0.483	0.629
LEF_1	-0.052	0.044	-1.186	0.236
LEF_15	-0.137	0.029	-4.762	0.000
LEF_29	0.172	0.044	3.894	0.000
LEF_2	-0.046	0.039	-1.181	0.238
LEF_16	-0.074	0.067	-1.115	0.265
LEF_30	0.170	0.046	3.678	0.000
LEF_3	-0.056	0.033	-1.696	0.090
LEF_17	-0.074	0.041	-1.823	0.068
LEF_31	0.127	0.036	3.554	0.000
LEF_5	-0.076	0.072	-1.060	0.289
LEF_18	-0.052	0.080	-0.654	0.513
LEF_32	0.227	0.058	3.921	0.000
LEF_6	-0.030	0.056	-0.547	0.585
LEF_20	-0.053	0.064	-0.826	0.409
LEF_33	0.221	0.072	3.056	0.002
LEF_7	-0.132	0.058	-2.296	0.022
LEF_21	-0.035	0.122	-0.286	0.775
LEF_35	0.086	0.057	1.519	0.129
LEF_9	-0.149	0.272	-0.549	0.583
LEF_22	0.072	0.184	0.392	0.695
LEF_36	0.118	0.255	0.463	0.643
LEF_10	-0.193	0.024	-8.069	0.000
LEF_24	0.056	0.028	1.974	0.048
LEF_37	0.096	0.037	2.579	0.010
LEF_11	-0.192	0.037	-5.149	0.000
LEF_25	0.032	0.025	1.273	0.203
LEF_39	0.090	0.085	1.058	0.290
LEF_13	-0.194	0.067	-2.892	0.004
LEF_26	0.074	0.110	0.666	0.506
LEF_40	0.117	0.058	2.013	0.044
LEF_14	-0.234	0.034	-6.895	0.000
LEF_28	0.065	0.053	1.228	0.219
LEF_41	0.119	0.066	1.793	0.073
COOPTEAM BY				
LEF_1	0.755	0.075	10.080	0.000
LEF_15	0.859	0.080	10.760	0.000
LEF_29	0.854	0.051	16.762	0.000
LEF_12	0.098	0.046	2.151	0.031
LEF_27	0.127	0.044	2.904	0.004
LEF_42	0.049	0.037	1.342	0.180
LEF_2	-0.016	0.040	-0.392	0.695
LEF_16	-0.002	0.022	-0.078	0.938
LEF_30	0.050	0.050	1.004	0.315
LEF_3	0.005	0.030	0.165	0.869
LEF_17	0.029	0.035	0.829	0.407
LEF_31	0.030	0.044	0.691	0.490
LEF_5	-0.021	0.039	-0.540	0.589
LEF_18	0.133	0.109	1.228	0.220
LEF_32	0.004	0.049	0.084	0.933
LEF_6	0.068	0.031	2.228	0.026

	Estimate	S.E.	Est./S.E.	P-Value
LEF_20	0.047	0.058	0.810	0.418
LEF_33	0.046	0.051	0.892	0.372
LEF_7	-0.030	0.042	-0.721	0.471
LEF_21	0.030	0.024	1.249	0.212
LEF_35	0.052	0.028	1.844	0.065
LEF_9	-0.004	0.055	-0.068	0.946
LEF_22	0.008	0.044	0.175	0.861
LEF_36	0.036	0.042	0.840	0.401
LEF_10	-0.010	0.026	-0.382	0.702
LEF_24	0.014	0.032	0.444	0.657
LEF_37	0.016	0.028	0.581	0.561
LEF_11	0.046	0.038	1.221	0.222
LEF_25	0.029	0.036	0.787	0.431
LEF_39	-0.011	0.031	-0.365	0.715
LEF_13	0.038	0.035	1.094	0.274
LEF_26	0.019	0.025	0.743	0.457
LEF_40	0.016	0.034	0.463	0.643
LEF_14	0.013	0.035	0.387	0.699
LEF_28	0.029	0.031	0.950	0.342
LEF_41	-0.016	0.033	-0.481	0.630
SELFEFFI BY				
LEF_2	0.695	0.175	3.981	0.000
LEF_16	1.095	0.112	9.806	0.000
LEF_30	0.552	0.173	3.182	0.001
LEF_12	0.098	0.038	2.567	0.010
LEF_27	0.098	0.032	3.034	0.002
LEF_42	0.122	0.042	2.888	0.004
LEF_1	-0.088	0.036	-2.423	0.015
LEF_15	0.055	0.035	1.564	0.118
LEF_29	0.016	0.036	0.456	0.649
LEF_3	-0.024	0.028	-0.884	0.377
LEF_17	0.040	0.047	0.853	0.393
LEF_31	0.028	0.042	0.670	0.503
LEF_5	-0.033	0.042	-0.781	0.435
LEF_18	0.052	0.072	0.723	0.470
LEF_32	0.072	0.053	1.358	0.175
LEF_6	-0.013	0.042	-0.320	0.749
LEF_20	0.116	0.043	2.708	0.007
LEF_33	0.054	0.049	1.118	0.264
LEF_7	0.028	0.050	0.568	0.570
LEF_21	0.013	0.025	0.511	0.610
LEF_35	-0.026	0.030	-0.855	0.392
LEF_9	0.098	0.045	2.208	0.027
LEF_22	0.085	0.044	1.935	0.053
LEF_36	0.079	0.038	2.053	0.040
LEF_10	0.005	0.028	0.190	0.849
LEF_24	0.002	0.024	0.090	0.928
LEF_37	-0.023	0.032	-0.721	0.471
LEF_11	0.006	0.034	0.180	0.857
LEF_25	0.031	0.049	0.637	0.524
LEF_39	0.044	0.036	1.214	0.225
LEF_13	-0.013	0.039	-0.333	0.739
LEF_26	0.028	0.026	1.049	0.294
LEF_40	0.064	0.041	1.541	0.123
LEF_14	0.025	0.030	0.831	0.406
LEF_28	-0.009	0.032	-0.270	0.787
LEF_41	0.025	0.032	0.768	0.443

	Estimate	S.E.	Est./S.E.	P-Value
LEAD BY				
LEF_3	0.991	0.051	19.444	0.000
LEF_17	0.840	0.107	7.820	0.000
LEF_31	0.583	0.077	7.532	0.000
LEF_12	0.071	0.049	1.447	0.148
LEF_27	0.035	0.039	0.907	0.364
LEF_42	0.073	0.045	1.626	0.104
LEF_1	0.004	0.059	0.064	0.949
LEF_15	0.043	0.043	0.999	0.318
LEF_29	-0.011	0.036	-0.314	0.754
LEF_2	0.096	0.066	1.448	0.148
LEF_16	-0.029	0.029	-1.009	0.313
LEF_30	-0.003	0.049	-0.065	0.948
LEF_5	0.060	0.069	0.873	0.382
LEF_18	-0.054	0.068	-0.788	0.430
LEF_32	0.177	0.051	3.492	0.000
LEF_6	0.015	0.033	0.450	0.653
LEF_20	-0.051	0.053	-0.961	0.337
LEF_33	0.094	0.048	1.949	0.051
LEF_7	0.011	0.044	0.244	0.807
LEF_21	-0.011	0.027	-0.415	0.678
LEF_35	0.078	0.034	2.324	0.020
LEF_9	0.041	0.035	1.164	0.244
LEF_22	0.080	0.040	1.981	0.048
LEF_36	0.050	0.036	1.392	0.164
LEF_10	-0.003	0.023	-0.144	0.886
LEF_24	0.019	0.036	0.532	0.595
LEF_37	0.043	0.025	1.668	0.095
LEF_11	0.032	0.040	0.804	0.422
LEF_25	-0.018	0.038	-0.476	0.634
LEF_39	0.022	0.029	0.771	0.441
LEF_13	0.039	0.032	1.203	0.229
LEF_26	0.027	0.025	1.063	0.288
LEF_40	-0.020	0.032	-0.645	0.519
LEF_14	0.015	0.026	0.570	0.569
LEF_28	0.032	0.030	1.060	0.289
LEF_41	0.007	0.037	0.199	0.842
ACTINV BY				
LEF_5	0.737	0.112	6.581	0.000
LEF_18	0.695	0.318	2.188	0.029
LEF_32	0.365	0.115	3.178	0.001
LEF_12	0.015	0.104	0.139	0.889
LEF_27	0.046	0.088	0.520	0.603
LEF_42	0.076	0.082	0.932	0.351
LEF_1	0.105	0.116	0.900	0.368
LEF_15	-0.045	0.059	-0.764	0.445
LEF_29	-0.010	0.067	-0.149	0.882
LEF_2	0.087	0.098	0.891	0.373
LEF_16	-0.076	0.033	-2.303	0.021
LEF_30	0.071	0.092	0.774	0.439
LEF_3	-0.015	0.066	-0.228	0.820
LEF_17	0.003	0.067	0.044	0.965
LEF_31	0.091	0.083	1.103	0.270
LEF_6	-0.012	0.048	-0.250	0.803
LEF_20	0.052	0.114	0.453	0.651
LEF_33	0.004	0.082	0.046	0.963
LEF_7	0.099	0.069	1.428	0.153
LEF_21	0.099	0.045	2.195	0.028

	Estimate	S.E.	Est./S.E.	P-Value
LEF_35	-0.026	0.059	-0.437	0.662
LEF_9	0.128	0.077	1.671	0.095
LEF_22	0.166	0.088	1.893	0.058
LEF_36	0.040	0.072	0.558	0.577
LEF_10	0.075	0.042	1.768	0.077
LEF_24	-0.017	0.070	-0.238	0.812
LEF_37	-0.029	0.046	-0.624	0.533
LEF_11	-0.058	0.061	-0.960	0.337
LEF_25	0.052	0.046	1.134	0.257
LEF_39	-0.009	0.049	-0.193	0.847
LEF_13	-0.027	0.057	-0.479	0.632
LEF_26	0.056	0.038	1.459	0.144
LEF_40	0.059	0.059	1.007	0.314
LEF_14	0.010	0.053	0.187	0.852
LEF_28	0.045	0.045	0.993	0.321
LEF_41	0.039	0.055	0.709	0.478
OPENTH BY				
LEF_6	0.774	0.085	9.065	0.000
LEF_20	0.552	0.096	5.719	0.000
LEF_33	0.336	0.100	3.354	0.001
LEF_12	0.066	0.054	1.238	0.216
LEF_27	0.053	0.044	1.212	0.226
LEF_42	0.009	0.054	0.167	0.868
LEF_1	0.119	0.040	2.943	0.003
LEF_15	-0.012	0.040	-0.312	0.755
LEF_29	-0.038	0.037	-1.031	0.303
LEF_2	0.073	0.045	1.641	0.101
LEF_16	-0.020	0.027	-0.738	0.461
LEF_30	0.010	0.056	0.182	0.855
LEF_3	0.030	0.035	0.840	0.401
LEF_17	0.005	0.035	0.136	0.891
LEF_31	-0.095	0.042	-2.246	0.025
LEF_5	0.103	0.058	1.759	0.079
LEF_18	-0.156	0.052	-2.992	0.003
LEF_32	0.083	0.084	0.984	0.325
LEF_7	0.294	0.045	6.492	0.000
LEF_21	0.022	0.050	0.435	0.664
LEF_35	0.012	0.052	0.229	0.819
LEF_9	0.063	0.040	1.553	0.120
LEF_22	0.063	0.040	1.582	0.114
LEF_36	0.031	0.042	0.730	0.465
LEF_10	0.001	0.025	0.026	0.979
LEF_24	0.049	0.032	1.542	0.123
LEF_37	0.006	0.031	0.211	0.833
LEF_11	0.010	0.038	0.268	0.789
LEF_25	0.003	0.038	0.088	0.930
LEF_39	-0.019	0.035	-0.534	0.593
LEF_13	-0.008	0.034	-0.249	0.803
LEF_26	0.040	0.030	1.321	0.187
LEF_40	-0.023	0.037	-0.625	0.532
LEF_14	0.012	0.036	0.325	0.745
LEF_28	0.040	0.035	1.127	0.260
LEF_41	0.069	0.040	1.714	0.087
QUALSEEK BY				
LEF_7	0.372	0.072	5.173	0.000
LEF_21	0.737	0.082	8.959	0.000
LEF_35	0.677	0.092	7.368	0.000

	Estimate	S.E.	Est./S.E.	P-Value
LEF_12	-0.044	0.069	-0.639	0.523
LEF_27	-0.052	0.047	-1.119	0.263
LEF_42	0.025	0.054	0.461	0.645
LEF_1	-0.128	0.050	-2.541	0.011
LEF_15	0.046	0.037	1.253	0.210
LEF_29	0.066	0.052	1.263	0.206
LEF_2	-0.053	0.054	-0.996	0.319
LEF_16	0.019	0.024	0.783	0.433
LEF_30	-0.034	0.062	-0.555	0.579
LEF_3	-0.038	0.041	-0.932	0.351
LEF_17	0.066	0.037	1.797	0.072
LEF_31	-0.015	0.050	-0.309	0.758
LEF_5	0.002	0.057	0.042	0.967
LEF_18	0.071	0.085	0.836	0.403
LEF_32	0.028	0.075	0.365	0.715
LEF_6	0.007	0.048	0.135	0.892
LEF_20	0.145	0.077	1.888	0.059
LEF_33	0.128	0.084	1.526	0.127
LEF_9	0.108	0.058	1.857	0.063
LEF_22	0.120	0.047	2.541	0.011
LEF_36	0.239	0.045	5.339	0.000
LEF_10	-0.044	0.045	-0.991	0.322
LEF_24	-0.033	0.044	-0.744	0.457
LEF_37	0.064	0.037	1.729	0.084
LEF_11	0.004	0.057	0.069	0.945
LEF_25	-0.027	0.039	-0.683	0.495
LEF_39	0.035	0.048	0.731	0.465
LEF_13	0.139	0.052	2.684	0.007
LEF_26	0.005	0.032	0.157	0.875
LEF_40	0.047	0.048	0.961	0.336
LEF_14	0.022	0.060	0.373	0.709
LEF_28	-0.043	0.036	-1.219	0.223
LEF_41	-0.003	0.054	-0.053	0.958
SELFCON BY				
LEF_9	0.465	0.129	3.598	0.000
LEF_22	0.336	0.075	4.488	0.000
LEF_36	0.438	0.100	4.384	0.000
LEF_12	0.348	0.102	3.414	0.001
LEF_27	0.328	0.086	3.835	0.000
LEF_42	0.312	0.106	2.952	0.003
LEF_1	-0.040	0.062	-0.647	0.518
LEF_15	-0.038	0.107	-0.357	0.721
LEF_29	-0.017	0.123	-0.139	0.890
LEF_2	0.033	0.070	0.473	0.636
LEF_16	-0.065	0.067	-0.973	0.331
LEF_30	0.017	0.128	0.129	0.897
LEF_3	-0.032	0.055	-0.589	0.556
LEF_17	-0.009	0.067	-0.138	0.890
LEF_31	-0.011	0.095	-0.116	0.908
LEF_5	-0.003	0.082	-0.032	0.974
LEF_18	0.002	0.101	0.017	0.987
LEF_32	0.056	0.160	0.348	0.728
LEF_6	0.018	0.041	0.438	0.662
LEF_20	-0.056	0.058	-0.953	0.340
LEF_33	0.058	0.158	0.369	0.712
LEF_7	0.023	0.101	0.229	0.819
LEF_21	0.059	0.053	1.109	0.268
LEF_35	0.099	0.078	1.260	0.208

	Estimate	S.E.	Est./S.E.	P-Value
LEF_10	-0.024	0.129	-0.182	0.856
LEF_24	0.003	0.052	0.056	0.956
LEF_37	-0.026	0.065	-0.391	0.696
LEF_11	0.019	0.120	0.160	0.873
LEF_25	0.063	0.046	1.370	0.171
LEF_39	-0.043	0.068	-0.629	0.529
LEF_13	-0.041	0.148	-0.277	0.782
LEF_26	-0.051	0.045	-1.120	0.263
LEF_40	-0.096	0.062	-1.560	0.119
LEF_14	-0.014	0.178	-0.080	0.937
LEF_28	0.031	0.043	0.734	0.463
LEF_41	-0.081	0.065	-1.242	0.214
SOCIALEF BY				
LEF_10	0.911	0.034	27.098	0.000
LEF_24	0.837	0.039	21.483	0.000
LEF_37	0.817	0.033	24.407	0.000
LEF_12	0.150	0.033	4.504	0.000
LEF_27	0.086	0.030	2.840	0.005
LEF_42	0.058	0.030	1.896	0.058
LEF_1	-0.015	0.030	-0.485	0.628
LEF_15	-0.008	0.023	-0.332	0.740
LEF_29	0.007	0.020	0.357	0.721
LEF_2	-0.098	0.027	-3.557	0.000
LEF_16	0.032	0.018	1.821	0.069
LEF_30	0.027	0.032	0.851	0.395
LEF_3	-0.043	0.028	-1.517	0.129
LEF_17	-0.009	0.026	-0.332	0.740
LEF_31	0.148	0.038	3.868	0.000
LEF_5	-0.018	0.039	-0.467	0.640
LEF_18	0.043	0.070	0.605	0.545
LEF_32	0.081	0.038	2.152	0.031
LEF_6	0.003	0.028	0.123	0.902
LEF_20	0.106	0.035	3.025	0.002
LEF_33	0.039	0.037	1.039	0.299
LEF_7	0.007	0.039	0.178	0.859
LEF_21	0.011	0.022	0.525	0.599
LEF_35	0.007	0.030	0.247	0.805
LEF_9	0.060	0.038	1.550	0.121
LEF_22	0.020	0.036	0.571	0.568
LEF_36	0.107	0.033	3.287	0.001
LEF_11	0.008	0.027	0.278	0.781
LEF_25	0.068	0.027	2.504	0.012
LEF_39	0.003	0.029	0.115	0.909
LEF_13	0.023	0.030	0.775	0.439
LEF_26	-0.012	0.022	-0.554	0.580
LEF_40	0.074	0.025	2.929	0.003
LEF_14	-0.004	0.022	-0.168	0.866
LEF_28	-0.006	0.023	-0.271	0.786
LEF_41	0.015	0.023	0.655	0.513
STRESSMG BY				
LEF_11	0.747	0.084	8.880	0.000
LEF_25	0.661	0.073	9.107	0.000
LEF_39	0.782	0.068	11.436	0.000
LEF_12	0.113	0.054	2.089	0.037
LEF_27	0.020	0.043	0.478	0.633
LEF_42	-0.027	0.043	-0.632	0.527
LEF_1	0.066	0.042	1.580	0.114

	Estimate	S.E.	Est./S.E.	P-Value
LEF_15	-0.041	0.036	-1.130	0.258
LEF_29	-0.003	0.032	-0.098	0.922
LEF_2	0.119	0.056	2.134	0.033
LEF_16	-0.058	0.031	-1.878	0.060
LEF_30	0.039	0.063	0.623	0.534
LEF_3	0.033	0.028	1.186	0.236
LEF_17	-0.044	0.034	-1.286	0.198
LEF_31	0.025	0.043	0.577	0.564
LEF_5	-0.032	0.032	-0.989	0.323
LEF_18	0.037	0.049	0.769	0.442
LEF_32	-0.045	0.046	-0.991	0.322
LEF_6	0.022	0.032	0.670	0.503
LEF_20	-0.041	0.039	-1.061	0.289
LEF_33	0.066	0.050	1.324	0.186
LEF_7	0.049	0.042	1.182	0.237
LEF_21	0.028	0.030	0.931	0.352
LEF_35	-0.014	0.038	-0.367	0.713
LEF_9	0.074	0.047	1.579	0.114
LEF_22	0.181	0.043	4.213	0.000
LEF_36	0.051	0.046	1.124	0.261
LEF_10	-0.052	0.030	-1.761	0.078
LEF_24	0.044	0.027	1.667	0.095
LEF_37	0.065	0.030	2.177	0.029
LEF_13	0.017	0.042	0.400	0.689
LEF_26	0.033	0.032	1.034	0.301
LEF_40	0.065	0.040	1.630	0.103
LEF_14	-0.002	0.052	-0.045	0.964
LEF_28	0.028	0.051	0.557	0.577
LEF_41	0.150	0.047	3.189	0.001
TIMEEFFI BY				
LEF_13	0.681	0.044	15.540	0.000
LEF_26	0.777	0.038	20.513	0.000
LEF_40	0.737	0.054	13.589	0.000
LEF_12	0.142	0.037	3.861	0.000
LEF_27	0.300	0.032	9.255	0.000
LEF_42	0.242	0.036	6.649	0.000
LEF_1	0.016	0.028	0.581	0.561
LEF_15	-0.018	0.033	-0.552	0.581
LEF_29	-0.023	0.031	-0.724	0.469
LEF_2	0.048	0.026	1.885	0.059
LEF_16	-0.031	0.022	-1.414	0.157
LEF_30	-0.010	0.040	-0.253	0.800
LEF_3	0.003	0.021	0.164	0.869
LEF_17	-0.060	0.025	-2.400	0.016
LEF_31	0.037	0.034	1.096	0.273
LEF_5	0.033	0.044	0.745	0.456
LEF_18	-0.035	0.045	-0.776	0.438
LEF_32	0.072	0.052	1.386	0.166
LEF_6	-0.021	0.025	-0.817	0.414
LEF_20	-0.026	0.032	-0.800	0.424
LEF_33	0.071	0.045	1.564	0.118
LEF_7	0.115	0.037	3.086	0.002
LEF_21	0.042	0.032	1.297	0.195
LEF_35	0.070	0.033	2.159	0.031
LEF_9	-0.071	0.031	-2.310	0.021
LEF_22	-0.018	0.033	-0.543	0.587
LEF_36	-0.061	0.031	-1.955	0.051
LEF_10	-0.002	0.034	-0.051	0.959

	Estimate	S.E.	Est./S.E.	P-Value
LEF_24	0.016	0.023	0.713	0.476
LEF_37	-0.010	0.024	-0.407	0.684
LEF_11	0.024	0.039	0.616	0.538
LEF_25	-0.016	0.024	-0.664	0.507
LEF_39	0.043	0.032	1.326	0.185
LEF_14	0.066	0.049	1.342	0.180
LEF_28	-0.035	0.027	-1.323	0.186
LEF_41	0.013	0.032	0.404	0.686
COPINGCH BY				
LEF_14	0.768	0.080	9.652	0.000
LEF_28	0.809	0.088	9.233	0.000
LEF_41	0.700	0.081	8.651	0.000
LEF_12	0.062	0.049	1.270	0.204
LEF_27	0.080	0.045	1.798	0.072
LEF_42	0.144	0.051	2.833	0.005
LEF_1	-0.107	0.044	-2.456	0.014
LEF_15	0.030	0.031	0.954	0.340
LEF_29	0.040	0.034	1.175	0.240
LEF_2	-0.129	0.051	-2.548	0.011
LEF_16	-0.011	0.035	-0.310	0.756
LEF_30	0.177	0.056	3.152	0.002
LEF_3	-0.043	0.044	-0.982	0.326
LEF_17	0.057	0.036	1.567	0.117
LEF_31	0.013	0.048	0.261	0.794
LEF_5	0.036	0.055	0.656	0.512
LEF_18	0.014	0.064	0.213	0.832
LEF_32	0.074	0.059	1.254	0.210
LEF_6	0.021	0.040	0.517	0.606
LEF_20	0.088	0.056	1.584	0.113
LEF_33	0.118	0.055	2.138	0.033
LEF_7	-0.035	0.043	-0.801	0.423
LEF_21	-0.032	0.031	-1.033	0.302
LEF_35	0.050	0.040	1.262	0.207
LEF_9	0.043	0.050	0.850	0.395
LEF_22	-0.016	0.044	-0.371	0.710
LEF_36	0.071	0.041	1.728	0.084
LEF_10	0.011	0.023	0.486	0.627
LEF_24	-0.037	0.030	-1.241	0.215
LEF_37	0.003	0.031	0.083	0.934
LEF_11	0.045	0.048	0.927	0.354
LEF_25	0.063	0.052	1.219	0.223
LEF_39	0.046	0.053	0.876	0.381
LEF_13	0.083	0.039	2.131	0.033
LEF_26	0.002	0.034	0.061	0.951
LEF_40	0.020	0.045	0.451	0.652
ILOC BY				
LEF_4	0.791	0.018	43.456	0.000
LEF_19	0.838	0.017	48.446	0.000
LEF_34	0.770	0.023	33.442	0.000
LEF_8	0.023	0.020	1.160	0.246
LEF_23R	0.059	0.017	3.395	0.001
LEF_38R	-0.078	0.017	-4.599	0.000
ELOC BY				
LEF_8	0.717	0.024	29.305	0.000
LEF_23R	0.741	0.023	32.287	0.000
LEF_38R	0.766	0.024	32.564	0.000

	Estimate	S.E.	Est./S.E.	P-Value
LEF_4	0.006	0.017	0.354	0.723
LEF_19	-0.007	0.014	-0.487	0.626
LEF_34	0.006	0.022	0.259	0.796
AUTONOM BY				
PSNA1	0.760	0.021	37.017	0.000
PSNA2	0.866	0.013	67.607	0.000
PSNA3	0.790	0.019	42.026	0.000
PSNA4	0.732	0.020	36.951	0.000
RELATED BY				
PSNR5	0.773	0.023	33.186	0.000
PSNR6	0.862	0.012	69.228	0.000
PSNR7	0.845	0.014	60.727	0.000
PSNR8	0.645	0.028	22.709	0.000
COMPET BY				
PSNC9	0.821	0.013	61.918	0.000
PSNC10	0.862	0.011	81.941	0.000
PSNC11	0.826	0.014	57.653	0.000
PSNC12	0.836	0.012	67.051	0.000
ACADEM BY				
SDQAC1	0.704	0.030	23.679	0.000
SDQAC2	0.832	0.022	37.843	0.000
SDQAC3	0.885	0.018	49.406	0.000
SDQAC4	0.778	0.022	34.626	0.000
SDQAC5	0.809	0.023	35.693	0.000
SDQPH1	0.010	0.034	0.299	0.765
SDQPH2	-0.030	0.026	-1.139	0.255
SDQPH3	-0.020	0.019	-1.066	0.286
SDQPH4	-0.004	0.023	-0.167	0.867
SDQPH5	0.055	0.029	1.906	0.057
SDQSE1	-0.087	0.024	-3.683	0.000
SDQSE2	0.015	0.022	0.655	0.512
SDQSE3	0.157	0.025	6.213	0.000
SDQSE4	0.087	0.030	2.927	0.003
SDQSE5	-0.099	0.026	-3.855	0.000
SDQSE6	-0.126	0.025	-5.124	0.000
SDQSE7	0.119	0.024	4.925	0.000
SDQPA1	-0.001	0.018	-0.068	0.946
SDQPA2	0.000	0.018	0.003	0.998
SDQPA3	0.047	0.020	2.332	0.020
SDQPA4R	-0.008	0.033	-0.234	0.815
SDQPS1	0.152	0.025	6.038	0.000
SDQPS2	-0.011	0.025	-0.418	0.676
SDQPS3	0.052	0.025	2.040	0.041
SDQPS4	-0.005	0.024	-0.215	0.830
SDQPS5	0.038	0.024	1.576	0.115
SDQPS6	-0.028	0.022	-1.269	0.204
SDQPS7	-0.075	0.021	-3.629	0.000
SDQPS8	-0.036	0.022	-1.642	0.101
SDQR2R	-0.007	0.026	-0.282	0.778
SDQR3	0.053	0.042	1.259	0.208
SDQR6R	0.006	0.019	0.322	0.747
SDQR7R	-0.005	0.020	-0.263	0.792
SDQR9	0.079	0.047	1.680	0.093
SDQR1R	0.034	0.036	0.943	0.345
SDQR4	0.026	0.017	1.504	0.133

	Estimate	S.E.	Est./S.E.	P-Value
SDQR5	-0.034	0.019	-1.793	0.073
SDQR8R	0.003	0.031	0.095	0.925
PHYSIC BY				
SDQPH1	0.676	0.047	14.508	0.000
SDQPH2	0.813	0.027	29.681	0.000
SDQPH3	0.945	0.021	44.763	0.000
SDQPH4	0.850	0.020	42.557	0.000
SDQPH5	0.650	0.032	20.267	0.000
SDQAC1	0.024	0.029	0.804	0.421
SDQAC2	-0.005	0.021	-0.260	0.795
SDQAC3	0.028	0.022	1.262	0.207
SDQAC4	0.019	0.029	0.658	0.511
SDQAC5	-0.053	0.023	-2.359	0.018
SDQSE1	-0.004	0.029	-0.128	0.898
SDQSE2	0.026	0.023	1.146	0.252
SDQSE3	0.006	0.022	0.267	0.790
SDQSE4	-0.008	0.026	-0.309	0.758
SDQSE5	0.023	0.031	0.734	0.463
SDQSE6	0.005	0.030	0.170	0.865
SDQSE7	-0.016	0.030	-0.521	0.602
SDQPA1	-0.002	0.016	-0.093	0.926
SDQPA2	0.004	0.021	0.202	0.840
SDQPA3	0.047	0.027	1.756	0.079
SDQPA4R	-0.118	0.037	-3.157	0.002
SDQPS1	0.041	0.029	1.394	0.163
SDQPS2	0.013	0.029	0.455	0.649
SDQPS3	-0.006	0.028	-0.206	0.837
SDQPS4	0.036	0.023	1.600	0.110
SDQPS5	0.022	0.022	0.999	0.318
SDQPS6	-0.024	0.026	-0.908	0.364
SDQPS7	-0.032	0.021	-1.508	0.132
SDQPS8	-0.017	0.021	-0.807	0.419
SDQR2R	0.020	0.027	0.745	0.456
SDQR3	0.151	0.062	2.454	0.014
SDQR6R	0.057	0.022	2.631	0.009
SDQR7R	0.064	0.021	3.070	0.002
SDQR9	0.144	0.062	2.320	0.020
SDQR1R	-0.057	0.042	-1.354	0.176
SDQR4	0.011	0.018	0.577	0.564
SDQR5	0.028	0.022	1.276	0.202
SDQR8R	-0.039	0.051	-0.770	0.442
SELFESTE BY				
SDQSE1	0.854	0.041	21.022	0.000
SDQSE2	0.787	0.037	21.222	0.000
SDQSE3	0.759	0.039	19.722	0.000
SDQSE4	0.746	0.042	17.585	0.000
SDQSE5	0.763	0.043	17.857	0.000
SDQSE6	0.814	0.042	19.529	0.000
SDQSE7	0.743	0.038	19.472	0.000
SDQAC1	-0.007	0.044	-0.157	0.875
SDQAC2	0.049	0.030	1.668	0.095
SDQAC3	-0.012	0.027	-0.446	0.656
SDQAC4	-0.116	0.029	-4.071	0.000
SDQAC5	0.137	0.033	4.124	0.000
SDQPH1	0.072	0.056	1.279	0.201
SDQPH2	-0.063	0.038	-1.637	0.102
SDQPH3	-0.039	0.029	-1.330	0.183

	Estimate	S.E.	Est./S.E.	P-Value
SDQPH4	-0.012	0.029	-0.423	0.672
SDQPH5	0.096	0.042	2.274	0.023
SDQPA1	0.036	0.023	1.577	0.115
SDQPA2	-0.007	0.028	-0.267	0.789
SDQPA3	0.002	0.034	0.068	0.946
SDQPA4R	-0.008	0.050	-0.165	0.869
SDQPS1	0.004	0.046	0.081	0.935
SDQPS2	0.039	0.035	1.136	0.256
SDQPS3	0.079	0.044	1.788	0.074
SDQPS4	-0.009	0.034	-0.271	0.786
SDQPS5	0.005	0.030	0.167	0.867
SDQPS6	-0.007	0.036	-0.191	0.848
SDQPS7	-0.046	0.026	-1.792	0.073
SDQPS8	0.046	0.029	1.607	0.108
SDQR2R	0.015	0.038	0.392	0.695
SDQR3	0.124	0.076	1.635	0.102
SDQR6R	0.107	0.031	3.435	0.001
SDQR7R	0.012	0.028	0.420	0.675
SDQR9	0.091	0.077	1.189	0.235
SDQR1R	-0.048	0.053	-0.910	0.363
SDQR4	0.044	0.023	1.903	0.057
SDQR5	0.006	0.028	0.211	0.833
SDQR8R	-0.007	0.059	-0.121	0.904
PARENTS BY				
SDQPA1	0.872	0.017	50.652	0.000
SDQPA2	0.906	0.021	42.762	0.000
SDQPA3	0.854	0.020	42.082	0.000
SDQPA4R	0.453	0.038	11.897	0.000
SDQAC1	0.028	0.025	1.108	0.268
SDQAC2	0.034	0.020	1.679	0.093
SDQAC3	0.021	0.018	1.125	0.261
SDQAC4	-0.023	0.018	-1.296	0.195
SDQAC5	0.000	0.019	-0.019	0.985
SDQPH1	-0.084	0.029	-2.879	0.004
SDQPH2	0.090	0.023	3.978	0.000
SDQPH3	0.037	0.017	2.110	0.035
SDQPH4	-0.050	0.020	-2.522	0.012
SDQPH5	-0.024	0.029	-0.841	0.400
SDQSE1	0.114	0.025	4.512	0.000
SDQSE2	-0.015	0.023	-0.626	0.532
SDQSE3	-0.054	0.023	-2.292	0.022
SDQSE4	-0.021	0.026	-0.822	0.411
SDQSE5	-0.012	0.027	-0.453	0.650
SDQSE6	0.071	0.028	2.510	0.012
SDQSE7	-0.048	0.021	-2.317	0.021
SDQPS1	0.039	0.025	1.593	0.111
SDQPS2	-0.015	0.020	-0.750	0.453
SDQPS3	0.000	0.022	0.012	0.991
SDQPS4	0.023	0.023	1.015	0.310
SDQPS5	0.014	0.021	0.660	0.509
SDQPS6	0.005	0.021	0.213	0.831
SDQPS7	-0.024	0.019	-1.255	0.210
SDQPS8	0.002	0.019	0.092	0.926
SDQR2R	0.010	0.022	0.438	0.661
SDQR3	0.073	0.038	1.915	0.056
SDQR6R	0.047	0.020	2.325	0.020
SDQR7R	0.058	0.018	3.277	0.001
SDQR9	0.025	0.039	0.633	0.527

	Estimate	S.E.	Est./S.E.	P-Value
SDQR1R	-0.005	0.027	-0.186	0.853
SDQR4	0.025	0.016	1.550	0.121
SDQR5	-0.018	0.016	-1.127	0.260
SDQR8R	0.061	0.029	2.130	0.033
PROB BY				
SDQPS1	0.642	0.034	18.828	0.000
SDQPS2	0.755	0.022	33.690	0.000
SDQPS3	0.730	0.033	22.040	0.000
SDQPS4	0.793	0.026	30.118	0.000
SDQPS5	0.785	0.026	30.633	0.000
SDQPS6	0.827	0.021	38.639	0.000
SDQPS7	0.896	0.021	42.483	0.000
SDQPS8	0.802	0.024	33.673	0.000
SDQAC1	0.088	0.033	2.661	0.008
SDQAC2	-0.063	0.022	-2.872	0.004
SDQAC3	-0.039	0.018	-2.092	0.036
SDQAC4	0.156	0.026	5.920	0.000
SDQAC5	-0.035	0.024	-1.473	0.141
SDQPH1	-0.018	0.036	-0.503	0.615
SDQPH2	0.053	0.030	1.773	0.076
SDQPH3	-0.015	0.020	-0.717	0.474
SDQPH4	0.009	0.025	0.350	0.726
SDQPH5	-0.004	0.032	-0.137	0.891
SDQSE1	-0.071	0.032	-2.217	0.027
SDQSE2	0.032	0.025	1.283	0.199
SDQSE3	0.017	0.025	0.686	0.493
SDQSE4	0.002	0.032	0.062	0.950
SDQSE5	0.066	0.028	2.345	0.019
SDQSE6	0.040	0.030	1.303	0.193
SDQSE7	0.054	0.028	1.927	0.054
SDQPA1	0.031	0.016	1.916	0.055
SDQPA2	0.035	0.018	1.951	0.051
SDQPA3	-0.013	0.024	-0.569	0.569
SDQPA4R	-0.046	0.033	-1.409	0.159
SDQR2R	0.009	0.028	0.327	0.743
SDQR3	-0.017	0.049	-0.353	0.724
SDQR6R	-0.023	0.022	-1.083	0.279
SDQR7R	0.042	0.019	2.244	0.025
SDQR9	-0.017	0.051	-0.337	0.736
SDQR1R	0.021	0.038	0.555	0.579
SDQR4	0.037	0.021	1.748	0.081
SDQR5	0.039	0.024	1.630	0.103
SDQR8R	0.023	0.033	0.703	0.482
SSEXRELS BY				
SDQR2R	0.659	0.026	25.333	0.000
SDQR3	0.269	0.049	5.441	0.000
SDQR6R	0.750	0.025	29.581	0.000
SDQR7R	0.839	0.019	43.177	0.000
SDQR9	0.226	0.051	4.460	0.000
SDQAC1	0.069	0.022	3.097	0.002
SDQAC2	0.004	0.018	0.240	0.810
SDQAC3	-0.002	0.017	-0.100	0.920
SDQAC4	-0.063	0.020	-3.203	0.001
SDQAC5	-0.020	0.016	-1.258	0.208
SDQPH1	-0.019	0.022	-0.883	0.377
SDQPH2	0.145	0.027	5.446	0.000
SDQPH3	0.047	0.023	2.047	0.041

	Estimate	S.E.	Est./S.E.	P-Value
SDQPH4	-0.046	0.024	-1.914	0.056
SDQPH5	-0.018	0.023	-0.774	0.439
SDQSE1	-0.010	0.023	-0.442	0.659
SDQSE2	0.062	0.020	3.169	0.002
SDQSE3	0.034	0.020	1.717	0.086
SDQSE4	-0.002	0.022	-0.103	0.918
SDQSE5	-0.058	0.025	-2.356	0.018
SDQSE6	0.011	0.023	0.497	0.619
SDQSE7	0.033	0.023	1.417	0.157
SDQPA1	-0.045	0.017	-2.638	0.008
SDQPA2	-0.061	0.018	-3.436	0.001
SDQPA3	-0.092	0.022	-4.256	0.000
SDQPA4R	0.294	0.034	8.666	0.000
SDQPS1	0.004	0.022	0.193	0.847
SDQPS2	-0.003	0.018	-0.172	0.864
SDQPS3	0.047	0.020	2.404	0.016
SDQPS4	0.007	0.018	0.381	0.703
SDQPS5	0.000	0.017	-0.018	0.985
SDQPS6	0.002	0.019	0.111	0.911
SDQPS7	-0.020	0.016	-1.278	0.201
SDQPS8	-0.029	0.016	-1.770	0.077
SDQR1R	0.343	0.039	8.902	0.000
SDQR4	-0.092	0.026	-3.543	0.000
SDQR5	-0.180	0.029	-6.124	0.000
SDQR8R	0.625	0.039	15.930	0.000
OSEXRELS BY				
SDQR1R	0.465	0.046	10.030	0.000
SDQR4	0.848	0.029	29.743	0.000
SDQR5	0.874	0.031	28.406	0.000
SDQR8R	0.316	0.049	6.405	0.000
SDQAC1	-0.021	0.028	-0.743	0.457
SDQAC2	0.001	0.038	0.015	0.988
SDQAC3	-0.003	0.029	-0.113	0.910
SDQAC4	0.002	0.025	0.065	0.948
SDQAC5	0.015	0.028	0.545	0.586
SDQPH1	0.009	0.034	0.275	0.784
SDQPH2	-0.089	0.028	-3.201	0.001
SDQPH3	-0.034	0.019	-1.808	0.071
SDQPH4	0.027	0.022	1.234	0.217
SDQPH5	0.018	0.031	0.582	0.560
SDQSE1	0.007	0.024	0.276	0.782
SDQSE2	-0.043	0.021	-2.075	0.038
SDQSE3	-0.011	0.020	-0.564	0.573
SDQSE4	-0.037	0.025	-1.473	0.141
SDQSE5	0.075	0.028	2.699	0.007
SDQSE6	-0.027	0.028	-0.950	0.342
SDQSE7	0.000	0.024	-0.002	0.999
SDQPA1	0.011	0.017	0.656	0.512
SDQPA2	-0.011	0.018	-0.605	0.545
SDQPA3	0.050	0.023	2.161	0.031
SDQPA4R	-0.102	0.035	-2.896	0.004
SDQPS1	-0.017	0.027	-0.632	0.527
SDQPS2	0.061	0.026	2.309	0.021
SDQPS3	-0.079	0.026	-3.097	0.002
SDQPS4	0.001	0.022	0.056	0.955
SDQPS5	0.002	0.023	0.086	0.931
SDQPS6	0.001	0.022	0.045	0.964
SDQPS7	0.042	0.021	1.978	0.048

	Estimate	S.E.	Est./S.E.	P-Value
SDQPS8	0.016	0.021	0.741	0.459
SDQR2R	0.173	0.030	5.754	0.000
SDQR3	0.299	0.056	5.304	0.000
SDQR6R	-0.007	0.024	-0.281	0.779
SDQR7R	-0.014	0.022	-0.667	0.505
SDQR9	0.257	0.054	4.764	0.000

Appendix 6: *Ms*, *SDs*, Skewness and Kurtosis Descriptive Statistics

Table 46 *Overall Ms, SDs, Skew and Kurt for the Treatment and WLC Conditions at T0, T1, T2 and T3 for ROPELOC, SDQ, BPNSFS, MES, GHQ-12, and One-Factor Scales*

		T0 (*N* =346)				T1 (*N* =378)				T2 (*N* =380)				T3 (*N* =156)			
Scale		*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*
ROPELOC																	
SC	Treat	6.12	1.25	-0.56	-0.22	6.81	1.16	-0.47	-0.15	6.48	1.11	-0.84	0.72	6.38	1.05	-0.51	0.35
	WLC	6.10	1.12	-0.75	0.62	6.06	1.33	-0.84	0.90	6.10	1.23	-1.12	2.48	--	--	--	--
TE	Treat	5.39	1.53	-0.47	-0.10	5.31	1.39	-0.45	0.13	5.65	1.43	-0.67	0.19	5.77	1.35	-0.64	0.84
	WLC	5.43	1.42	-0.53	0.06	5.50	1.40	-0.33	-0.50	5.36	1.49	-0.74	0.28	--	--	--	--
AI	Treat	6.20	1.23	-0.77	1.10	6.30	1.15	-1.00	2.10	6.59	1.05	-0.85	0.36	6.43	1.17	-0.92	1.52
	WLC	6.35	1.21	-0.74	0.47	6.14	1.31	-0.70	0.53	6.35	1.25	-1.01	1.61	--	--	--	--
CT	Treat	5.97	1.40	-1.06	1.36	6.00	1.38	-0.80	0.72	6.16	1.17	-0.75	0.82	6.21	1.24	-0.69	0.13
	WLC	6.13	1.40	-1.06	1.27	6.01	1.44	-1.16	1.72	6.09	1.45	-1.03	0.98	--	--	--	--
OT	Treat	6.17	1.08	-0.39	0.22	6.14	1.03	-0.33	0.26	6.29	0.97	-0.67	1.15	6.32	1.09	-0.99	1.61
	WLC	6.13	1.14	-0.81	1.06	6.11	1.12	-0.43	0.66	6.23	1.12	-0.87	2.06	--	--	--	--
QS	Treat	6.51	1.10	-0.75	0.21	6.42	1.16	-0.77	0.56	6.66	1.07	-0.67	-0.16	6.52	1.12	-0.58	-0.44
	WLC	6.57	1.10	-0.81	-0.04	6.45	1.19	-0.97	1.70	6.48	1.12	-0.69	0.17	--	--	--	--
SF	Treat	5.49	1.41	-0.42	0.03	5.51	1.30	-0.39	-0.06	5.82	1.28	-0.52	-0.13	5.99	1.28	-0.70	0.42
	WLC	5.50	1.32	-0.59	0.24	5.47	1.34	-0.50	0.28	5.68	1.40	-0.71	0.86	--	--	--	--
SE	Treat	5.80	1.36	-1.05	1.42	5.76	1.33	-0.72	0.81	6.02	1.34	-0.85	0.88	5.99	1.25	-0.50	0.64
	WLC	5.87	1.26	-0.89	0.98	5.77	1.38	-0.87	1.05	5.91	1.30	-0.80	1.03	--	--	--	--
SM	Treat	5.45	1.57	-0.48	-0.26	5.42	1.36	-0.38	-0.30	5.84	1.37	-0.77	0.86	5.82	1.39	-0.89	0.96
	WLC	5.49	1.38	-0.48	0.02	5.51	1.31	-0.53	0.31	5.65	1.39	-0.91	1.12	--	--	--	--
LA	Treat	5.90	1.41	-0.74	0.71	5.91	1.37	-0.60	0.28	6.03	1.43	-0.66	0.16	6.02	1.45	-0.71	0.26
	WLC	6.12	1.49	-1.13	1.41	5.90	1.47	-0.89	0.86	6.14	1.41	-1.14	1.80	--	--	--	--
CH	Treat	5.66	1.38	-0.25	-0.56	5.59	1.27	-0.37	-0.04	6.03	1.33	-0.71	0.19	6.05	1.30	-0.83	0.90
	WLC	5.71	1.33	-0.83	0.61	5.65	1.25	-0.37	-0.28	5.83	1.38	-0.71	0.44	--	--	--	--
Overall	Treat	5.88				5.92				6.14				6.14			
	WLC	5.95				5.87				5.98				--			

Note. ROPELOC factors are: SC = Self Confidence, TE = Time Efficiency, AI = Active Involvement, CT = Cooperative Teamwork, OT = Open Thinking, QS = Quality Seeking, SF = Self Efficacy, SE = Social Effectiveness, SM = Stress Management, LA = Leadership Ability, CH = Coping with Change.

Table 46 *(Continued)*

Scale		T0 (*N* =346)				T1 (*N* =378)				T2 (*N* =380)				T3 (*N* =156)			
		M	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*
EL	Treat	5.00	1.68	-0.32	-0.36	5.38	1.37	-0.16	-0.07	5.67	1.38	-0.45	-0.31	4.91	1.62	-0.02	-0.55
	WLC	4.74	1.45	-0.08	-0.10	5.03	1.50	0.01	-0.35	5.02	1.59	0.02	-0.55	--	--	--	--
IL	Treat	6.51	1.19	-0.76	-0.17	6.60	1.56	-0.82	0.27	6.86	1.23	-1.48	3.55	6.59	1.08	-0.65	-0.24
	WLC	6.46	1.19	-0.51	-0.46	6.52	1.27	-1.05	2.07	6.59	1.18	-1.12	2.35	--	--	--	--
Overall	Treat	5.76				5.99				6.27				5.75			
	WLC	5.60				5.78				5.81				--			
SDQ																	
Acad	Treat	4.51	0.88	-0.76	1.37	4.38	0.86	-0.59	0.24	4.40	0.92	-0.73	0.53	4.57	0.88	-0.86	1.65
	WLC	4.16	1.01	-0.72	0.54	4.19	0.93	-0.71	0.99	4.21	1.01	-0.64	0.37	--	--	--	--
Esteem	Treat	4.62	0.82	-0.59	0.73	4.66	0.80	-0.59	0.52	4.86	0.73	-0.66	0.59	4.85	0.76	-1.33	4.27
	WLC	4.54	0.86	-0.76	1.20	4.57	0.87	-1.01	2.36	4.59	0.82	-0.93	2.86	--	--	--	--
Physical	Treat	4.23	1.17	-0.71	0.43	4.31	1.09	-0.65	0.33	4.60	1.01	-0.86	0.84	4.55	1.15	-0.98	0.98
	WLC	4.46	1.13	-0.83	0.40	4.32	1.15	-0.70	0.35	4.35	1.15	-0.75	0.42	--	--	--	--
Prnt	Treat	5.17	0.87	-1.17	1.16	5.24	0.86	-1.24	1.26	5.50	0.74	-1.77	3.00	5.13	0.91	-0.78	-0.46
	WLC	4.99	1.08	-1.20	1.41	5.05	1.05	-1.22	1.31	4.98	1.08	-1.15	0.99	--	--	--	--
Sm Sx	Treat	4.66	0.94	-1.01	1.74	4.73	0.82	-0.77	0.56	4.87	0.84	-0.90	1.03	4.58	0.96	-0.49	-0.21
	WLC	4.78	0.88	-0.82	1.16	4.76	0.85	-0.71	0.17	4.72	0.81	-0.70	0.49	--	--	--	--
Op Sx	Treat	4.22	0.96	-0.48	0.44	4.29	0.93	-0.64	0.84	4.43	0.97	-0.43	0.08	4.38	0.94	-0.08	-0.31
	WLC	4.26	0.98	-0.56	0.22	4.40	0.97	-0.57	0.39	4.31	0.94	-0.39	0.11	--	--	--	--
Prob	Treat	4.34	0.78	-0.04	0.05	4.31	0.85	-0.25	0.13	4.47	0.75	-0.30	0.04	4.64	0.81	-0.69	1.03
	WLC	4.33	0.83	-0.29	0.09	4.28	0.89	-0.75	1.85	4.42	0.82	-0.66	1.98	--	--	--	--
Overall	Treat	4.54				4.56				4.73				4.67			
	WLC	4.50				4.51				4.51				--			

Note. ROPELOC factors are: ELoC = External Locus of Control, and ILoC = Internal Locus of Control. SDQ factors are: Acad = General Academic, Esteem = Global Esteem, Physical = Physical Ability, Sm-Sex = Same-Sex Relations, Op-Sex = Opposite-Sex Relations, Prob = Problem Solving.

Table 46 *(Continued)*

Scale		T0 (*N* =346)				T1 (*N* =378)				T2 (*N* =380)				T3 (*N* =156)			
		M	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*
BPNSFS																	
Auton	Treat	4.19	1.01	-0.77	1.01	4.27	0.92	-0.58	0.66	4.34	0.95	-0.66	0.76	4.51	0.89	-1.22	2.99
	WLC	4.30	0.89	-0.54	-0.05	4.40	0.94	-0.98	1.44	4.34	0.90	-0.66	0.74	--	--	--	--
Relat	Treat	4.42	0.92	-0.94	1.20	4.45	0.88	-0.66	0.38	4.68	0.81	-1.19	3.28	4.60	0.79	-0.59	0.94
	WLC	4.49	0.85	-0.53	0.25	4.46	0.94	-1.02	1.61	4.32	0.94	-1.18	1.94	--	--	--	--
Comp	Treat	4.73	0.84	-0.67	0.83	4.67	0.83	-0.44	-0.33	4.87	0.81	-1.03	2.32	4.84	0.74	-0.39	0.07
	WLC	4.67	0.84	-0.58	0.06	4.65	0.90	-0.81	1.07	4.65	0.85	-1.06	2.73	--	--	--	--
Overall	Treat	4.45				4.46				4.63				4.65			
	WLC	4.48				4.50				4.44				--			
MES																	
Thought	Treat	4.78	0.76	-0.61	0.41	4.78	0.73	-0.63	0.66	4.88	0.71	-0.47	0.07	4.82	0.75	-0.91	1.67
	WLC	4.82	0.79	-0.88	1.49	4.85	0.75	-1.09	3.02	4.77	0.85	-1.16	2.37	--	--	--	--
Behaviour	Treat	4.31	0.97	-0.29	-0.08	4.29	0.95	-0.54	0.53	4.36	1.01	-0.68	0.71	4.48	0.91	-0.53	0.03
	WLC	4.38	1.02	-0.45	-0.47	4.36	0.95	-0.44	0.38	4.25	1.02	-0.56	0.41	--	--	--	--
Hampering	Treat	3.59	0.99	-0.37	-0.07	3.72	0.93	-0.24	0.00	3.19	0.87	0.25	-0.38	3.46	1.05	0.12	-0.39
	WLC	3.36	0.97	-0.03	-0.18	3.49	0.90	0.08	-0.05	3.54	0.94	-0.22	-0.12	--	--	--	--
Overall	Treat	4.23				4.26				4.14				4.26			
	WLC	4.18				4.24				4.19				--			
GHQ																	
Positive	Treat	2.08	0.42	0.66	1.22	2.13	0.32	0.74	2.85	2.35	0.38	0.39	-0.04	2.20	0.90	-0.62	0.34
	WLC	2.08	0.41	1.04	4.14	2.14	0.38	0.36	1.62	2.01	0.43	0.71	2.52	--	--	--	--
Negative	Treat	2.12	0.65	0.97	0.87	2.32	0.58	1.12	1.38	2.46	0.50	1.56	3.90	2.19	0.83	0.41	-0.25
	WLC	2.13	0.63	0.77	0.46	2.28	0.57	0.62	0.04	2.13	0.66	0.91	0.88	--	--	--	--
Overall	Treat	2.10				2.23				2.41				2.20			
	WLC	2.11				2.21				2.07				--			

Note. BPNSFS scales are: Auton = Autonomy, Relat = Relatedness and Comp = Competence. MES scales are: Behav = Behaviour, and Hamp = Hampering.

Table 46 *(Continued)*

		T0 (*N* =346)				T1 (*N* =378)				T2 (*N* =380)				T3 (*N* =156)			
Scale		*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*	*M*	*SD*	*Skew*	*Kurt*
1-Fac																	
Well Being	Treat	3.55	0.65	-0.81	2.20	3.70	0.60	-0.68	1.52	3.88	0.65	-0.60	0.73	3.83	0.68	-0.24	-0.12
	WLC	3.63	0.62	-0.19	0.10	3.62	0.66	-0.29	0.70	3.56	0.75	-0.52	0.86	--	--	--	--
LSat	Treat	7.51	2.16	-1.16	0.96	7.51	2.16	-1.33	2.56	8.21	1.64	-1.58	3.06	8.08	1.72	-1.37	2.58
	WLC	7.54	1.93	-0.96	0.66	7.65	1.87	-1.01	1.25	7.40	2.12	-1.15	1.08	--	--	--	--
Gratitide	Treat	4.90	0.75	-0.60	0.13	4.87	0.77	-0.64	1.13	5.20	0.75	-0.92	0.44	4.78	0.70	0.08	-0.61
	WLC	4.73	0.84	-0.90	1.84	4.86	0.83	-1.02	2.41	4.84	0.80	-0.86	1.09	--	--	--	--
Resilience	Treat	4.30	0.95	-0.66	0.54	4.30	0.95	-0.48	0.23	4.52	0.92	-0.75	0.99	4.62	0.86	-0.75	0.87
	WLC	4.28	1.06	-0.93	1.12	4.41	0.90	-0.85	1.71	4.38	0.97	-0.85	0.83	--	--	--	--

Note. 1-Fac = Single Factor scales. Single Factor scales are: LSat – Life Satisfaction.

Appendix 7: Overall Change in Basic Psychological Needs Satisfaction

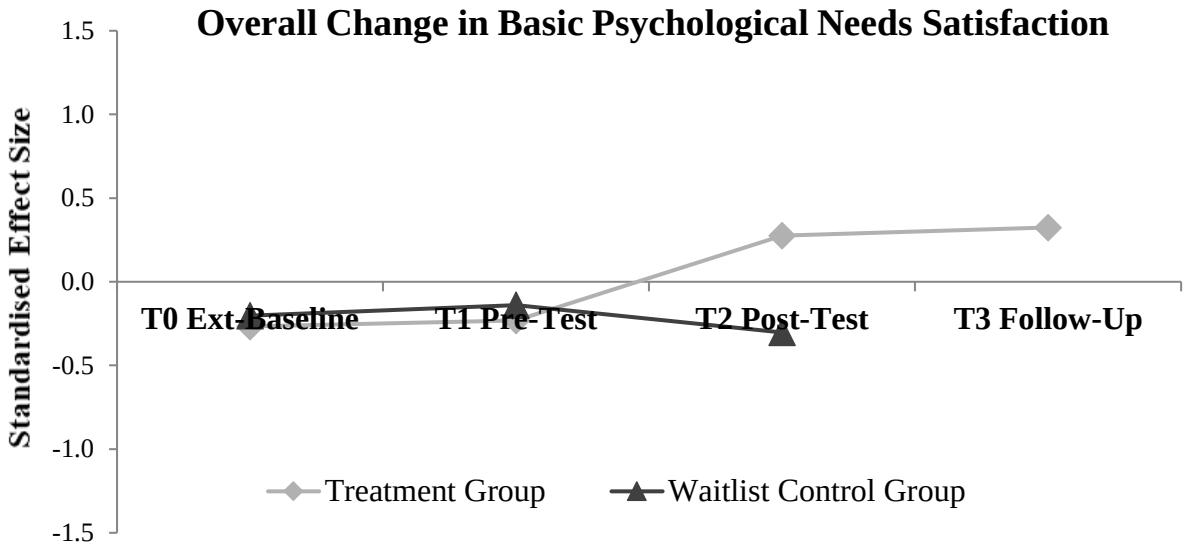


Figure 24. Change in overall basic psychological needs satisfaction for the treatment group and WLC over the four time-waves.

Note. WLC data at T3 was derived from the WLC T2 scores. *Scale* M*s were calculated from BPNSFS factors: Autonomy, Relatedness and Competence.*

Appendix 8: Long-Term Positive Change Analysis

An Alternative Analytic Method, Within-Groups Comparison.

Long-Term Positive Effects (T1 to T3). Hypothesis 2.2 long-term effects analysis followed a between-groups design using the WLC T2 to act as the reference against follow-up T3 data for the treatment group. The following analysis was intended as an alternative and complementary approach to envelop a strong representation of the long-term OAE program effects. The analysis ran regression models that compared treatment group overall scores on each outcome variable between T1 (pre-test) and T3 (6-month follow-up), an approximate 10-month period. The results from this analysis have been presented in Table 47.

Results. Descriptive statistics are shown in Table 36 and ESs with 95% CIs shown in Table 47. For the seven-factor SDQ, statistically significant small to moderate effect sizes were observed for General Academic (ES = .219, $p < .01$), Global Esteem (ES = .262, $p < .01$), Physical Ability (ES = .093, $p < .01$), and Problem Solving (ES = .390, $p < .001$) domains, while a significant decrease was observed in Same-Sex Relations (ES = -.244, $p < .01$). For the ROPELOC, small to medium effects were evident in personal effectiveness areas of Time Efficiency (ES = .302, $p < .001$), Cooperative Teamwork (ES = .173, $p < .05$), Self-Efficacy (ES = .358, $p < .001$), Social Effectiveness (ES = .173, $p < .01$), Stress Management (ES = .299, $p < .001$), Leadership Ability (ES = .158, $p < .05$), and Coping with Change (ES = .346, $p < .001$). In contrast, the ELoC scale evidenced an unfavourable decrease over the T1 to T3 interval (ES = -.322, $p < .001$). Statistically significant small to medium effects were observed in BPNSFS domains of Autonomy (ES = .211, $p < .001$), Relatedness (ES = .175, $p < .01$) and Competence (ES = .169, $p < .01$); and both Well Being (ES = .197, $p < .01$) and Resilience (ES = .297, $p < .001$) single-factor scales. The MES Behaviour (ES = .093, $p < .01$) factor further reflected beneficial

gains overtime, however a statistically significant decrease was seen for Hampering (ES = -.260, $p < .01$).

Conclusion. The alternative within-groups analysis of long-term positive change reflected small to moderate positive changes in areas of self-concept, life effectiveness, psychological needs satisfaction and personal psychological resources. Compared to the primary analysis using the WLC reference group, the alternative within-groups method showed greater sensitivity to detect significant findings, whereby 20 effects were observed (three in a negative direction), compared to the 19 (two in a negative direction) using between-groups methods. The within-subject change in effect sizes over the post-treatment interval (T2 to T3) will now be presented.

Post-Treatment Within-Person Change (T2 to T3). Longitudinal OAE research is largely characterised by pre-post repeated-measures designs absence of reference-group comparisons. The following section replicates this method so to provide an alternative post hoc method of identifying changes in treatment effects over the post-treatment (T2) to follow-up (T3) interval.

Results. Descriptive statistics are shown in Table 36, while standardized ESs with 95% confidence intervals (CIs) for post-treatment and follow-up post hoc analysis for treatment-only participant overall scores from T2 to T3 are presented in Table 48. Within-groups analysis of the T2 to T3 interval for SDQ factors found statistically significant increases in General Academic ($\beta = .178$, $SE = .068$, $p < .01$) and Problem Solving ($\beta = .235$, $SE = .071$, $p < .01$) domains. Significant decreases were observed in Same-Sex Relations ($\beta = -.407$, $SE = .087$, $p < .001$) and Parent-Relations ($\beta = -.301$, $SE = .070$, $p < .01$) domains with moderate to large effect sizes, while non-significant effects for time ($p > .05$) on Global Esteem and Physical Ability scales indicated maintenance of experimental effects.

Table 47 *Effect Sizes and Descriptive Statistics for Pair-Wise Comparisons for the Treatment-Group Overall Scores between Pre-Test (T1) and Follow-Up (T3)*

Within-Group Pairwise Comparisons (T1 to T3, *N* = 194)					
Scale	*ES*	*SE*	Scale	*ES*	*SE*
SDQ			ROPELOC		
General Academic	.219**	.068	Self Confidence	.094	.066
Global Esteem	.262**	.076	Time Efficiency	.302***	.086
Physical Ability	.093**	.035	Active Involvement	.114	.064
Parent Relations	-.012	.075	Cooperative Teamwork	.173*	.071
Same-Sex Relations	-.244**	.091	Open Thinking	.037	.071
Opposite-Sex Relations	.127	.069	Quality Seeking	.010	.070
Problem Solving	.390***	.078	Self-Efficacy	.358***	.074
Mean	.118		Social Effectiveness	.173**	.066
MES			Stress Management	.299***	.072
Thought	.030	.035	Leadership Ability	.158*	.069
Behaviour	.260**	.082	Coping with Change	.346***	.082
Hampering	-.260**	.079	Mean	.188	
Mean	.010		External LoC	-.322***	.080
GHQ-12			Internal LoC	.043	.066
Positive	.058	.044	Mean	-.140	
Negative	-.007	.036	BPNSFS		
Mean	.026		Autonomy	.211***	.059
One-Factor Scales			Relatedness	.175**	.056
Wellbeing	.197**	.072	Competence	.169**	.051
Life Satisfaction	.088	.133	Mean	.185	
Gratitude	.013	.072			
Resilience	.297***	.072			

Note. β = standardised effect sizes. ROPELOC scales are: External LoC = External Locus of Control; Internal LoC = Internal Locus of Control. Hampering and GHQ-12 scales have been reverse-scored. The Opposite-Sex Relations scale has been shaded 'grey' to represent the research control scale and has not been included in the computation of the SDQ overall mean score. * $p < .05$, ** $p < .01$, *** $p < .001$.

Statistically significant decreases were observed on ROPELOC scales Quality Seeking (β = -.181, *SE* = .070, $p < .05$), ELoC (β = -.494, *SE* = .082, $p < .001$), and ILoC (β = -.200, *SE* = .070, $p < .01$), while an increase in Autonomy (β = .134, *SE* = .060, $p < .05$) was apparent for the BPNSFS scale.

For secondary outcome variables, comparisons of the T2 to T3 interval revealed a statistically significant gain for treatment participants in MES Behaviour (β = .201, *SE* = .080, $p < .05$), while a diminishing in effects was observed in domains of MES Thought (β = -.154, *SE* = .067, $p < .05$) and Hampering (β = -.400, *SE* = .073, $p < .001$); GHQ-12 Positive (β = -.095, *SE* = .020, $p < .001$) and Negative (β = -.145, *SE* = .034, $p < .001$); and Gratitude (β = -.291, *SE* = .073, $p < .001$).

Conclusion. An alternative approach to post hoc analysis examined changes in outcome effects from post-test (T2) to follow-up (T3) for treatment-only participants. There were four statistically significant increases in effect size over the post-treatment interval, while 10 outcomes showed statistically meaningful decreases. These results found the within-person methods to reflect consistently improved retention of post-treatment gains (T2 to T3) compared to the between-groups analysis (self-concept ES = -.033 versus ES = -.048; life effectiveness ES = -.004 versus ES = -.024; LoC ES = -.347 versus ES = -.422).

Table 48 *Changes in ES from Post-Test (T2) to Follow-Up (T3) for Treatment-Group-Only Participants*

	Within-Group Post-Treatment Change (T2 to T3, *N* = 192)				
Scale	*ES*	*SE*		*ES*	*SE*
SDQ			ROPELOC		
General Academic	.178**	.068	Self Confidence	-.116	.060
Global Esteem	.004	.037	Time Efficiency	.139	.081
Physical Ability	-.011	.062	Active Involvement	-.111	.062
Parent Relations	-.301***	.070	Coop Team	.024	.061
Sm-Sex Rels	-.407***	.087	Open Thinking	.037	.071
Op-Sex Rels	.072	.068	Quality Seeking	-.181*	.070
Problem Solving	.235**	.071	Self-Efficacy	.105	.075
Mean	-.033		Social Effectiveness	-.016	.068
MES			Stress Management	.040	.069
Thought	-.154*	.067	Leadership Ability	.011	.067
Behaviour	.201*	.080	Coping with Change	.029	.080
Hampering	-.400***	.073	Mean	-.004	
Mean	-.118		External LoC	-.494***	.082
GHQ-12			Internal LoC	-.200**	.070
Positive	-.095***	.020	Mean	-.347	
Negative	-.145***	.034	BPNSFS		
Mean	-.120		Autonomy	.134*	.060
One-Factor Scales			Relatedness	-.053	.055
Wellbeing	-.102	.074	Competence	-.033	.049
Life Satisfaction	-.126	.139	Mean	.061	
Gratitude	-.291***	.073			
Resilience	.108	.071			

Note. β = standardised effect sizes. ROPELOC scales are: Coop Team = Cooperative Teamwork, External LoC = External Locus of Control; and Internal LoC = Internal Locus of Control.

* $p < .05$, ** $p < .01$, *** $p < .001$.