COGNITION AND EMOTION
2023, VOL. 37, NO. 5, 990–996
https://doi.org/10.1080/02699931.2023.2214349

Routledge
Taylor & Francis Group

Our individual order of things directs how we think we feel

Thomas Goetz[a], Katarzyna Gogol[b,c], Reinhard Pekrun[d,e,f], Anastasiya A. Lipnevich[g], Eva S. Becker[h], Maike Krannich[h] and Fabio Sticca[i]

[a]Department of Developmental and Educational Psychology, University of Vienna, Vienna, Austria; [b]Department of Empirical Educational Research, University of Konstanz, Konstanz, Germany; [c]Thurgau University of Teacher Education, Kreuzlingen, Switzerland; [d]Department of Psychology, University of Essex, Colchester, UK; [e]Institute for Positive Psychology and Education, Australian Catholic University, Sydney, Australia; [f]Department of Psychology, University of Munich, Munich, Germany; [g]Queens College and the Graduate Center, The City University of New York, New York, NY, USA; [h]Institute of Education, University of Zurich, Zurich, Switzerland; [i]University of Teacher Education in Special Needs, Zurich, Switzerland

ABSTRACT

Our work draws upon Foucault's idea that the order of things, defined as the way we categorise our world, matters for how we think about the world and ourselves. Specifically, and drawing upon Pekrun's control-value theory, we focus on the question of whether the way we individually order our world into categories influences how we think about our typically experienced emotions related to these categories. To investigate this phenomenon, we used a globally accessible example, namely, the categorisation of knowledge based on school subjects. In a longitudinal sample of high school students (grades 9–11), we found that judging academic domains as similar led to judging typical emotions related to those domains as more similar than experienced in real life (assessed via real-time assessment of emotions). Our study thus shows that the order of things matters in how we think we feel with respect to those things.

ARTICLE HISTORY
Received 1 November 2021
Revised 28 April 2023
Accepted 9 May 2023

KEYWORDS
Emotion; experience sampling; Foucault; similarity judgements; achievement

> Thinking without the positing of categories and concepts in general would be as impossible as breathing in a vacuum.
>
> Albert Einstein: Philosopher, Scientist

Per definition, categorisation is a process of ordering various phenomena (objects, events, actions, processes, qualities, etc.) into different groups according to their perceived similarity. Forming categories is a basic cognitive process that guides our interaction with the environment. Categorisation gives meaning to otherwise isolated details in an overwhelmingly complex world.

Throughout their lives, individuals are bound to adopt a set of culturally accepted categories that represent a seemingly objective, ordered world. Thus, *the order of things* (Foucault, 2005) is a cultural code of structuring and interpreting the world. In addition to simply accepting culturally imposed categories, individuals also categorise the surrounding phenomena in an idiosyncratic manner grounded in their subjective judgements of similarities (Goldstone, 1994). In other words, each of us creates his or her own distinctive system of categorizations.

It is safe to assume that perceiving the world through the lens of an idiosyncratic categorisation system has an impact on our thinking and feelings about the world and ourselves. This assumption is based on Pekrun's (2006) control-value theory, which implies that the characteristics of a social environment impact cognitive appraisals (e.g. perceived control and value) related to the environment,

CONTACT Thomas Goetz thomas.goetz@univie.ac.at Department of Developmental and Educational Psychology, University of Vienna, Universitätsstraße 7, 1010 Vienna, Austria

Supplemental data for this article can be accessed online at https://doi.org/10.1080/02699931.2023.2214349.

which, in turn, impact emotional experiences. For example, if we perceive two environments as being similar to each other (e.g. due to culturally adopted beliefs), our thinking (i.e. appraisals) and consequently our feelings (e.g. enjoyment) related to these categories may be similar. With respect to feelings, our individual system of categorisation might strongly impact how different we judge our typical emotional response to be across different categories. Thus, it is plausible to presume that when people reflect upon their emotions in a particular life domain, they apply their unique categorical lens and report their specific, individually conceived realities.

In line with these considerations, the current study tested the assumption that individuals' idiosyncratic approaches to categorising the world have an impact on how similarly they think they typically feel related to various categories, above and beyond what they actually feel in relation to those categories. Actually experienced emotions are highly important, for example, with respect to well-being or learning behaviours (Pekrun, 2006). However, individuals' beliefs about how they typically feel in certain contexts or situations have also been found to be of great importance. For example, beliefs about one's emotions are predictive of future behaviours and choices (Levine et al., 2009; Wirtz et al., 2003). Consequently, individuals' beliefs about how they feel with regard to different categories may have an impact on their decision making about those categories (e.g. educational or career choices). Thus, it is important to understand how beliefs about our typical emotional experiences related to certain categories emerge and how they could potentially be changed.

Our study focused on an important instance of categorisation that preoccupied minds of scientists and philosophers for centuries, namely, categorisation of knowledge into domains of study. The initial attempts to organise knowledge into disciplines can be traced back to Aristotle (Deng & Luke, 2008) who worked on defining domains of inquiry as we know them today. Such organisation is familiar to most people who go through formal schooling, where categorisation of knowledge by subject domain is well-accepted (e.g. mathematics, physics, languages, sports, arts). Based on the aforementioned considerations, it can be assumed that people make unique judgments about similarities among subject domains and thus create their own categories. Such variable categorisation system of subject domains, in turn, can be assumed to have an impact on judgments of emotions typically experienced across different school subjects.

An earlier study by Goetz et al. (2014) suggested that certain school subjects tended to be perceived as more similar than others due to similarities in their constituent aspects (e.g. content difficulty). The study also showed that the similarity of reported typical emotions in those school subjects reflected the degree of perceived similarity between school subjects. However, in this study, individual judgments of the similarity of academic domains were not directly assessed. Instead they were inferred from the between-person correlations of various facets of the domains. In addition, this study did not consider the actual emotions experienced across subjects.

To our knowledge, links between similarity judgements of different categories and the similarity of reports of typically experienced emotions across these categories have not been empirically investigated. Therefore, the aim of this study was to examine the influence of individuals' categorisation of school domains on the similarity of reports of typically experienced emotions across these domains, while controlling for the similarity of the actually experienced emotions across these school subjects. We hypothesised that when two school subjects were judged to be more similar (e.g. French vs English), compared to other school-subject pairs (e.g. mathematics vs English), the similarity of the typical emotions reported across these two school subjects would be higher (compared to other school-subject pairs), above and beyond the influence of the similarity of the actually experienced emotions.

Method

Study design

This study reports a secondary analysis of an existing data set from a longitudinal study (grade levels 9–11) entitled "Structure and antecedents of academic emotions: Longitudinal analyses on habitual and state emotions across and within school domains" funded by the Swiss National Science Foundation (SNSF; grant number 100014_131713/1). The data are publicly available on OSF: https://osf.io/5e23h/?view_only = b39ff601975948d4830c527eefffa5b3. Results of the relations between similarity judgments of school subjects and their relations with typically and actually experienced emotions have not yet been published. A sample of students from eight

upper-track schools (Gymnasium) across the German-speaking parts of Switzerland were surveyed once a year for two consecutive weeks over a three-year period. The study was conducted in compliance with ethical standards (Ethical Principles if the WMA Declaration of Helsinki) and the procedures were deemed appropriate by the Institutional Review Board of the University of Konstanz.

Participants

Assessments were conducted in three waves: Wave 1 (Grade 9, $N = 149$, 54.4% female), Wave 2 (Grade 10, $N = 128$, 57% female), and Wave 3 (Grade 11, $N = 108$, 51.9% female). A total of 102 students participated in all three assessments. Attrition was primarily due to students relocating to another school, or to students being absent during data collection. Our analyses were based on $N = 163$ Swiss students ($M_{age} = 16.52$ years, $SD = 1.00$) who participated in at least one of the three assessment waves in Grade 9, Grade 10, and Grade 11. Students were randomly selected from 45 classrooms (two to four students from each classroom). In each classroom, there were, on average, 3.31 students who participated. The student composition of the classrooms was the same across the four school subjects examined in our study. Students were taught by different teachers across those four subjects.

Assessment design

The procedure at each assessment wave was as follows. First, students completed a paper-and-pencil questionnaire, which focused on demographic information, and reported pairwise judgments of the similarity of the four school subjects (mathematics, German, French, and English).

Second, Experience Sampling Method (ESM) data were collected for a period of ten school days using handheld devices to record students' real-time reports on immediate emotional experiences (i.e. enjoyment, pride, anxiety, shame, anger, and boredom) in their mathematics, German, French, and English classes (3–4 class periods per school subject). iPod Touch 4G devices programmed with experience-sampling-software (iDialog Pad; Kubiak & Krog, 2012) were used for the ESM assessment. Students were instructed to activate the devices at the beginning of their lessons (each lesson lasted 45 min). After activating the device, students were asked to indicate the subject and to confirm that the lesson was taking place. Subsequently, if a lesson was reported to be taking place, the device randomly emitted one audible signal during the next 10-35 min. At that point, students were presented with a questionnaire that they immediately completed. If a student did not respond within four minutes of hearing a signal or viewing a question, the device timed out and recorded a missed signal or a questionnaire time-out. Questions were displayed one at a time. In total, the devices were activated in 5,990 lessons and 5,365 questionnaires were completed (10.43% missed signals, questionnaire time-outs, or application crashes). On average, each student completed 17.38 questionnaires throughout the entire period of ESM-assessment in Wave 1 ($SD = 6.61$; minimum = 1, maximum = 34), 11.54 questionnaires in Wave 2 ($SD = 5.36$; minimum = 1, maximum = 26), and 11.10 questionnaires in Wave 3 ($SD = 5.16$; minimum = 1, maximum = 27). Across all assessment waves, on average, 29.2%, 23.63%, 24.4%, and 22.77% of the questionnaires were related to mathematics, German, French, and English, respectively.

Third, immediately after the ESM data were collected, students completed a paper-and-pencil questionnaire assessing their judgements of their typically experienced emotions in the four school subjects under investigation.

Measures

For judging the similarity of the four school subjects (i.e. mathematics, German, French, and English), we used six 5-point Likert-type items (1 = *not at all similar* to 5 = *very similar*) with one item per pair of school subjects; e.g. "How similar are German and mathematics for you?".

Students' reports on immediate emotional experiences (i.e. enjoyment, pride, anxiety, shame, anger, and boredom) in their mathematics, German, French, and English classes were assessed with one 5-point Likert-type item per emotion (1 = *strongly disagree* to 5 = *strongly agree*; e.g. "At this moment I am experiencing anger"). For each assessment wave, scores across different lessons were aggregated into a single discrete emotion score for each student. This was done separately for each of the four school subjects. This procedure resulted in individual scores for six emotions in four school subjects at three points of data collection (3 waves x 6 emotions x 4 subjects = 72 real-time emotion scores per student). Furthermore, for each wave, emotion, and pair of

school subjects, we calculated the absolute difference between students' real-time emotion scores for the four subjects, which resulted in six difference scores for each of the six emotions (in total $3 \times 6 \times 6 = 108$ difference scores) representing the dissimilarity of the real-time emotions across school subjects. These dissimilarity scores were subtracted from the maximum possible dissimilarity (i.e. four) to represent similarity.

Students' judgements of their typically experienced emotions (i.e. enjoyment, pride, anxiety, shame, anger, and boredom) in each school subject were assessed with one 5-point Likert-type item per emotion (1 = *strongly disagree* to 5 = *strongly agree*; e.g. "In French classes I usually experience anger"). The reliability and validity of single-item measures have been supported by findings from various studies (e.g. Gogol et al., 2014). For each typical emotion, we calculated pairwise absolute differences between scores related to each school-subject pair. At each wave this resulted in six difference scores per student for each of the six emotions (in total $3 \times 6 \times 6 = 108$ difference scores). Similar to the real-time emotions, these scores were inverted so that difference values represented the similarity of reports of typically experienced emotions.

To ensure that the observed relations were not mere artifacts of academic achievement, we controlled for the similarity of academic achievement across different school-subject pairs in our analyses. Academic achievement was included because it had been previously shown to influence students' emotions (see e.g. Pekrun et al., 2017). Academic achievement was indexed through students' mid-year grades (i.e. the last grade obtained before the annual assessments of the present study) in mathematics, German, English, and French, which were obtained from student records. Grades ranged from 1 (*poor*) to 6 (*excellent*). For each wave and pair of school subjects, we calculated the absolute difference between students' grades related to the respective two subjects, which resulted in six difference scores for each wave (in total $3 \times 6 = 18$ difference scores). These scores were also inverted so that resulting difference scores represented the similarity of academic achievement in pairs of school subjects.

Statistical analysis

Our data had a two-level structure with similarity scores, emotions, and achievement, respectively, at Level 1 ($N_{Level1} = 13{,}842$) clustered within students at Level 2 ($N_{Level2} = 163$; see Table S1 in the supplemental material for the visual representation of the data structure). The average number of available scores per student was 84.92. A multilevel random intercepts multiple regression model with level-1 predictors was estimated to examine the judgement of school-subject similarity as predictors of the similarity of the reports on typically experienced emotions, while taking the similarity of the reports on real-time emotions into account. We mean-centered the predictors within persons as we were interested in within-person effects only (Wang & Maxwell, 2015).

Apart from controlling for the similarity of academic achievement across different school-subject pairs in the analysis, we controlled for the effects of time (i.e. wave), type of emotion, and school-subject pair by using dummy variables, because the observations were nested within waves, emotions, and school-subject pairs which represent fixed rather than random factors. This approach is consistent with previously reported inquiries dealing with similar data structure (e.g. Huang, 2016; Möhring, 2012). We included 2 dummy variables (coded e.g. Dummy Variable 1: Wave 2 = 1, other waves = 0) representing the 3 waves (Wave 1 as reference), 5 dummy variables representing the 6 emotions (with enjoyment as reference), and 5 dummy variables representing the 6 school-subject pairs (with Math-German as reference). The analyses were conducted with Mplus 8.1 (Muthén & Muthén, 2017).

The percentage of missing values in the variables included in our analyses was 5.65% for the similarity of typically experienced emotions, 13.48% for the similarity of real-time emotions, 2.25% for pairwise school-subject similarity judgements, and 1.65% for similarity of academic achievement. We used the full information maximum likelihood procedure (FIML) implemented in Mplus to account for cases with missing values on all exogenous variables (Muthén & Muthén, 2017). The model parameters were estimated using the MLR estimator, which is robust against violations of the normality assumption.

Results

Descriptive results

Means, standard deviations and ICCs (*M*/*SD*/ICC) were 3.20/0.94/.120 for similarity of typically experienced emotions, 3.39/0.66/.116 for similarity of real-time

emotions, and 2.30/1.19/.128 for school-subject similarity. The ICCs represent the proportion of the variance of a variable that is accounted for by the clusters (i.e. students). The ICCs for all variables were similar in magnitude indicating that most of the variance in all constructs originated from variation within students (88.4% for similarity of typically experienced emotions, 88% for similarity of real-time emotions, and 87.2% for school-subject similarity). Similarity judgements (*M*/*SD*) were 3.34/1.07 for English-French, 2.88/1.09 for German-English, 2.78/1.00 for German-French, 1.77/0.85 for mathematics-German, 1.68/0.90 for mathematics-English, and 1.32/0.69 for mathematics-French. There were positive within-person correlations (Level 1) between the school-subject similarity and the similarity of typically experienced emotions ($r = .07$) and the similarity of real-time emotions ($r = .06$), as well as between the similarity of typically experienced emotions and the similarity of real-time emotions ($r = .25$).

Multilevel multiple regression results

The results from the multilevel random intercepts multiple regression model showed that, intraindividually, the similarity of real-time emotions between two school subjects positively predicted the similarity of the reports on typically experienced emotions at Level 1 (i.e. within persons; $b = 0.26$; ß = 0.18, $p < .001$; see Table 1 and Figure 1). Over and above this effect, the pairwise similarity judgements of different school subjects positively predicted the similarity of these reports as well ($b = 0.04$; ß = 0.05, $p < .01$)[1]. We also found a few significant effects of the covariates.[2.]

Table 1. Two-level multiple regression analyses (within-level) for school-subject similarity and similarity of real-time emotions as predictors of similarity of typically experienced emotions.

Predictor	*b*	β
School-subject similarity	0.04** [0.01; 0.08]	.05** [0.01; 0.09]
Similarity of real-time emotions	0.26***[0.21; 0.32]	.18***[0.15; 0.22]
Similarity of academic achievement	0.13***[0.07; 0.19]	.07***[0.03; 0.10]
Emotion dummies		
Pride	0.23***[0.15; 0.31]	.10***[0.06; 0.13]
Anxiety	0.42***[0.32; 0.51]	.18***[0.14; 0.22]
Shame	0.53***[0.45; 0.62]	.22***[0.19; 0.26]
Anger	0.07 [−0.01; 0.14]	.03 [−0.01; 0.06]
Boredom	0.03 [−0.04; 0.11]	.01 [−0.02; 0.05]
Wave dummies		
Wave 2	0.02 [−0.05; 0.09]	.01 [−0.03; 0.05]
Wave 3	−0.02 [−0.09; 0.05]	−.01 [−0.04; 0.02]
Subject-pair dummies		
German-English	0.06 [−0.02; 0.13]	.02 [−0.01; 0.06]
German-French	−0.03 [−0.11; 0.05]	−.01 [−0.04; 0.02]
English-French	−0.06 [−0.15; 0.02]	−.03 [−0.06; 0.01]
Math-French	−0.01 [−0.07; 0.06]	.00 [−0.03; 0.02]
Math-English	−0.07* [−0.12; −0.02]	−.03* [−0.05; −0.01]

Note. Values in brackets are 95% confidence intervals. Emotions, waves of data collection, and school-subject pairs were dummy coded with *enjoyment*, *Wave 1*, and *Math-German* as reference category, respectively. $N_{Level\ 1} = 13{,}842$ (assessments within students), $N_{Level\ 2} = 163$ (students). *b* = unstandardized coefficients; β = standardized coefficients (in case of the dummy variables these are β_{StdY} coefficients, which should be used for binary covariates; see Muthén & Muthén, 2017); * $p < .05$, ** $p < .01$, *** $p < .001$.

Discussion

Our results indicate that individuals' subjective categorisation of phenomena has an impact on how they think they typically feel in relation to different categories, independently of how they actually feel. Specifically, when we judge two categories as being more similar, we tend to think that we experience more similar emotions with regard to these categories than we actually do. In our study, this is demonstrated by the significant effect of perceived similarity of school subjects on the similarity of beliefs about typically experienced emotions in the subjects, controlling for the similarity of actually experienced emotions. These results contribute to our understanding of how individual beliefs about similarity of typically experienced emotions across categories emerge.

In addition to informing theory, these findings are also of practical relevance. Individuals' beliefs about how they typically feel have been found to be highly predictive of their future behaviour and choices (Levine et al., 2009; Wirtz et al., 2003). Our results demonstrated that independently of the similarity among actually experienced emotions, our own system of categorising subject domains has an impact on how similar we think we typically feel across these categories. By implication, our judgements of how similar categories are can influence our belief-based decision-making processes regarding these categories.

School subjects, used as an example of categorisation in the current study, are an important domain that has clear practical implications. For instance, educational and occupational career choices are thought to be influenced by academic emotions (Eccles, 1994; Wigfield et al., 2002), which could explain the existing underrepresentation of women in post-secondary studies and careers in mathematics (see research on

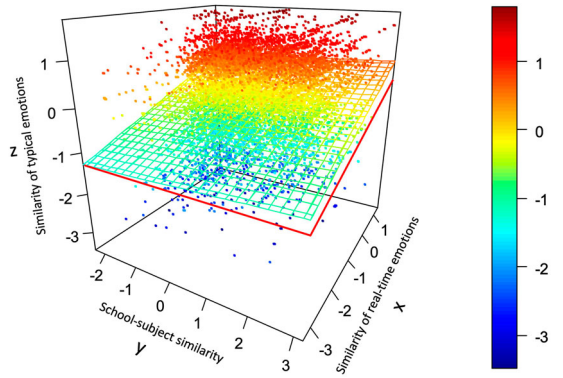

Figure 1. Intraindividual similarity of typical emotions as a function of school-subject similarity and similarity of real-time emotions.

Notes. The colours of the data points depict the similarity of typically experienced emotions (the outcome variable in our analyses, displayed on the z-axis). For instance, blue dots represent low similarity scores, whereas red dots represent high similarity scores. The coloured plane represents the regression plane and shows the estimated similarity of typical emotions as a function of school-subject similarity and similarity of real-time emotions. This regression plane represents a model, in which the similarity of typical emotions is positively predicted by the similarity of real-time emotions as well as by school-subject similarity. The two red lines indicate a second plane for which the parameters of the regression plane were changed in a way that the effect of school-subject similarity (y-axis) on the similarity of typical emotions (z-axis) was set to zero. Thus, this hypothetical plane represents a model in which the similarity of typical emotions is positively predicted by the similarity of real-time emotions exclusively, while school-subject similarity has no additional predictive power. The comparison of the regression plane to the hypothetical plane shows that with increasing perceptions of school-subject similarity the distance between the two planes increases. This increase is equal at all levels of similarity of real-time emotions (i.e. along the x-axis) and shows the effect of school-subject similarity on similarity of typical emotions above the effect of similarity of real-time emotions.

gender in STEM domains; Cheryan et al., 2016; Wang et al., 2013). Our results suggest that one fruitful avenue for reducing the existing gender gap in mathematics, and possibly in other STEM subjects, may be through the influence of judgements about the similarity of school subjects. Teachers and parents may focus on communicating commonalities between, for example, mathematics and humanities – an approach that is likely to result in similarities in individuals' beliefs about emotions they typically experience in these subjects.

Extrapolating our findings to other areas of categorisation offers important implications of high social relevance. For example, investing efforts into changing individuals' similarity judgements about different ethnic, racial, or gender groups could change their beliefs about how they think they feel about these groups and, consequently, alter decisions and future behaviour regarding the groups. This may be a key application and area of future research on categorisation and emotions. In such research, different discrete emotions and their specific effects can be considered. For example, pointing out similarities between racial groups (e.g. the human pursuit of a fulfilling and meaningful life) could lead to improving the corresponding similarity judgments and, consequently, reducing fear of such groups (as a result of an appraisal of otherness) with all its negative effects on oneself and members of those groups.

In addition to the causal links we have proposed, future studies could examine whether relations between emotions experienced in the context of specific subjects may also have implications for how similarly we evaluate those subjects. Experimental studies could be helpful to explore the directionality of this causal link. Moreover, future research could examine the generalizability of our findings to other school subjects (e.g. including other STEM fields and subjects, such as physical education, music, and arts) as well as domains of categorizations beyond school subjects. Finally, future studies could investigate the psychological mechanisms that lead from beliefs about domains to beliefs about typically experienced emotions in these domains.

Notes

1. A random slope model was estimated to test the robustness of the results, which indicated variance in the slope (mean = 0.032, p <.001; variance = 0.014, p < .001) and intercept (mean = 3.204, p < .001; variance = 0.112, p < .001) of the effect of the school-subject similarity on the similarity of typically experienced emotions, controlling for all other predictors in the model. The prediction results were largely unchanged when allowing for the slopes to vary across individuals. Specifically, the following effects on typically experienced emotions were found: b = 0.03, p < .001 (ß = .04, p < .001) for the school-subject similarity; b = 0.26, p < .001 (ß = .18, p < .001) for similarity of real-time emotions; b = 0.11, p < .001 (ß = .06, p < .001) for the similarity of academic achievement.
2. First, the similarity of academic achievement across school subjects positively predicted the similarity of reports on typically experienced emotions (b = .13; ß = .07, p < .001). Second, students' reports on typically experienced emotions were more similar across school subjects for pride, anxiety, and shame than for enjoyment. This finding is in line with results from previous studies (e.g., Goetz et al., 2006, 2007). Moreover, students reported more similar typical emotions in the school-subject pair Math-English than in Math-German.

Acknowledgement

We thank V. Morger for help with the conceptualisation of the field study-design and establishing contact to the participating schools.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the Swiss National Science Foundation (SNSF) [project "Structure and Antecedents of Academic Emotions: Longitudinal Analyses on Habitual and State Emotions Across and Within School Domains"; project grant number 100014_131713/1] and supported by a Research Chair grant awarded to Reinhard Pekrun from the University of Munich (VII.1-H172.10).

References

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2016). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, *143*(1), 1–35. https://doi.org/10.1037/bul0000052

Deng, Z., & Luke, A. (2008). Subject matter: Defining and theorizing school subjects. In F. M. Connelly, M. F. He, & J. Phillion (Eds.), *The SAGE handbook of curriculum and instruction* (pp. 66–87). Sage Publications.

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. Model of Achievement-Related Choices. *Psychology of Women Quarterly*, *18*(4), 585–609. https://doi.org/10.1111/j.1471-6402.1994.tb01049.x

Foucault, M. (2005). *The order of things: An archeology of the human sciences*. Routledge.

Goetz, T., Frenzel, A. C., Pekrun, R., Hall, N. C., & Lüdtke, O. (2007). Between- and within-domain relations of students' academic emotions. *Journal of Educational Psychology*, *99*(4), 715–733. https://doi.org/10.1037/0022-0663.99.4.715

Goetz, T., Haag, L., Lipnevich, A. A., Keller, M. M., Frenzel, A. C., & Collier, A. P. (2014). Between-domain relations of students' academic emotions and their judgments of school domain similarity. *Frontiers in Psychology*, *5*, 1153. https://doi.org/10.3389/fpsyg.2014.01153

Goetz, T., Pekrun, R., Hall, N. C., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology*, *76*(2), 289–308. https://doi.org/10.1348/000709905X42860

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., Fischbach, A., & Preckel, F. (2014). "My Questionnaire is Too Long!" The assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology*, *39*(3), 188–205. https://doi.org/10.1016/j.cedpsych.2014.04.002

Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, *52*(2), 125–157. https://doi.org/10.1016/0010-0277(94)90065-5

Huang, F. L. (2016). Alternatives to multilevel modeling for the analysis of clustered data. *The Journal of Experimental Education*, *84*(1), 175–196. https://doi.org/10.1080/00220973.2014.952397

Kubiak, T., & Krog, K. (2012). Computerized sampling of experiences and behavior. In M. R. Mehl, & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 124–e143). The Guilford Press.

Levine, L. J., Lench, H. C., & Safer, M. A. (2009). Functions of remembering and misremembering emotion. *Applied Cognitive Psychology*, *23*(8), 1059–1075. https://doi.org/10.1002/acp.1610

Möhring, K. (2012). The fixed effects as an alternative to multilevel analysis for cross-national analyses. *GK Soclife Working Papers Series*, *16*. https://doi.org/10.31235/osf.io/3xw7v

Muthén, L. K., & Muthén, B. O. (1998-2018). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, *88*(5), 1653–1670. https://doi.org/10.1111/cdev.12704

Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods*, *20*(1), 63–83. https://doi.org/10.1037/met0000030

Wang, M.-T., Eccles, J. S., & Kenny, S. (2013). Not lack of ability but more choice. *Psychological Science*, *24*(5), 770–775. https://doi.org/10.1177/0956797612458937

Wigfield, A., Battle, A., Keller, L. B., & Eccles, J. S. (2002). Sex differences in motivation, self-concept, career aspiration, and career choice: Implications for cognitive development. In A. McGillicuddy-De Lisi, & R. de Lisi (Eds.), *Biology, society, and behaviour. The development of sex differences in cognition* (pp. 93–124). Ablex.

Wirtz, D., Kruger, J., Scollon, C. N., & Diener, E. (2003). What to Do on spring break? *Psychological Science*, *14*(5), 520–524. https://doi.org/10.1111/1467-9280.03455