Chapter 6
Insights from Motivational Profiles in TIMSS Mathematics

Abstract A person-centered cluster analysis approach to the study of motivation in IEA's Trends in International Mathematics and Science Study (TIMSS) mathematics has revealed interesting profiles of students across key motivational constructs. Between four and six different clusters were extracted from each sample analyzed. Unsurprisingly, some clusters had consistent motivation scores, but in almost every jurisdiction, there were clusters of students with inconsistent score distributions between the contributing motivational constructs. The clusters were systematically different on various external variables, such as mean mathematics achievement, gender composition, and the level of home resources available to students. The study also presents a novel way of looking at the relative importance of enjoyment of, confidence in, and value for mathematics, and the association of these motivation variables with achievement and other demographic characteristics at the cluster level. When motivation scores were mixed rather than consistent, there was a uniform achievement advantage enjoyed by the groups of students who had higher scores for confidence in mathematics over enjoyment of, or value for mathematics. This approach revealed that gender and socioeconomic background are not independent of cluster membership. Typically, clusters with high confidence values were comprised of more boys than girls, and students from better resourced homes. The findings can be linked to relevant literature on motivation in mathematics. Educational efforts to develop student motivation need to take into account differential student profiles and prioritize techniques that target skill and competence in mathematics.

Keywords Cluster analysis · Educational achievement · Family characteristics · Mathematics competence · Mathematics motivation · Student characteristics · Student motivation profiles

6.1 Examining the Role of Motivation in Educational Achievement

From the early endeavors of IEA to study educational achievement cross-culturally (such as the Six Subject Survey, and the First International Mathematics Study; see https://www.iea.nl/other-iea-studies), a broad objective was to understand the

relationships between inputs and outputs in education more fully (Wagemaker 2014). With insights from this research, innovations and reforms could then be designed to assist in the improvement of educational systems. Over the past 60 years, ILSA programs have developed, expanded, and increased to encompass multiple school subjects and skills, a range of populations, diverse item formats and modes of administration, advanced technical and analytic methods, and detailed contextual and demographic measures. Hypotheses and research questions can be addressed through primary and secondary analyses of ILSA data. Moreover, successive administrations of each ILSA program facilitate the examination of stability or change in the measured characteristics.

In our study of student motivation profiles, we analyzed grade four and grade eight data from the TIMSS 1995, 2007, and 2015 administrations for 12 jurisdictions, focusing specifically on the mathematics assessments. We explored three dimensions of student motivation: enjoyment of mathematics (enjoyment), self-confidence in mathematics ability (confidence), and perceived value of mathematics (value, which was assessed at grade eight only). The number of items in the student background questionnaires that tap motivational, self-concept, and affective constructs related to mathematics have more than doubled over time, from 10 items in TIMSS 1995 to 27 items in TIMSS 2015.

Motivational variables are considered important predictors of achievement, and are even discussed as curricular goals by themselves (Hooper et al. 2013). Research on the relationship between motivation and achievement is abundant in the fields of educational psychology and mathematics learning. It has been examined extensively within various theoretical frameworks, such as self-determination theory, expectancy-value theory, self-efficacy, self-concept, and achievement goal theory. It has been addressed with correlational and experimental data, data from ILSA programs, and from TIMSS in particular. Reviews and meta-analytic studies (Hattie 2009; Lee and Shute 2010) have shown that motivational characteristics are associated with achievement. Similar research endeavors have additionally tried to compare those predictors; self-efficacy and confidence have been found to correlate more strongly with achievement than other variables (Richardson et al. 2012; Stankov 2013). Nonetheless, the strength of the relationship of these motivational and affective variables to achievement tends to be weak ($r < 0.30$). This suggests that while there is a general positive association between achievement and confidence, and interest and value, it may not be universal. It may be that, in looking at the combinations of the motivational constructs, some insight might be shed on factors that limit the association. The traditional correlational approach would suggest that all motivation variables correlate with achievement, and, if modeled together, their relative importance would vary and that they might interact; but the person-centered approach has revealed that groups of students, often large in size, have very distinctive profiles that would remain undetected in a variable-centered analysis.

6.2   Clusters of Students Using Motivation Variables: A Person-Centered Approach

Using a person-centered approach implemented via two-step clustering is a novel approach to the problem in the context of TIMSS motivational characteristics. Instead of looking at linear relationships between each motivation variable and achievement, as is common in the literature, we aimed to identify meaningful clusters of students; clusters with distinct profiles based on varying combinations of enjoyment, confidence, and value scores. These constructs are positively inter-correlated, but they are not identical, hence students need not score consistently on all three. Clustering is an empirical method for grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other clusters. As such, we used the two (enjoyment and confidence at grade four) or three (enjoyment, confidence, and value at grade eight) motivational scores provided by student self-reports to group those with similar response patterns, independent of their background variables or their mathematics performance. From each sample within 12 jurisdictions, three administrations, and two grades, we extracted between four and six clusters.

We found two types of patterns in the results. The most obvious pattern, especially in light of the intercorrelation of the variables, is consistent. As expected, there were clusters with consistently strong endorsement for both (in grade four) or all three (in grade eight) variables, clusters with consistently moderate scores, and clusters with consistently low scores for the motivation variables. Given the positive relationship reported in the literature between motivational variables and educational performance, the mean achievement in these clusters was related to the level of motivation. As such, these clusters conform to the results expected from the variable-centered type of analysis.

A more interesting pattern was revealed by the inconsistent clusters: groups of students who were inconsistent in their motivation variable responses. If a student is inconsistent (e.g., scoring high on one variable, but medium or low on other motivation variables), the natural question is whether endorsement of one motivational construct over the others is associated with greater learning success. Thus, rather than presuming that each contributing motivational variable works the same way for all participants, this clustering approach shows that the motivational variables work in quite different ways for subsections of the student population, and potentially explains why the overall effect of the variable toward achievement is relatively modest. The most common inconsistent profile was that of high value for mathematics and lower enjoyment and confidence for the subject. Inconsistent standing on confidence and enjoyment was also detected in multiple samples, although to a lesser extent; the two variables had distributions that usually aligned. Within the inconsistent clusters, mean achievement was higher when the score distribution for confidence, and often enjoyment, was high, while scores for value had less of an effect.

We also examined the clusters for differences in student characteristics and family background, (i.e., gender composition, home resources or parental education, and homework engagement). With respect to student gender, there were more girls than boys in clusters with low motivation, while there were more boys than girls in clusters with high motivation, or clusters that had high confidence but lower scores on other variables. With respect to family background, clusters with higher motivation scores and greater achievement had more home resources. The effect of more privileged homes was also reflected by the relative strength of confidence; students from more privileged homes tended to have greater confidence in mathematics. These findings validate the person-centered approach as theoretically worthwhile, and highlight the usefulness of identifying clusters with well-defined characteristics.

6.3   Motivation Clusters and Achievement

The role of the value for mathematics variable was a consistent finding across jurisdictions and administrations at grade eight. Value has to do with recognizing the importance, worth, or usefulness of something. This is inherently extrinsic, since the value of mathematics lies in how contemporary society uses mathematical skills to solve technical and technological problems. Because of that, the value of mathematics often lies in its utility and its contribution to future educational and employment opportunities for school students. Hence, as well as lying within the individual, value in mathematics also lies external to the individual in how it is viewed and treated by society. Recognition of the value of mathematics is meant to inspire students to persevere with learning it and prioritizing it in their schoolwork. But it is fundamentally an external motivator, and may even be accompanied with dislike for the subject (Ryan and Deci 2000). External motivators have been found to be less strongly associated with achievement (Eklöf 2007; Lee and Stankov 2018; Marsh et al. 2006).

In consistent clusters, the value score distribution would overlap with the confidence and enjoyment distributions. In inconsistent clusters, value would diverge; value would be typically higher than the confidence and enjoyment variables, implying that there are groups of students who value mathematics and consider the subject important in their life, but do not feel capable or intrinsically motivated. The cluster analysis identified such distinct groups of students. However, their motivational profile was usually less academically successful, since ascribing a higher value was not associated with higher achievement. When dissociated from enjoyment and confidence, a high value for mathematics could not compensate for low interest and self-competence, and was associated with comparatively low levels of performance. A recent multi-level analysis by Lee and Stankov (2018) with multiple predictors of achievement from TIMSS and PISA has also shown that value has a negative effect size, much smaller than self-beliefs, affect, and other motivation variables. A similar suggestion was made by Ratelle et al. (2007) after identifying motivational profiles of students within a SDT framework; students with an autonomous profile had similar performance but higher perseverance than students with a combined autonomous and extrinsically-oriented motivational profile.

The comparison of enjoyment of and confidence in mathematics was also systematic across countries, administrations, and the two grades. Confidence in personal ability to perform learning tasks has been found to be a stronger predictor for achievement (e.g., Stankov 2013; Lee and Stankov 2018; Marsh et al. 2006). Enjoyment and positive affect towards mathematics are desirable attitudes influencing engagement with the subject and fostering learning. Students with greater interest in a domain tend to have greater prior knowledge and learn more from instruction (Murphy and Alexander 2002). Score distributions for the two variables were largely overlapping in most of the clusters (with or without consistent scores on the value variable). This pattern is consistent with research among older students (Marsh et al. 2006); strong covariation has been found between affect that intrinsically motivates students and belief in their own mathematical competence. However, in the current study, when they were not consistent within a cluster, it was confidence that was more closely associated with achievement.

An implication from the current analysis is that a positive attitude towards mathematics is adaptive (i.e., associated with greater achievement) if coupled with confidence and efficacy beliefs. Nonetheless, this is not the naïve notion of believing in confidence itself or in oneself; rather, this is an example of justified confidence that is associated with actual performance. The students with stronger endorsement of confidence rightly believed they could do the mathematics in the TIMSS tests: they achieved higher scores than their peers who prioritized value or enjoyment, but lacked strong beliefs in their capabilities. Thus, it would seem that mathematics teaching should seek to develop justified confidence in doing mathematics tasks. As students gain competence in the domain of mathematics it seems plausible that they will develop confidence in their capabilities. Trying to give students a sense of confidence, independent of real capability, is unlikely to be effective (see Pajares 2008). Perhaps the challenge for education involves moving classrooms from the practice of teachers making students interested in mathematics or knowing its value, to one in which teachers focus on helping students become competent. It may be sensible with novices to invoke situational interest, but the goal is to lead them to intrinsic interest as a consequence of greater competence, expertise, and knowledge (Murphy and Alexander 2002).

6.4 Motivation Clusters, and Student and Family Characteristics

Consistent gender differences were found between clusters at grades four and eight, across all three administrations and in nearly all samples. It should be noted that we did not compare the mean achievement by gender in this analysis, but the composition of clusters, which also happened to differ in mean achievement. The gender composition differences were typically observed in the extreme clusters, namely the high versus low motivation clusters. There were more boys than girls

in clusters with high motivation students, while there were more girls than boys in clusters with low scores on motivation variables. It is generally reported that boys report greater confidence than girls in mathematics, while in terms of interest the difference is equivocal (Ganley and Lubienski 2016). Hence, the current results from high achieving clusters, whether with consistently high motivation scores or in mixed clusters with high confidence, seem to reflect the advantage boys have around confidence, rather than enjoyment.

The Iranian sample was an exception, in that gender differences were either non-significant, or, when significant, the trend favored girls. It is worth mentioning that, in Iran, students attend single-sex schools, which may reduce sex differences in school achievement. Although Marsh et al. (2013) did not include Iran, a non-Arabic-speaking nation, in their study of four Arab countries, they found comparable results; gender differences on motivation variables and achievement in TIMSS favored girls, in contrast to four Anglo countries where the trend favored boys. Marsh et al. (2013) suggested that in Arab cultures girls have higher achievement because they put more time and effort toward schoolwork compared to boys. Recent studies comparing a large number of countries have found that in more gender equal and more economically developed countries, differences between males and females increased along more traditional lines with respect to (a) occupational and educational preferences (Falk and Hermle 2018), and (b) relative academic strengths and pursuit of science, technology, engineering, and medicine (STEM) degrees (Stoet and Geary 2018). A plausible explanation is that in less gender-equal countries, pressures surrounding the quality of life promote the engagement of females with STEM subjects (Stoet and Geary 2018). In contrast, economic development and gender equality in other countries may allow differential, albeit gender-specific, preferences by males and females to be manifested (Falk and Hermle 2018).

Indicators of socioeconomic background are routinely included in studies of the determinants of educational achievement. Measures such as parental education level, occupation, and income, as well as resources available to students at home, have been used as proxies for SES. Items measuring such characteristics have been included in background questionnaires in TIMSS administered to the students or the parents from the first cycle of the program. When used individually, such indicators are not necessarily stronger than self-belief measures (Lee and Stankov 2018). In the TIMSS 2015 administration, relevant indicators were combined to generate a scale score for home educational resources (grade eight) and home resources for learning (grade four). A multidimensional composite is more likely to result in stronger relationships to achievement (Van Ewijk and Sleegers 2010). Mean comparisons revealed statistically significant differences in home resources across clusters in nearly all samples.

Motivation levels, and achievement levels were not independent of the socioeconomic resources measures. Higher resources scores were found in clusters consisting of students high on all motivation variables, or in inconsistent clusters with high confidence. The pattern was evident even though differences in the cluster resources scores within jurisdictions were smaller than between jurisdictions. The important message here is that home resources were positively associated with

competence and motivation in general, as well as achievement. Parents in higher SES homes are more educated themselves, so they are better able to assist with early school learning, model and socialize academic success, and afford extra tuition and other educational resources. These social determinants of achievement are well-established (Coleman 1966) and it would be unreasonable to expect teachers or individual students to take responsibility for such factors.

The differences in home resources across clusters in grade four samples, while statistically significant, appeared to be smaller and not always systematic compared to the differences in grade eight samples. However, it was evident that socioeconomic disparities are present from the elementary school years (see, e.g., Duncan et al. 2011). School policies and teaching should probably prioritize methods that result in faster development of learning for students from lower resourced homes. For example, widespread use of techniques such as spaced practice, interleaving, retrieval practice, elaboration, concrete examples, and dual coding (Weinstein et al. 2018), frequent reviews, asking questions, and providing models and scaffolding (Rosenshine 2010) is likely to reduce knowledge deficits associated with reduced home resources. Policies that aim to reduce gaps associated with socioeconomic factors are warranted on multiple grounds. Social programs and policies designed to ensure that all children are raised in homes with sufficient socioeconomic resources is a responsibility that goes well beyond that of teaching, schools, or education.

Results about cluster differences in time invested on homework were mixed. In the TIMSS 2015 administration, engagement in homework was not consistent across countries: more motivated and higher achieving students reported more homework in some countries, less homework in others, or there was no systematic pattern between homework and cluster membership in other samples. Descriptive statistics from the TIMSS 2007 grade eight and TIMSS 1995 grade four administrations gave a similar picture. Systematic patterns could be detected in a few jurisdictions where more motivated student clusters reported doing more homework (e.g., TIMSS 1995 grade eight) or less homework (e.g., TIMSS 2007 grade four) than their peers. Moreover, the percentages of homework engagement within samples suggest that there are large differences between jurisdictions. Policies for homework are highly variable across cultures. Total time spent in school, type of school or class, prevalence of afternoon programs or private tutoring, and parental involvement in students' learning are also factors that affect the amount of time devoted to homework, and such factors are not homogeneous across cultures (Chen and Stevenson 1989; Dettmers et al. 2009). This result seems consistent with the research on the value of homework, which suggests that it is highly variable depending on a number of factors, including age of students and style of work done at home (Marzano and Pickering 2007). Hence, the relationship between motivation and homework engagement remains complex and contextualized, at least in terms of the cluster-level approach employed in this study.

6.5   Methodological Concerns

The cluster analysis was implemented using student self-reports on items measuring motivation in the TIMSS background questionnaires. As described in Chap. 3, the measurement of the enjoyment, confidence, and value variables has gradually evolved from a few items intermixed in a list of items measuring attitudes towards mathematics in TIMSS 1995, to separately-presented, multi-item scales with evidence supporting their factorial structure in TIMSS 2015. Score estimation was also different across administrations, with the use of item averages in TIMSS 1995 and 2007, and latent variable methods in TIMSS 2015. Adopting a "construct-level" approach, the cluster analyses in the TIMSS 1995, 2007, and 2015 samples were conducted with the motivation variables as inputs, despite differences in their operationalization and, as a consequence, in their psychometric properties. With these limitations in mind, the cluster solutions have not been compared across time (i.e., in terms of relative size or mean mathematics achievement). Beyond the number of extracted clusters for each jurisdiction, which varied across administrations, graphical comparison was restricted to cluster differences within a grade/administration/jurisdiction sample. Cohort changes in motivation would be an interesting research question to explore, particularly when important reforms are implemented in a country. Invariant methods for obtaining data and modeling scores would enable such comparisons.

Enjoyment and confidence variables were available in TIMSS grades four and eight, while items measuring value for mathematics were administered only at grade eight. Using different input variables in cluster analysis for the two grades prevented the direct comparison of motivation clusters across grades. This difference meant that the hypothesis of developmental decline in motivation could not be examined in the context of cluster analysis. Since similar items for enjoyment and confidence were administered to grade four and grade eight students, provided measurement invariance holds, mean comparisons could be a fruitful further study to pursue.

The cohort findings reported in this book are grounded in visual examination of boxplots. The visual display of the distributions of the variables by cluster facilitated the interpretations for each cluster. However, we did not attempt to validate the cluster results. Discriminant analysis may be used to validate the clustering method, but this would not prove that the students in a cluster are actually different to those in a different cluster. It would only demonstrate the feasibility of using predictor variables for matching the atheoretical empirical method of clustering. A potentially useful extension of this work would be to design field studies to describe different clusters in more detail with methodologies beyond survey self-reports. Research in the real world of school classrooms might identify whether students who respond differently on the TIMSS questionnaire items differ in some way in their practice, strategy usage, or learning, and so on.

Finally, the promising results generated by these cluster analyses suggest that similar approaches can be employed in future research in a number of ways. For example, instead of the classic cluster analysis employed in this study, which

uses data-based agglomeration techniques, latent class analysis might provide more subtle and sophisticated student groups, provided there is a valid model for generating latent classes (Vermunt and Magidson 2002). This research may provide a theoretical basis for modeling motivational classes in learning mathematics; further studies could seek evidence to validate these profiles. Moreover, it would be potentially useful to include more time points in a hypothesis-driven framework, to establish better warranted claims regarding longitudinal changes in student motivation and related outcomes. For example, motivational profiles in relation to achievement can be compared before—and after certain policy, curriculum, or pedagogical innovations to study the impact of such changes. Incorporating more than one country that implemented similar policies (e.g., the formal introduction of assessment for learning, or standardized testing for school accountability) would allow the study of the effectiveness of policies across culturally or geographically diverse jurisdictions. Comparison of mean scores for the motivation variables across clusters could also be explored. Assuming measurement invariance holds, clusters could be compared within countries, and across countries which differ in average performance (e.g., high achieving versus low achieving), culture (e.g., individualistic versus collectivistic), language (e.g., Indo-European versus Asian), performance trends in TIMSS (e.g., increasing versus stable versus decreasing), and so on. Studies that focus on clusters of students provide interesting and potentially different results from studies conducted at the individual-student level. However, inferences from a group-level analysis, such as the current analyses, should not be deduced for individuals, to avoid ecological fallacies.

6.6 Concluding Remarks

We used an innovative person-centered, multiple-sample approach, which provided informative findings and some surprising results. While the inconsistent profiles among the motivational variables were anticipated, we were surprised as to how consistent the results were across TIMSS administrations, jurisdictions, and grades. Although jurisdictions varied in location, culture, population size, mean performance, and socioeconomic level, the motivational clusters looked similar, both in terms of motivational patterns and demographic characteristics; this was evident despite the differences in how TIMSS measured motivation in 1995, 2007, and 2015. The similarity of the clusters and their patterns across such a diverse collection of jurisdictions, administrations, and grades strengthen the generality of the findings and add new empirical knowledge to the literature on motivation.

Clusters with students high on all motivation variables typically had the highest achievement. It was revealing though that, when motivation variables were inconsistent in their distributions, higher levels of confidence, often aligned with enjoyment, were always associated with higher mean performance; this was not as pronounced with value for mathematics. This combined result suggests strongly that student groups that have a mixture of motivational characteristics can be

identified. Moreover, gender and socioeconomic background are not independent of the motivation clusters. Accordingly, educational efforts to develop motivation need to take into account differential student profiles, and prioritize techniques that target skill and competence in mathematics over other motivational dimensions.

References

Chen, C., & Stevenson, H. W. (1989). Homework: A cross-cultural examination. *Child Development, 60*(3), 551–561.

Coleman, J. (1966). *Equality of educational opportunity*. Washington, DC: National Center for Educational Statistics.

Dettmers, S., Trautwein, U., & Lüdtke, O. (2009). The relationship between homework time and achievement is not universal: Evidence from multilevel analyses in 40 countries. *School Effectiveness and School Improvement*, *20*(4), 375–405.

Duncan, G. J., Morris, P. A., & Rodrigues, C. (2011). Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments. *Developmental Psychology, 47*(5), 1263–1279.

Eklöf, H. (2007). Self-concept and valuing of mathematics in TIMSS 2003: Scale structure and relation to performance in a Swedish setting. *Scandinavian Journal of Educational Research, 51*(3), 297–313.

Falk, A., & Hermle, J. (2018). Relationship of gender differences in preferences to economic development and gender equality. *Science*, *362*(6412), eaas9899.

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences, 47*, 182–193.

Hattie, J. A. (2009). *Visible learning: A synthesis of 800+ meta-analyses on achievement*. Abingdon, UK: Routledge.

Hooper, M., Mullis, I. V. S., & Martin, M. O. (2013). TIMSS 2015 context questionnaire framework. In I. V. S. Mullis, & M. O. Martin (Eds.), *TIMSS 2015 assessment frameworks* (pp. 61–82). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://timssandpirls.bc.edu/timss2015/frameworks.html.

Lee, J., & Shute, V. J. (2010). Personal and social-contextual factors in K–12 academic performance: An integrative perspective on student learning. *Educational Psychologist*, *45*(3), 185–202.

Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences*, *65*, 50–64.

Marsh, H. W., Hau, K. T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing*, *6*(4), 311–360.

Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J., Abdelfattah, F., Leung, K. C., … & Parker, P. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology*, *105*(1), 108–128.

Marzano, R. J., & Pickering, D. J. (2007). The case for and against homework. *Educational Leadership, 64*(6), 74–79.

Murphy, P. K., & Alexander, P. A. (2002). What counts? The predictive powers of subject-matter knowledge, strategic processing, and interest in domain-specific performance. *The Journal of Experimental Education, 70*(3), 197–214.

Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 111–139). New York, NY: Lawrence Erlbaum Associates.

Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S., & Senécal, C. (2007). Autonomous, controlled, and amotivated types of academic motivation: A person-oriented analysis. *Journal of Educational Psychology*, *99*(4), 734–746.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387.

Rosenshine, B. (2010). *Principles of instruction*. Brussels, Belgium: International Academy of Education.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67.

Stankov, L. (2013). Noncognitive predictors of intelligence and academic achievement: An important role of confidence. *Personality and Individual Differences*, *55*(7), 727–732.

Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological Science*, *29*(4), 581–593.

Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, *5*(2), 134–150.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. Hagenaars, & A. McCutcheon (Eds.), *Applied latent class analysis* (pp. 89–106). Cambridge, UK: Cambridge University Press.

Wagemaker, H. (2014). International large-scale assessments: From research to policy. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 11–36). Boca Raton, FL: CRC Press.

Weinstein, Y., Madan, C. R., & Sumeracki, M. A. (2018). Teaching the science of learning. *Cognitive Research: Principles and Implications, 3*(1), 1. Retrieved from https://doi.org/10.1186/s41235-017-0087-y.