Check for updates

Unraveling the relation between representational competence and conceptual knowledge across four samples from two different countries

Peter Adriaan Edelsbrunner[1]*and Sarah Isabelle Hofer[2]

[1]ETH Zürich, Zurich, Switzerland, [2]Faculty of Psychology and Educational Sciences, Ludwig-Maximilians-University Munich, Munich, Germany

Introduction: Whereas it is commonly assumed that in learning science, representational competence is a critical prerequisite for the acquisition of conceptual knowledge, comprehensive psychometric investigations of this assumption are rare. We undertake a step in this direction by re-analyzing the data from a recent study that found a substantial correlation between the two constructs in undergraduates in the context of field representations and electromagnetism.

Methods: We re-analyze the data (N=515 undergraduate students; M_{age}=21.81, SD_{age}=4.04) to examine whether the relation between representational competence and conceptual knowledge, both measured with psychometrically validated test instruments, is similar or varies between four samples from two countries. To this end, we will employ correlational analysis and scatter plots. Employing these methods, we will examine whether a positive relation between representational competence and conceptual knowledge can be found and is of similar magnitude in all samples. We will also employ multiple-group latent profile analysis to examine how the more detailed association between the two constructs varies or is similar across samples. Finally, we will examine how commonalities and differences between samples relate to aspects of learners' gender, topic-specific learning opportunities, and individual preferences for specific learning content. The aim is to unravel the generalizability of this relation and thereby derive hypotheses for potential moderating factors that can be further examined in future research.

KEYWORDS

representational competence, conceptual knowledge, undergraduates, stem education, latent profile analysis

1. Introduction

Representational competence, that is, the ability to interpret and translate between different representations of scientific concepts (Kozma and Russell, 2005), is commonly portrayed as a critical prerequisite for developing conceptual knowledge in science (Ainsworth, 2008; Corradi et al., 2012; Treagust et al., 2017). Consequently, much research has examined how these two constructs interact during inquiry activities (e.g., Kohl et al., 2007; Nieminen et al., 2013; Scheid et al., 2019).

Recently, more targeted studies have tried to measure representational competence with elaborate psychometric instruments (Klein et al., 2017; Scheid et al., 2018). Klein et al. (2017) developed a two-tier instrument to assess representational competence in Kinematics, and Scheid et al. (2018) an open-answer instrument to assess representational competence in the topic of Ray Optics.

In the present study, we are concerned with the empirical relationship between representational competence and conceptual knowledge and, in particular, with the conditions that might influence the magnitude of this relationship. Although a strong positive relationship has been assumed by many researchers, it has been noted that the quantitative empirical evidence on this relationship is rather sparse and struggles with methodological issues (Chang, 2018; Edelsbrunner et al., 2022). As Edelsbrunner et al. (2022) note, instruments that have been used to measure these two constructs and their interrelations often have not been psychometrically validated, or suffer from contextual bias. Specifically, if instruments such as pen-and-paper tests that are assumed to measure the two constructs are both contextualized within the same topic (e.g., electromagnetism; Nieminen et al., 2013; Nitz et al., 2014; Scheid et al., 2019), then the common relation between the two constructs cannot be disentangled from such common topical context.

A recent study by Edelsbrunner et al. (2022) tried to overcome these issues by using an assessment instrument for representational competence with field representations (such as vector-field plots and field lines) that does not employ an explicit topical context. Using this instrument, the authors investigated the relation between representational competence with fields and conceptual knowledge about electromagnetism in university undergraduates from Germany and Switzerland. Encompassing students from four different universities, the authors found a substantial positive relationship, with

a Pearson correlation estimate of $r = 0.54$, $p < 0.001$. In addition, the authors found that in a scatter plot, there were almost no students with high conceptual knowledge but low representational competence, whereas there appeared to be more students with high representational competence but low conceptual knowledge (see Figure 1). From these results, the authors inferred the hypothesis that representational competence is a necessary yet insufficient prerequisite for developing conceptual knowledge (Edelsbrunner et al., 2022).

In the present study, we follow up on this hypothesis with an alternative model-based analytical approach. Specifically, whereas a scatter plot is an important and powerful visual tool, it does not prevent getting false impressions about patterns that might be driven by sampling error. Another reason for a model-based re-analysis is that the sample by Edelsbrunner et al. (2022) encompassed four different student samples from different courses (teacher education, STEM and non-STEM study programs) at three different universities. The authors analyzed the students from all four samples within the same models, potentially hiding important information regarding the generalizability of their findings. The overall pattern might be unreliable if, for example, it is underlain by Simpson's paradox (e.g., (Kievit et al., 2013). This paradox describes situations in which, on the level of a whole population [i.e., the common population underlying all four samples by Edelsbrunner et al. (2022)], a pattern is visible that might actually not exist or even be reversed within the distinct sub-samples forming the larger population. More specific analyses on the level of sub-samples are needed to unravel the generalizability of the findings.

In the context of physics education, such sub-samples might be defined by variables such as *gender*, *topic-specific learning opportunities in the classroom*, or *individual preferences for specific learning content* that can be expected to affect learning in a number of
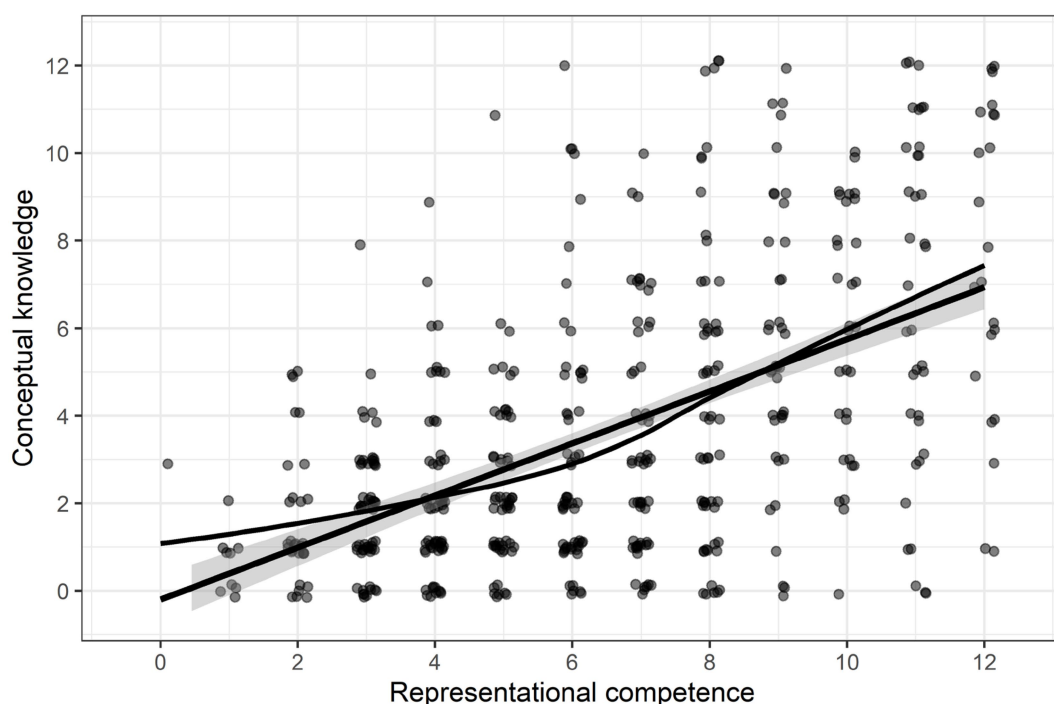

FIGURE 1
Scatter plot of representational competence and conceptual knowledge in data by Malone et al. (2021).

ways. Regarding gender, females seem to perform worse than boys on tests of conceptual knowledge across various physics topics (e.g., OECD, 2009; Madsen et al., 2013; Hofer et al., 2018). Among the manifold explanations provided for this gender effect are missing female role models (e.g., Mullis et al., 2016; Chen et al., 2020) or underlying subject-specific gender differences in motivational-affective variables such as interest and self-concept (e.g., Jansen et al., 2014; Patall et al., 2018; Kang et al., 2019). Although, so far, we do not know much about gender differences in representational competence, female students seem to struggle more with visual graphical representations (e.g., Hegarty and Kriz, 2008; Chan and Wong, 2019; Tam et al., 2019) and different types of mathematical-graphical tasks (e.g., axis tasks; Lowrie and Diezmann, 2011) than male students do. Gender differences in spatial abilities (e.g., Reinhold et al., 2020) might at least in parts explain such findings (see Heo and Toomey, 2020). Concerning the relation between conceptual knowledge and representational competence, Nieminen et al. (2013) reported more problems on the part of female students to infer the same facts from tasks differing in the representational formal.

As regards topic-specific learning opportunities in the classroom, we can expect experiments to be especially effective in promoting understanding of fields and electromagnetism (see National Research Council, 2012; Sandoval et al., 2014; de Jong, 2019; Rodriguez et al., 2020). In physics education, student experiments allow learners to observe physical phenomena and explore their dependence on physical quantities. An important tool helping students to acquire the physical concepts underlying observations are external representations. They can help students in acquiring knowledge by visualizing non-visible fundamentals and causes of the observed phenomena (Olympiou et al., 2013). Usually, external representations of physical concepts such as vector fields are presented before or after experimentation. Teachers provide students with explanations and visual-graphical models which represent aspects that cannot be directly observed. Such learning opportunities involving guided experimentation have proven successful in terms of conceptual understanding (Hardy et al., 2006; van der Graaf et al., 2020). Whether students in the different samples did guided student experiments in the field of electromagnetism, whether they watched the teacher conduct such experiments or whether they did not have this learning opportunity at secondary school at all can hence be expected to influence students' conceptual knowledge and representational competence as well as their relation.

Finally, sub-samples might be determined by systematic differences in individual preferences for specific learning content as reflected in the choice of a specific study program. There is evidence that interest and prior knowledge are substantially and linearly related (Tobias, 1994). In line with this finding, various studies have documented systematic differences on cognitive variables between students in specific study programs. Comparing (among others) physical sciences, math/computer science, engineering, humanities, and social sciences majors, Lubinski and Benbow (2006), for instance, found considerable differences in terms of mathematical, verbal, and spatial ability, with physical sciences students being the only ones with positive manifestations on all three measures.

The four samples are from two different countries, namely Germany and Switzerland, which have similar track-based educational systems. Country will not be examined as a covariate in the present analysis because there is no reason to expect country-level effects on

the relation between and magnitude of students' conceptual knowledge and representational competence, since they are assumed to depend on learning opportunities varying on the teacher- or school-level as well as on individual experiences and preferences. We do not expect systematic differences between Germany and Switzerland on these variables.

2. The present study

In the present study, we examine the generalizability of the findings by Edelsbrunner et al. (2022) across all four samples from their study. Building on the fact that they published their data set for re-use (the published data set is available from the repository of Malone et al., 2021), we will re-analyze these data. The major reason to use this data set is that the present study and its research questions have been triggered by this study and data. In addition, it is the only study so far that has used psychometrically validated instruments to assess both representational competence and conceptual knowledge that are not embedded within the same context, preventing topical bias (Edelsbrunner et al., 2022). Whereas this implies that our study and research questions are bound to the variables available within this data set, the data set appears to offer potential for various informative research questions which to the best of our knowledge have not been addressed so far. Based on this data set, we will address the following research questions:

1. Do we find a positive linear correlation between representational competence and conceptual knowledge in each of the four samples?

To answer this research question, we will estimate linear correlations of the relation between the two constructs in all four samples separately. Based on extensive literature emphasizing the importance of representational competence for acquiring conceptual knowledge (e.g., Nitz et al., 2014), we expect that there is a positive linear correlation in all four samples:

> *H1*: A positive linear association between representational competence and conceptual knowledge exists in all four samples.

2. Do we find comparable linear associations between representational competence and conceptual knowledge in all four samples?

To answer this research question, we will compare the four samples with regard to their linear correlation estimates of the relation between the two constructs. This is an exploratory research question aiming at unraveling potential similarities and differences in the magnitude of the association between the two constructs across samples, so we do not have hypotheses regarding this question.

After examining research questions (1) and (2), we will produce scatter plots individually within all four samples. These will help us visualize the data patterns underlying the estimated correlations to discuss differences and similarities across samples.

Research questions (3) to (5) are also exploratory, with the aim to generate hypotheses for future research. Here, we apply a method to identify different combinations of conceptual knowledge and representational competence that systematically occur in each of the

four samples. These different combinations, for example, high representational competence and low conceptual knowledge, are referred to as latent profiles. The corresponding research questions are as follows:

3. To what degree do latent profiles of representational competence and conceptual knowledge show similar or different patterns across the four different samples?
4. To what degree is the pattern of low representational competence but high conceptual knowledge, but not vice versa, visible in all four samples?
5. Do gender, topic-specific learning opportunities, and individual preferences for specific learning content (as reflected in the choice of study program) explain differences and similarities between the four samples?

To answer research questions (3)–(5), we will apply multiple group latent profile analysis (e.g., Morin et al., 2016). In this analysis, latent profiles of representational competence and conceptual knowledge can be extracted and compared across the four samples.

Although this research is predominantly exploratory, a preregistration can be helpful for exploratory research to ensure that researcher degrees on freedom in selecting and presenting analysis and their results are controlled (Nosek et al., 2018; Dirnagl, 2020). In addition, although latent profile analysis is predominantly used as an exploratory tool, there are central decisions such as the process leading to determining the number of latent profiles that are guided by multiple fit criteria and subjective considerations (Edelsbrunner et al., in press). In preregistering the approach to determining the number of latent profiles, we can ensure that we keep to *a priori* criteria that have been agreed upon in peer-review and clearly mark and discuss any deviations thereof.

3. Method

3.1. Sample

This will be a secondary analysis based on the sample by Edelsbrunner et al. (2022). The sample encompassed $N=515$ undergraduate students from three different universities in Germany and Switzerland. The main sample characteristics are provided in Table 1. For detailed descriptions of the sampling strategy, see Küchemann et al. (2021).

3.2. Assessment instruments

The authors used a newly developed measure of conceptual knowledge about electromagnetism, encompassing 12 items (internal consistency of $\omega=0.92$), and another newly developed measure of representational competence with fields that also encompassed 12 items ($\omega=0.86$). More detailed psychometric characteristics of the two instruments are provided in Edelsbrunner et al. (2022), and more details regarding the instrument for representational competence in Küchemann et al. (2021). We will use sum scores of items solved from each instrument, which in both cases reach from 0 to 12 points.

As covariates, we will use gender, topic-specific learning opportunities, and individual preferences for specific learning content. All variables were assessed in online questionnaires. For gender, participants could indicate female, male, or diverse. We will not include participants with diverse gender in the statistical models, as these were only $n=6$ across all four samples, undermining reliable parameter estimation, but we will include them in all descriptive analysis for which this is possible.

As an indicator of topic-specific learning opportunities, we will use a question on which participants indicated whether their teachers used the conductor swing-experiment in their Physics classes, or not. In this experiment, a conducting piece is put within the magnetic field of a magnet. Current is then activated that flows through the conductor, initiating a second electromagnetic field. Through the two magnetic fields' crossing directions, a Lorentz Force results, causing the conductor piece to swing. The Lorentz Force is a standard topic in German and Swiss Physics education. The conductor swing-experiment, by providing students with a visible phenomenon relating to magnetic fields and the Lorentz force, is supposed to foster students' conceptual understanding of electromagnetism and fields (Donhauser et al., 2020). We will therefore use information from a question asking participants whether in high school, their teachers used the conductor swing-experiment as a demonstration experiment (implemented by the teacher), as a student experiment (implemented by the students themselves), or not at all. In a fourth answer option, the participants could indicate that they could not remember whether this experiment was part of their Physics education. This variables thus has four categorical answer options: Experiment not implemented, implemented as student experiment, implemented as teacher experiment, or cannot remember.

As a final covariate, we will consider students' fields of study. This will serve as an indicator of individual preferences for specific learning

TABLE 1 Characteristics of participants in the four samples.

	Sample 1 (*n*=188)	Sample 2 (*n*=149)	Sample 3 (*n*=98)	Sample 4 (*n*=80)
Gender (% m/f/diverse)	29/71/0	85/14/1	28/29/3	67/29/3
Mean age (*SD*)	20.75 (3.82)	20.72 (2.36)	21.10 (1.60)	19.74 (2.96)
Years of physics at school (*SD*)	4.45 (1.69)	5.57 (1.90)	3.11 (1.27)	4.08 (1.48)
University	Saarland University	TU Kaiserslautern	ETH Zurich	ETH Zurich
Study semester (*SD*)	1.19 (0.97)	1.58 (1.38)	3.12 (0.51)	1.18 (1.59)
Most frequent field of study	Teacher education (100%)	Mechanical engineering (40%), electrical engineering (28%)	Environmental sciences (57%)	Physics (55%)

content. We will use two dummy variables to group students into the non-exclusive categories of teacher education student ($n = 422$) vs. non-teacher education ($n = 93$), and STEM ($n = 305$) vs. non-STEM student ($n = 210$).

3.3. Analytical approach

Analyses will be conducted in the MPlus and R environments. To examine our research questions, we will use descriptive methods and the approach of multiple group latent profile analysis. We will interpret statistical tests at 10% alpha error-levels or 90% confidence intervals, to prevent too large beta error probabilities particularly in the smaller samples. We will not apply corrections for multiple tests, as all but the first research question are exploratory and meant to bring up hypothesis for future research. We will also not use corrections for multiple tests for research question 1 to prevent substantial inflation of beta-errors, which we judge more severe than alpha-errors for the aim of this research question.

3.3.1. Research question 1

To examine the first research question, whether a positive linear association between representational competence and conceptual knowledge exists within all samples, we will estimate bivariate correlations through maximum likelihood-estimation within the Mplus software package. In order to standardize students' mean scores for both constructs, we will define dummy-latent variables with unit variance and a fixed factor loading of one. By constraining the error variance in students' scores to 0, all their variance will be represented in the respective latent variable standardized with a variance of 1. We will estimate the covariance of these two latent variables within each group, which through the standardization will represent the correlation estimate between the two constructs. To test hypothesis 1 and decide whether correlations are present in all six samples, we will inspect bootstrapped 90% confidence intervals based on 10,000 bootstrap draws for the covariance between the two constructs (DiCiccio and Efron, 1996). If a bootstrapped 90% confidence interval within a sample lies fully above 0, we will conclude that a positive correlation between representational competence and conceptual knowledge is present within the respective sample. If the confidence intervals in all four samples are above 0, we will interpret this as support of hypothesis 1. If either or all confidence intervals include 0 or are fully below, we will interpret this as a lack of evidence for hypothesis 1.

3.3.2. Research question 2

To examine the second research question, whether we find comparable linear associations between representational competence and conceptual knowledge in all four samples, we will use parameter constraints to test the covariance parameters between the two constructs for equality between the four samples. We will first set all four covariance parameters to equality and compare the fit of this model to that of the model in which the covariance is allowed to vary between all four samples. We will use $p < 0.10$ as a cut-off to decide whether the assumption of parameter equality holds. If the likelihood ratio test is significant at $p < 0.10$, we will inspect which of the samples contributes most to the significant result and free that sample's

parameter to deviate from the others. This will be done until we yield a model with a non-significant likelihood ratio test. After examining research question 2, we will produce scatter plots to examine the exact nature of the relation between representational competence and conceptual knowledge within the four samples. This will help interpret similarities and differences in this correlation across samples. For example, in these scatter plots we might see different patterns of floor- or ceiling-effects in the samples, which might help explain differences in correlations.

3.3.3. Research questions 3–5

To examine research questions 3–5, we will conduct (multiple group) latent profile analyses. This statistical method allows capturing patterns such as the one observed by Edelsbrunner et al. (2022; i.e., low representational competence and high conceptual knowledge) in explicit model parameters. A latent profile analysis allows clustering individuals based on observed patterns of means and variances on one or more variables (Hickendorff et al., 2018). The two clustering variables in our study will be students' mean scores on the representational competence-test and on the conceptual knowledge-test. In a latent profile analysis, systematically different patterns of means and variances across variables are modelled usually in a mostly data-driven manner. These patterns are then represented in a latent categorical variable that represents the patterns as different latent profiles. We will first determine the number of latent profiles in each sample individually by increasing the number of profiles from 1 to 7 in a step-wise manner. We will specificy latent profiles differing in means and variances across the two indicator variables. The number of profiles will be determined based on the AIC, AIC3, BIC, aBIC, and the VLMR-likelihood ratio test with a significance criterion of $p < 0.10$ (Edelsbrunner et al., in press; Nylund-Gibson and Choi, 2018). The AIC in many cases points to a higher number of profiles than the BIC, with the AIC3 and the aBIC in between (Edelsbrunner et al., in press). In these cases or in case the VLMR-test is in disagreement with the other indices, we will determine whether the additional profiles shown by the AIC, or by one of the other indices, is informative by relying on our content knowledge (Marsh et al., 2009). More in-depth descriptions of the different steps to determine the number of profiles in latent profile analyses are provided by Ferguson et al. (2020) as well as by Hickendorff et al. (2018).

After determining the number of profiles within each sample, we will extend the latent profile analyses to a multiple group-model. In a multiple group-extension, profiles can be extracted within each sample individually but within the same model estimation process (Morin et al., 2016). This allows fixing parameters or letting them vary across samples, to test whether one or more parameters (mean- and variance- estimates of profiles, or relative profile sizes) differ or are similar across samples. Based on this analytic approach, we will examine the comparability and specific patterns of latent profiles found in the four samples. We will inspect which profile parameters appear similar or different between samples and then fix those to equality that we judge to be similar from a theoretical perspective. After fixing parameters to equality, we will test the respective restriction through a likelihood ratio test. We will interpret the resulting profiles to judge to which extent we can find the pattern described in

Edelsbrunner et al. (2022) of students showing high conceptual knowledge only with high representational competence.

After extracting and comparing the profiles, we will add the three covariates to the model *via* Lanza's approach (Lanza et al., 2013; Asparouhov and Muthén, 2014) to examine whether the profiles in the four samples relate to different mean values on students' gender, topic-specific learning experiences, and individual preferences for specific learning content. We will again rely on p-values <0.10 to draw hypotheses for future research regarding the different profiles' correlates.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: https://osf.io/p476u/.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ainsworth, S. (2008). "The educational value of multiple representations when learning complex scientific concepts" in *Visualization: Theory and practice in science education*. eds. J. K. Gilbert, M. Reiner and A. Nakama (Dordrecht: Springer), 191–208.

Asparouhov, T., and Muthén, B. (2014). Auxiliary variables in mixture modeling: using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model. *Mplus Web Notes* 21, 1–22.

Chan, W. W. L., and Wong, T. T.-Y. (2019). Visuospatial pathways to mathematical achievement. *Learn. Instr.* 62, 11–19. doi: 10.1016/j.learninstruc.2019.03.001

Chang, H. Y. (2018). Students' representational competence with drawing technology across two domains of science. *Sci. Educ.* 102, 1129–1149. doi: 10.1002/sce.21457

Chen, C., Sonnert, G., and Sadler, P. M. (2020). The effect of first high school science teacher's gender and gender matching on students' science identity in college. *Sci. Educ.* 104, 75–99. doi: 10.1002/sce.21551

Corradi, D., Elen, J., and Clarebout, G. (2012). Understanding and enhancing the use of multiple external representations in chemistry education. *J. Sci. Educ. Technol.* 21, 780–795. doi: 10.1007/s10956-012-9366-z

de Jong, T. (2019). Moving towards engaged learning in STEM domains; there is no simple answer, but clearly a road ahead. *J. Comput. Assist. Learn.* 35, 153–167. doi: 10.1111/jcal.12337

DiCiccio, T. J., and Efron, B. (1996). Bootstrap confidence intervals. *Stat. Sci.* 11, 189–228. doi: 10.1214/ss/1032280214

Dirnagl, U. (2020). Preregistration of exploratory research: learning from the golden age of discovery. *PLoS Biol.* 18:e3000690. doi: 10.1371/journal.pbio.3000690

Donhauser, A., Küchemann, S., Kuhn, J., Rau, M., Malone, S., Edelsbrunner, P., et al. (2020). Making the invisible visible: visualization of the connection between magnetic field, electric current, and Lorentz force with the help of augmented reality. *Phys. Teach.* 58, 438–439. doi: 10.1119/10.0001848

Edelsbrunner, P.A., Flaig, M., and Schneider, M. (in press). A simulation study on latent transition analysis for examining profiles and trajectories in education: Recommendations for fit statistic. *Journal of Research on Educational Effectiveness*. doi: 10.1080/19345747.2022.2118197

Edelsbrunner, P. A., Lichtenberger, A., Malone, S., Küchemann, S., Stern, E., Brünken, R., et al. (2022). *The relation of representational competence and conceptual knowledge in female and male undergraduates*. Preprint available from PsyArXiv, under https://psyarxiv.com/m4u5j

Ferguson, S. L., Moore, G., E, W., and Hull, D. M. (2020). Finding latent groups in observed data: a primer on latent profile analysis in Mplus for applied researchers. *Int. J. Behav. Dev.* 44, 458–468. doi: 10.1177/0165025419881721

Hardy, I., Jonen, A., Möller, K., and Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking.". *J. Educ. Psychol.* 98, 307–326. doi: 10.1037/0022-0663.98.2.307

Hegarty, M., and Kriz, S. (2008). "Effects of knowledge and spatial ability on learning from animation" in *Learning with animation: Research implications for design*. eds. R. Lowe and W. Schnotz (New York, NY: Cambridge University Press), 3–29.

Heo, M., and Toomey, N. (2020). Learning with multimedia: the effects of gender, type of multimedia learning resources, and spatial ability. *Comput. Educ.* 146:103747. doi: 10.1016/j.compedu.2019.103747

Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., and Trezise, K. (2018). Informative tools for characterizing individual differences in learning: latent class, latent profile, and latent transition analysis. *Learn. Individ. Differ.* 66, 4–15. doi: 10.1016/j.lindif.2017.11.001

Hofer, S. I., Schumacher, R., Rubin, H., and Stern, E. (2018). Enhancing physics learning with cognitively activating instruction: a quasi-experimental classroom intervention study. *J. Educ. Psychol.* 110, 1175–1191. doi: 10.1037/edu0000266

Jansen, M., Schroeders, U., and Lüdtke, O. (2014). Academic self-concept in science: multidimensionality, relations to achievement measures, and gender differences. *Learn. Individ. Differ.* 30, 11–21. doi: 10.1016/j.lindif.2013.12.003

Kang, J., Hense, J., Scheersoi, A., and Keinonen, T. (2019). Gender study on the relationships between science interest and future career perspectives. *Int. J. Sci. Educ.* 41, 80–101. doi: 10.1080/09500693.2018.1534021

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., and Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in psychology*. 4:513.

Klein, P., Müller, A., and Kuhn, J. (2017). Assessment of representational competence in kinematics. *Phys. Rev. Phys. Educat. Res.* 13:10132. doi: 10.1103/PhysRevPhysEducRes.13.010132

Kohl, P. B. (2007). Towards an understanding of how students use representations in physics problem solving. (Boulder: Doctoral dissertation, University of Colorado).

Kozma, R., and Russell, J. (2005). "Students becoming chemists: Developing representational competence" in *Visualization in science education*. ed. J. K. Gilbert (Dordrecht: Springer Netherlands), 121–145.

Küchemann, S., Malone, S., Edelsbrunner, P., Lichtenberger, A., Stern, E., Schumacher, R., et al. (2021). Inventory for the assessment of representational competence of vector fields. *Phys. Rev. Phys. Educ. Res.* 17:020126. doi: 10.1103/PhysRevPhysEducRes.17.020126

Lanza, S. T., Tan, X., and Bray, B. C. (2013). Latent class analysis with distal outcomes: a flexible model-based approach. *Struct. Equ. Model. Multidiscip. J.* 20, 1–26. doi: 10.1080/10705511.2013.742377

Lowrie, T., and Diezmann, C. (2011). Solving graphics tasks: gender differences in middle-school students. *Learn. Instr.* 21, 109–125. doi: 10.1016/j.learninstruc.2009.11.005

Madsen, A., McKagan, S. B., and Sayre, E. C. (2013). Gender gap on concept inventories in physics: what is consistent, what is inconsistent, and what factors influence the gap? *Phys. Rev. ST Phys. Educ. Res.* 9:020121. doi: 10.1103/PhysRevSTPER.9.020121

Malone, S., Küchemann, S., Edelsbrunner, P. A., Lichtenberger, A., Altmeyer, K., Schumacher, R., et al. (2021). *CESAR 0 data*. Available at: https://osf.io/p476u/

Marsh, H. W., Lüdtke, O., Trautwein, U., and Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: synergy of person- and variable-centered approaches to theoretical models of self-concept. *Struct. Equ. Model. Multidiscip. J.* 16, 191–225. doi: 10.1080/10705510902751010

Morin, A. J. S., Meyer, J. P., Creusier, J., and Biétry, F. (2016). Multiple-group analysis of similarity in latent profile solutions. *Organ. Res. Methods* 19, 231–254. doi: 10.1177/1094428115621148

Mullis, I. V. S., Martin, M. O., Foy, P., and Hooper, M. (2016). *TIMSS advanced 2015 international results in advanced mathematics and physics*. Available at Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2015/international-results/advanced/

National Research Council (2012), *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.

Nieminen, P., Savinainen, A., and Viiri, J. (2013). Gender differences in learning of the concept of force, representational consistency, and scientific reasoning. *Int. J. Sci. Math. Educ.* 11, 1137–1156. doi: 10.1007/s10763-012-9363-y

Nitz, S., Ainsworth, S. E., Nerdel, C., and Prechtl, H. (2014). Do student perceptions of teaching predict the development of representational competence and biological knowledge? *Learn. Instr.* 31, 13–22. doi: 10.1016/j.learninstruc.2013.12.003

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., and Mellor, D. T. (2018). The preregistration revolution. *Proc. Natl. Acad. Sci.* 115, 2600–2606. doi: 10.1073/pnas.1708274114

Nylund-Gibson, K., and Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Transl. Issues Psychol. Sci.* 4, 440–461. doi: 10.1037/tps0000176

OECD (2009), *Equally prepared for life?: How 15-year-old boys and girls perform in school*, PISA, OECD Publishing, Paris

Olympiou, G., Zacharias, Z., and Dejong, T. (2013). Making the invisible visible: enhancing students' conceptual understanding by introducing representations of abstract objects in a simulation. *Instr. Sci.* 41, 575–596. doi: 10.1007/s11251-012-9245-2

Patall, E. A., Steingut, R. R., Freeman, J. L., Pituch, K. A., and Vasquez, A. C. (2018). Gender disparities in students' motivational experiences in high school science classrooms. *Sci. Educ.* 102, 951–977. doi: 10.1002/sce.21461

Reinhold, F., Hofer, S., Berkowitz, M., Strohmaier, A., Scheuerer, S., Loch, F., et al. (2020). The role of spatial verbal numerical and general reasoning abilities in complex word problem solving for young female and male adults. *Math. Educ. Res. J.* 32, 189–211. doi: 10.1007/s13394-020-00331-0

Rodriguez, L. V., van der Veen, J. T., Anjewierden, A., van den Berg, E., and de Jong, T. (2020). Designing inquiry-based learning environments for quantum physics education in secondary schools. *Phys. Educ.* 55:065026. doi: 10.1088/1361-6552/abb346

Sandoval, W. A., Sodian, B., Koerber, S., and Wong, J. (2014). Developing children's early competencies to engage with science. *Educ. Psychol.* 49, 139–152. doi: 10.1080/00461520.2014.917589

Scheid, J., Müller, A., Hettmannsperger, R., and Schnotz, W. (2018). "Representational competence in science education: from theory to assessment" in *Towards a framework for representational competence in science education*. ed. K. Daniel (Cham: Springer), 263–277.

Scheid, J., Müller, A., Hettmannsperger, R., and Schnotz, W. (2019). Improving learners' representational coherence ability with experiment-related representational activity tasks. *Phys. Rev. Phys. Educ. Res.* 15:010142. doi: 10.1103/PhysRevPhysEducRes.15.010142

Tam, Y. P., Wong, T. T. Y., and Chan, W. W. L. (2019). The relation between spatial skills and mathematical abilities: the mediating role of mental number line representation. *Contemp. Educ. Psychol.* 56, 14–24. doi: 10.1016/j.cedpsych.2018.10.007

Tobias, S. (1994). Interest, prior knowledge, and learning. *Rev. Educ. Res.* 64, 37–54. doi: 10.3102/00346543064001037

Treagust, D. F., Duit, R., and Fischer, H. E. (Eds.). (2017). *Multiple representations in physics education* (Vol. *10*). Cham Springer International Publishing

van der Graaf, J., Segers, E., and de Jong, T. (2020). Fostering integration of informational texts and virtual labs during inquiry-based learning. *Contemp. Educ. Psychol.* 62:101890. doi: 10.1016/j.cedpsych.2020.101890