Rosenberg *et al.*
International Journal of STEM Education (2023) 10:68
https://doi.org/10.1186/s40594-023-00449-0

International Journal of STEM Education

RESEARCH

Open Access

Using intensive longitudinal methods to quantify the sources of variability for situational engagement in science learning environments

Joshua M. Rosenberg[1*], Patrick N. Beymer[2], Vicky Phun[3] and Jennifer A. Schmidt[3]

Abstract

Background Situational engagement in science is often described as context-sensitive and varying over time due to the impact of situational factors. But this type of engagement is often studied using data that are collected and analyzed in ways that do not readily permit an understanding of the situational nature of engagement. The purpose of this study is to understand—and quantify—the sources of variability for learners' situational engagement in science, to better set the stage for future work that measures situational factors and accounts for these factors in models.

Results We examined how learners' situational cognitive, behavioral, and affective engagement varies at the situational, individual learner, and classroom levels in three science learning environments (classrooms and an out-of-school program). Through the analysis of 12,244 self-reports of engagement collected using intensive longitudinal methods from 1173 youths, we found that the greatest source of variation in situational engagement was attributable to individual learners, with less being attributable to—in order—situational and classroom sources. Cognitive engagement varied relatively more between individuals, and affective engagement varied more between situations.

Conclusions Given the observed variability of situational engagement across learners and contexts, it is vital for studies targeting dynamic psychological and social constructs in science learning settings to appropriately account for situational fluctuations when collecting and analyzing data.

Keywords Intensive longitudinal methods, Engagement, Science education, Multivariate models, Mixed effects models

Many science teachers know that small but important moments can make a notable difference for learners. This sentiment is noted in theoretical frameworks that recognize that dynamic, moment-to-moment factors can be important. For instance, interest in science can be sparked (and sustained) by serendipitous occurrences (Azevedo, 2018; Hidi & Renninger, 2006). Though it is both theoretically ground and common sense that the constructs studied by educational researchers can vary across specific moments, there is occasionally a mismatch regarding how such constructs are measured in research, and the study of situational engagement is one area where this mismatch can be seen.

Situational engagement is regularly defined as dynamic—situation-specific and varying in response to time and features of the context (Appleton et al., 2008;

*Correspondence:
Joshua M. Rosenberg
jmrosenberg@utk.edu
[1] University of Tennessee, 1122 Volunteer Blvd., Knoxville, TN 37919, USA
[2] University of Cincinnati, Cincinnati, USA
[3] Michigan State University, East Lansing, USA

Skinner & Pitzer, 2012; Skinner et al., 2008)—like related constructs, including interest, emotional states, and motivation. As a recent report from the National Academy of Sciences, Engineering, and Medicine (2018) notes, "There is... strong evidence for the view that engagement and intrinsic motivation develop and change over time." Additionally, while scholars have underscored the importance of measuring situational engagement at multiple levels (Eccles & Wang, 2012; Reschly & Christenson, 2012; Sinatra et al., 2015; Xie et al., 2023), this construct is often studied using single time-point surveys—which may not account for the dynamic nature of learners' situational engagement. Indeed, Fredricks and McColskey (2012) reviewed how with extant measures of engagement, "most current methods do not adequately capture the dynamic and interactive nature of engagement" (p. 779). At the same time, scholars called for new data collection approaches—such as intensive longitudinal methods—when measuring engagement (Sinatra et al., 2015). The potential downsides of not accounting for the situational nature of engagement include over-confident (i.e., biased) estimates for the effects of antecedents or the outcomes of situational engagement and missed opportunities to understand contextual—and perhaps malleable—factors and their effects on learners.

In addition to being dynamic, situational engagement can be seen as a dynamic construct that is a function of factors at multiple levels, including learner characteristics and inclinations, situational or momentary factors, and classroom factors (Fredricks et al., 2004; Schmidt et al., 2018; Sinatra et al., 2015; Skinner & Pitzer, 2012). Though theory and past research suggest that situational engagement is a product of factors at these levels (Skinner & Pitzer, 2012; Skinner et al., 2008; Strati et al., 2017), there has been limited empirical work exploring the relative degree to which each of these levels contributes to situational engagement. As a result, we have a limited understanding of how much of learners' engagement can be explained by their learning situation.

At the same time that there exists a limited understanding of the situational impacts on science learners', another issue concerns how engagement in science is measured: using self-report surveys at relatively few time points or using measurement techniques that afford more frequent assessments of the state of students' engagement (Azevedo, 2015), such as *intensive longitudinal methods* (Bolger & Laurenceau, 2013; Hektner et al., 2007; Mehl & Conner, 2013). Intensive longitudinal methods allow for the study of situational engagement in a way that arguably accords better with how it is conceptualized—as dynamic and sensitive to features of the context (Sinatra et al., 2015). These methods also allow for the study of how engagement varies across multiple levels, including not only the teacher and learner level but also the situation-to-situation level.

The purpose of this study, then, is to understand—and quantify—the sources of variability for learners' situational engagement in science. This is to better set the stage for future work that measures situational factors and accounts for these factors in models. If situational factors explain relatively little variability in learners' engagement relative to the factors at the individual learner and class or instructor levels, for instance, then we can gain confidence in the value of the many approaches that emphasize factors at the learner and class level. But, if situational factors account for ample variation in engagement—or, for specific dimensions of engagement (e.g., behavioral or affective engagement, but not cognitive engagement), then this study can point to the need to better understand situational factors and the situational nature of engagement in science. We leverage a unique dataset of with a combined 13,716 reports of situational engagement collected using intensive longitudinal methods from 1173 youth in 50 science learning environments. In addition, we use a relatively novel Bayesian analytic approach suited to the multi-level and multivariate nature of engagement.

Literature review

Scholars generally agree that engagement is a multidimensional construct with multiple components (Christenson et al., 2012; Fredricks et al., 2004). Here, we focus on the three most cited components of engagement: behavioral, cognitive, and affective. *Behavioral engagement* is defined as one's level of effort and participation in academic, social, or extracurricular activities. *Cognitive engagement* is defined as one's value and mental investment in their learning (Fredricks et al., 2004; Greene, 2015; Sinatra et al., 2015). *Affective engagement* is defined as the positive or negative feelings a student has towards their teachers, classmates, learning activities, and/or school more generally (Pekrun & Linnenbrink-Garcia, 2012). As noted earlier, we review domain-general and domain-specific findings on the antecedents, nature, and outcomes of situational engagement but highlight and focus on science education-specific findings.

Sources of variability for science learners' situational engagement

One source of variation in learners' situational engagement is at the situational level. Situational factors are the instructional and classroom events and learners' experiences of them at a moment-to-moment level. These include instructional activities or forms of instructional support (Schmidt et al., 2018; Schneider et al., 2016; Shernoff et al., 2000; Shumow & Schmidt, 2014), such as the specific science and engineering practices in which learners engage (Inkinen et al., 2019, 2020; Rosenberg, 2018; Schmidt et al., 2018). These factors are especially relevant to the present study as they can be experienced by all learners at specific moments. Each learning experience that members of a classroom participate in together (e.g., Tuesday's scaffolded exploration of the permeability of cell membranes relative to Thursday's examination on cell membranes) is likely to have some common association with the engagement of students who share this experience. Some specific situational features have been identified in past science education research, including the importance of the science and engineering practice of developing and using models for learners' engagement (Inkinen et al., 2019) and the role of the choices learners make during laboratory activities in which they engage (Schmidt et al., 2018). Additionally, research has shown how learners' psychological state in different moments can impact their engagement in science: when (male, as discussed further in the next section) learners perceive there to be greater challenge and relevance (Schmidt et al., 2020) and when they value a specific task more and have a greater expectation that they will be successful at that particular task (Upadyaya et al., 2021), they are more likely to be highly engaged. Despite these past research findings on situational factors that impact engagement, the extent to which these kinds of features matter in general (and relative to individual learner sources) is not yet clear.

Another source of variation can be observed at the individual student level. Understanding individual differences may help better explain learners' engagement (Eccles, 2016). Individual factors may include—but are not limited to—gender differences and students' self-beliefs or their beliefs about science and other STEM disciplines. For example, past research concerns how science as a discipline has typically encouraged male learners to see themselves as capable while discouraging women in implicit or explicit ways (Carlone, 2004). Past education research has documented gender disparities in engagement in science and STEM: Schmidt et al. (2020) found that while engagement was positively associated with a perception of the challenge of learning activities in out-of-school STEM programs, these relationships were solely for male students—for female students, the relationship was negative. Additionally, interest in the subject area has been shown to be a strong predictor of engagement in science when in combination with learners being able to make choices about their learning (Beymer et al., 2020). Lastly, learners' mastery orientation has been found to be associated with science learners' engagement (and achievement), when approached in terms of *profiles* of learners' engagement (Caberera et al., 2023). Thus, several learner characteristics are salient factors in terms of science learners' situational engagement. While research has not conclusively identified individual-level factors that explain engagement, what has not been established is how much of the variation in situational engagement occurs at the individual level.

In addition to situational and individual factors relating to engagement, classrooms—and the instructors in them—can shape learners' situational engagement. One key set of findings pertains to how teachers support students' autonomy and learning: when science students' autonomy is supported by their teachers in particular instances in science classes, their engagement is likely higher (Patall et al., 2018; Strati et al., 2017). These effects could accumulate over time, as particular instructors are likely to provide greater or lesser support in these ways. Many studies have attempted to account for instructor or classroom differences, but most of the time, these effects are small (Patall et al., 2018; Schmidt et al., 2020) or they are not the focus and so their magnitudes are not explicitly reported (Inkinen et al., 2019). Relative to findings on situational and learner effects, we know less about how classrooms and instructors shape learners' situational engagement in science. Thus, the extent to which instructor or classroom effects matter relative to individual learner and situational factors has not been established, and understanding the general magnitude of these effects relative to others can guide researchers' efforts in understanding the dynamic nature of situational engagement in science.

Using intensive longitudinal methods to study situational engagement

Exploring the extent to which situational engagement varies across situations, persons, and classrooms requires data collection techniques and research designs. *Intensive longitudinal methods*, a term that encompasses a variety of techniques with names that vary by field (e.g., ESM [Experience Sampling Method], ambulatory assessment, daily diary studies, and ecological momentary assessment), are needed to collect multiple reports of learners' situational engagement across a variety of situations (Bolger & Laurenceau, 2013; Mehl & Conner, 2013).

Intensive longitudinal methods have several affordances for studying situational engagement. First, they allow for the within-learner dynamics of situational engagement—and other constructs—to be studied (Zirkel et al., 2015). Second, they provide a more ecologically valid means of assessing a construct because it can be difficult to recall previous situations long after one was engaged in it (Schwarz, 2012). Thus, these intensive longitudinal methods are both sensitive to changes in engagement over time and between learners, allowing us to understand the nature of situational engagement (Sinatra et al., 2015; Turner & Meyer, 2000; Zirkel et al., 2015).

Research has shown that intensive longitudinal methods can allow researchers to pose and answer new questions. For instance, Shernoff et al. (2003) examined engagement using ESM measures aligned with flow theory, namely, using measures of concentration, interest, and enjoyment (Csikszentmihalyi, 1997). In a study using the same measures of engagement, Shernoff et al. (2016) used an observational measure of challenge and control (or environmental complexity) and found that it significantly predicted engagement, as well as self-esteem, intrinsic motivation, and academic intensity. Schneider et al. (2016) and Linnansaari et al. (2015) examined features of optimal learning moments or moments in which students report high levels of interest, skill, and challenge. In a variant of the ESM, researchers have also used repeated *End-of-Class Reports*, ESM-like surveys conducted at the end of specific classes—rather than randomly selected moments as in ESM—to explore relationships between teachers' instructional practices during a lesson and students' engagement during that lesson (Schmidt et al., 2019). Similarly, scholars have used daily diary studies to examine situational engagement as a function of autonomy-supportive classroom practices (Patall et al., 2016, 2017). Because intensive longitudinal methods involve asking learners about their experience at the time they were signaled in the moment, this technique may be especially useful when the goal is to understand and model factors or variation at the situational level, for which traditional self-report surveys yield limited insight, as well as the effects at the individual learner and classroom levels.

Modeling variation in situational engagement

Despite the large and growing body of research on student engagement, there are not many other examples of studies that show that there is variation in students' engagement at the situational level or how much of the variation in students' engagement is attributable to sources at this level, for several reasons. For example, many studies using intensive longitudinal methods do not collect responses from students in the same situation; for instance, learners within science classes may be asked to respond to ESM responses at different times specific to each student. In such cases, there are no situational dependencies that can be modeled, which may—for many research questions—be an appropriate condition. However, when situational factors are included in a study (e.g., the instrumental and emotional support provided by teachers), it is important to model these effects to obtain unbiased estimates (Strati et al., 2017).

Another issue concerns how situational sources of variation are identified and interpreted. In some past research, scholars have considered the situational and residual sources of variation to be combined, even though they are not analytically separable. In other words, while the residual values can include variation due to specific situations, they may be entirely random noise. In a recent study, van Braak et al. (2021) acknowledged this but argued that, given its malleable nature, it may not be necessary to model variation in engagement at the situational level. However, modeling situational variation as separate from residual variation would allow for empirically disentangling situational and unexplained sources of variation.

Finally, some studies have situational dependencies—but do not model them. These studies are characterized by students within classes responding at the same time, for instance, 15 min into their science classes. In such cases, there are unmodeled dependencies among all of the responses that were generated in the same moment—which means that it is not possible to quantify how much variation in situational engagement is between situations. This also introduces bias in the estimation of other effects (Judd et al., 2012; Raudenbush & Bryk, 2002; West et al., 2014). Examples of studies in which students within the same class responded to the same end-of-class report are the work of Patall et al. (2018) and Schmidt et al. (2017).

If data are collected from students simultaneously, another consideration concerns how to model the data. When present, modeling situational sources of variation can be difficult as doing so requires the use of cross-classified multi-level models (Raudenbush & Bryk, 2002), which some software does not estimate (West et al., 2014). In addition, some analyses may be further complicated by the presence of multiple dependent variables measured through intensive longitudinal methods that may be correlated (Lishinski et al., 2022). In such cases, new methods can facilitate the estimation and interpretation of both the sources of variation in outcomes as well as the effects of predictors, Bayesian statistical methods, or methods that use Markov Chain Monte Carlo (MCMC) estimation (Hadfield, 2010). One affordance of such methods is purely pragmatic: they can be used to estimate highly complex models, including, for instance, those with a cross-classified and nested structure and those with multivariate outcomes. In addition, these methods have some conceptual and *interpretive* benefits, benefits that align with calls to understand variation in the constructs we study first before we study what relates to or impacts these constructs (Yarkoni, 2021). As Levy (2016) describes in an overview of Bayesian methods in the context of educational research, they "allow us to use the language of probability to directly discuss what

Table 1 Summary of datasets

Dataset	Type of learning environment	Data collection method	*N* Responses	*N* Individuals	*N* Moments	*N* Classrooms
Moments in Science	Formal	ESM	4136	244	485	12
Science Learning Environments	Formal	End of Class Report	6610	726	315	29
OST STEM	Informal	ESM	2970	203	248	9
Total			13,716	1173	1048	50

is of inferential interest—parameters, hypotheses, models, and so on—rather than indirectly as in frequentist approaches" (p. 370). In the context of our interest in the variation in situational engagement, these methods allow us to summarize how much variation in engagement can be attributed to different levels (i.e., learner, situation, and classroom).

The present study

This study is intended to explore how learners' situational engagement varies across the levels of individuals, situations, and classrooms using a collection of three data sets that all involve dependencies in learners' reports of their situational engagement—and which are amenable to the kind of analyses using cross-classified, multi-level modeling that can be used to understand where variation in learners' engagement is. Two research questions guide this study:

1. How much variability within each dimension of situational engagement can be attributed to *individual*, *situational*, and *classroom* sources?
2. How do the sources of variability differ across the *behavioral*, *cognitive*, and *affective* dimensions of engagement?

Based on prior research on situational, learner, and classroom effects, we hypothesize that each of these sources of variability for learners' situational engagement in science is appreciable (i.e., non-zero). In the absence of prior research on differences between behavioral, cognitive, and affective dimensions of engagement, we hypothesize that any differences between these dimensions are comparable.

Method

Data sources

Three extant projects with intensive longitudinal datasets were used in this study. First, the *Science in the Moment* project (Schmidt & Shumow, 2012) provided a descriptive account of learners' subjective experiences in high school science classrooms. Data from this project were gathered in 2008–2009 from 244 learners in 12 regular-track high school science classrooms taught by 13 teachers.[1] A more complete description is provided in Additional file 1: Supplementary Material A (which also contains brief descriptions of the other two projects).

The second project, (Schmidt et al., 2015), tested whether classroom interventions targeting mindset and utility value had measurable effects on learning and situational engagement. Data from the second project were gathered in 2011–2012 from 726 learners in 7th-grade general science and 9th-grade integrated science classrooms (29 classrooms total with four 7th-grade teachers and six 9th-grade teachers).

The third project, *STEM Interest and Engagement* (Schmidt et al., 2020), examined how interest and situational engagement developed for underserved youth during summer science programs and created an educational toolkit for youth leaders. Data for the third project were collected during the summer of 2015 from 203 youth enrolled in nine different 4-week summer programs focused on science as well as broader STEM content. For most of the programs, students participated in both fieldwork activities and STEM instruction.

Summary of data collection procedures

As summarized in Table 1, the three extant studies drawn upon for the present analysis comprise a large sample from which to examine the variation in learner engagement in science learning environments.

Two of the three sources involved signaling learners multiple times within their classes, while one used a *daily diary approach* (via end-of-class reports described earlier). See Additional file 1: Supplementary Material A for a detailed description of the specific data collection procedures associated with each study. As we describe later, we account for differences across studies in our data analysis by modeling the unique effect of the data being associated with each of the three extant studies from which we utilized data.

[1] A new teacher was assigned to one of the general science classrooms due to staff changes.

Table 2 Demographic characteristics of participants across studies

	Science in the moment ($n = 244$)	Incremental mindset and utility for science learning and engagement ($n = 726$)	STEM interest and engagement study ($n = 203$)
Sex			
Male	53%	49%	50%
Female	47%	51%	50%
Race/ethnicity			
American Indian	1%	< 0.5%	
Asian/Pacific Islander	2%	2%	7%
Black	12%	12%	36%
White	37%	19%	6%
Hispanic	42%	61%	48%
Multi-racial	6%	6%	3%
Age			
10			4%
11		< 0.5%	28%
12		39%	31%
13		12%	21%
14	32%	35%	12%
15	26%	13%	3%
16	29%	1%	1%
17	11%		
18	2%		
Free/reduced lunch eligible	43%	71%	90%
Parental education			
High school or below	34%	36%	79%
Some college	16%	12%	
Graduated from college	20%	12%	21%
Advanced degree	15%	10%	
Do not know	15%	30%	

Learners considered at least one parent's degree attainment for parent education

Participants

As Table 2 illustrates, the 1173 participants in this study were diverse in terms of their sex, ethnicity, eligibility for subsidized lunch programs, and parents' highest level of education attained. Learners were middle- and high school-aged, ranging from 10 to 18 years old, with most being between 12 and 16.

Measures

For the measures of situational engagement, we used items with four-point Likert-type scales (Hektner et al., 2007) and created composite variables (comprised two items each) for each dimension. We note that many studies utilizing ESM utilize shorter scales that are administered more frequently (Beymer et al., 2022; Zirkel et al., 2015). The trade-off between the time it takes participants to respond and the number of items in a scale represents a tension that must be balanced in research using intensive longitudinal methods (Hektner et al., 2007).

We operationalized behavioral engagement as hard work and concentrated effort (Conner & Pope, 2013; Schmidt et al., 2018). Behavioral engagement was computed by taking the mean of learners' responses to the questions, "How hard were you working?" and "How well were you concentrating?" ($\alpha = 0.82$).

We operationalized cognitive engagement as the value and importance a learner places on the task he or she is completing (Conner & Pope, 2013; Schmidt et al., 2018). Cognitive engagement was computed by taking the mean of learners' responses to the questions, "How important was what you were doing to you?" and "How important was it to your future?" ($\alpha = 0.83$).

We operationalized affective engagement as a learner's interest and enjoyment in a given task (Conner & Pope, 2013; Schmidt et al., 2018). Affective engagement was computed by taking the mean of learners' responses to the questions, "Was this activity interesting?" and "Did you enjoy what you were doing?" ($\alpha = 0.84$)?

Data analysis

To model the variation that occurs at the individual, situational, and classroom levels, we used Bayesian mixed-effects (or multi-level or hierarchical linear) models, wherein the groups (each of the levels) are modeled with random effects (Gelman & Hill, 2006; West et al., 2014). Additionally, the model was multivariate, as the covariance between the three dimensions of situational engagement was modeled. Thus, a multivariate, cross-classified mixed-effects (or multi-level) model for the three dimensions (cognitive, behavioral, and affective) of situational engagement as continuous outcomes was used. We used the *brms* software package (Bürkner, 2017), which provides an interface to the *Stan* software (Carpenter et al., 2017) for estimating multivariate, multi-level models via *R* (R Core Team, 2021).

Using the notation outlined by Gelman and Hill (2006), our model was specified as follows, where α, on the first line, represents the model intercept, which randomly varied between individuals, situations, and classrooms—and which was predicted by the fixed effect for the dataset.

$$\begin{aligned}
\text{Dimension of engagement}_i &\sim N(\alpha_{j[i]} + \beta_1(\text{Dataset}), \sigma^2) \\
\alpha_j &\sim N(\mu_{\alpha_j}, \sigma^2_{\alpha_j}), \text{for Individual } j=1,\ldots \text{J} \\
\alpha_j &\sim N(\mu_{\alpha_k}, \sigma^2_{\alpha k}), \text{for Situation } k=1,\ldots \text{K} \\
\alpha_j &\sim N(\mu_{\alpha l}, \sigma^2_{\alpha_l}), \text{for Classroom } l=1,\ldots \text{L}
\end{aligned}$$

The second line of the equation represents within-student variation, which was assumed generated by a normal distribution; the third line of the equation represents between-student variation, which was similarly assumed to be generated from a normal distribution; and the fourth line represents between-classroom variation. Finally, the random effects and the residuals were modeled with a multivariate normal distribution and as is common for multivariate multi-level models (Hadfield, 2010), the correlations between the residuals at each level were estimated from the data. We note that analysts might consider adding an auto-regressive term (Lishinski et al., 2022; Patall et al., 2018). We think this is a worthwhile consideration, though the magnitude of auto-regressive effects likely depends a great deal on the specifics of the data collection process.

After estimating the model, we used the region of practical equivalence (ROPE) method, a hypothesis testing technique used as a part of Bayesian methods (Kruschke, 2015; Makowski et al., 2019). The ROPE technique involves constructing an interval around a null hypothesis as the ROPE, and then determining what proportion of the distribution for the parameter estimate overlaps with the ROPE. If less than 5% of the distribution overlaps with zero, then it can be concluded that the parameter differs from zero. We used a ROPE of 0.00 to 0.03, as we considered a parameter value for an intraclass correlation (ICC) within this ROPE to represent an *ICC* that may be practically and substantively equal to zero.

Using the ROPE, for RQ #1, we tested whether individual, situational, or classroom sources of variability differed within the dimension of situational engagement; for RQ #2, we tested whether the sources of variability differed between dimensions. We also accounted for any mean level differences in the mean levels of the outcomes that might be present by including a fixed effect, dummy-coded variable for the three studies.

Table 3 Means, standard deviations, and correlations with confidence intervals for the dimensions of engagement

Variable	*M* (SD)	1	2
1. Behavioral engagement	1.96 (0.88)	–	
2. Cognitive engagement	1.42 (0.99)	0.49 [0.48, 0.50]	–
3. Affective engagement	1.59 (0.95)	0.60 [0.59, 0.61]	0.57 [0.56, 0.58]

Each variable was measured on a four-point scale (0–3). Values in square brackets indicate the 95% confidence interval for each correlation

Results

Descriptive statistics

We first report descriptive statistics for the study variables (Table 3), including the means, standard deviations, and correlations. These descriptive statistics show that learners' engagement across the dimensions was around the midpoint of the scale—1.5). Learners generally reported higher behavioral engagement than affective and cognitive engagement, and these variables were moderately correlated, with behavioral and cognitive engagement demonstrating a correlation smaller in magnitude than the other pairs of correlations.

The variation in situational engagement explained by source (RQ #1)

For this question on the proportion of the variation in learners' situational engagement explained by source, we interpreted the ICCs to understand the total amount of variability associated with each source, as presented in Fig. 1. The specific values are presented in Additional file 1: Supplementary Material B. To interpret this figure, consider affective engagement. This pane of the figure represents the estimated ICCs for classroom, situation, and individual sources for this dimension of engagement. Considering classroom sources, the estimated ICC of 0.354 indicates that 35.4% of the variability in affective engagement can be attributed to individual learner differences. 10.1% can be attributed to situational sources, and 4.9% to classroom sources. The distributions of each of these suggest that they all differ from each other—and from zero. The sources of variation for the other

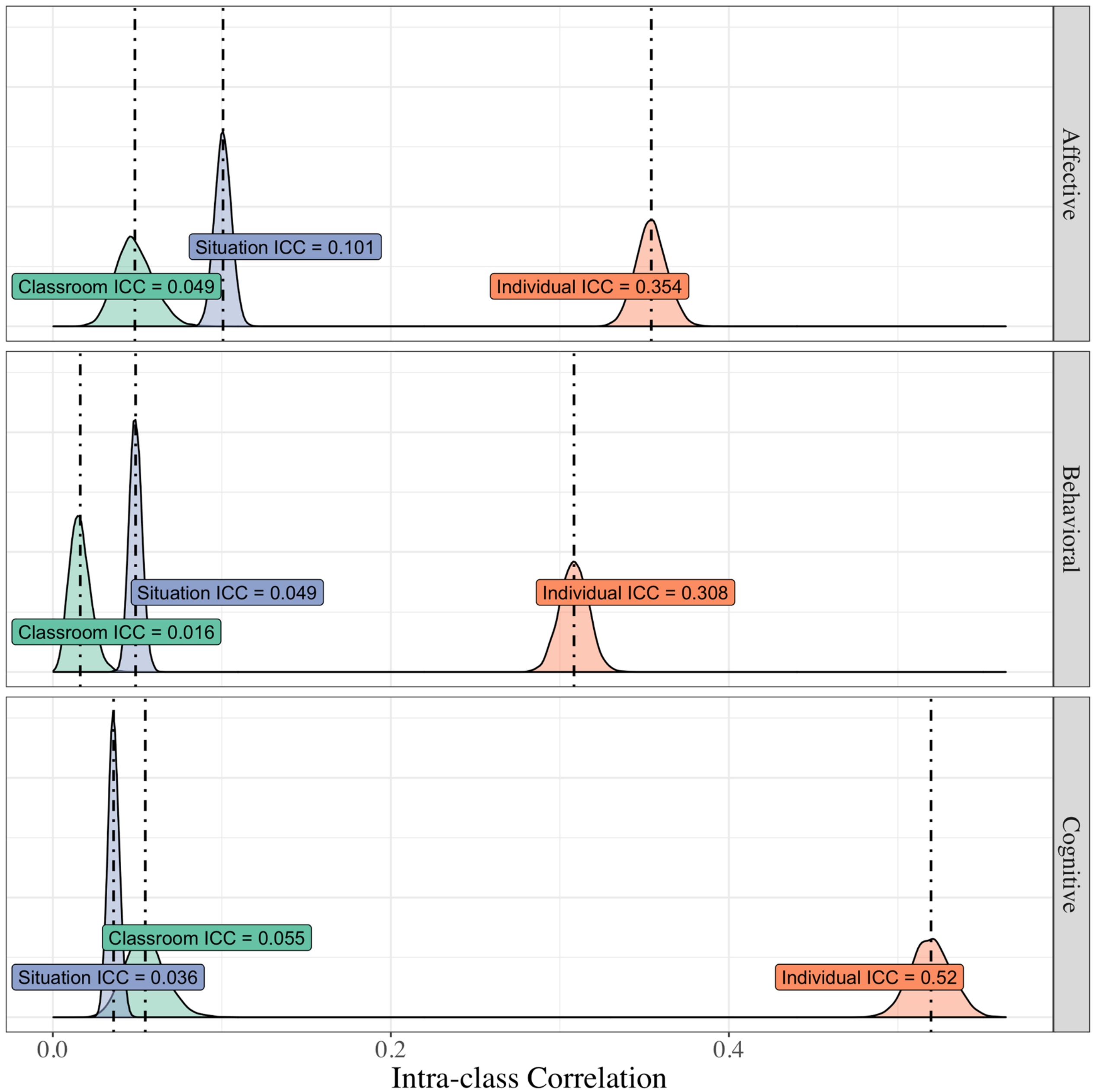

Fig. 1 Intraclass correlations (ICCs) for each source of variability for engagement by dimension

dimensions of engagement can be interpreted in the same manner.

The magnitude of the individual source of variation for the affective engagement dimension compares to the magnitude of behavioral engagement (ICC = 0.308). Individual learners explain even more of the variation in the cognitive (ICC = 0.52) dimension. Generally, the situation was associated with more modest ICC values, ranging from 0.036 for cognitive engagement to 0.101 for affective engagement. In short, more variation in affective engagement can be attributed to specific situations than for behavioral and especially cognitive engagement. The classroom was, overall, the smallest source of variation for engagement, being associated with ICC values that ranged from 0.016 for behavioral engagement to around 0.05 for both affective and cognitive engagement. Indeed, the only ICC that we determined was not statistically significant was that for classroom sources of variability for behavioral engagement, which had a 95% credible interval that overlapped with our ROPE, indicating that this

Table 4 Hypothesis tests for RQ #1 (differences in engagement by the source of variation)

Outcome	Hypothesis about level of engagement	*M* difference in variance (SE)
Affective	Individual–situation	0.257 (0.017), $p<0.001$
Affective	Individual–classroom	0.348 (0.034), $p<0.001$
Affective	Situation–classroom	0.091 (0.032), $p=0.004$
Behavioral	Individual–situation	0.31 (0.016), $p<0.001$
Behavioral	Individual–classroom	0.397 (0.029), $p<0.001$
Behavioral	Situation–classroom	0.088 (0.028), $p=0.001$
Cognitive	Individual–situation	0.494 (0.018), $p<0.001$
Cognitive	Individual–classroom	0.451 (0.037), $p<0.001$
Cognitive	Situation–classroom	0.043 (0.033), $p=0.089$

Table 5 Hypothesis tests for RQ #2 (differences in engagement by the dimension of engagement)

Level	Hypothesis about dimension of engagement	*M* difference in variance (SE)
Individual	Affective–behavioral	0.037 (0.012), $p=0.002$
Individual	Behavioral–cognitive	−0.154 (0.017), $p<0.001$
Individual	Affective–cognitive	−0.117 (0.015), $p<0.001$
Situation	Affective–behavioral	0.09 (0.011), $p<0.001$
Situation	Behavioral–cognitive	0.03 (0.011), $p=0.003$
Situation	Affective–cognitive	0.119 (0.013), $p<0.001$
Classroom	Affective–behavioral	0.087 (0.028), $p=0.001$
Classroom	Behavioral–cognitive	−0.101 (0.029), $p<0.001$
Classroom	Affective–cognitive	0.014 (0.032), $p=0.323$

estimate was not statistically significantly different from an ICC of 0.

In addition to determining whether the estimated *ICCs* for each source of variation differed from zero, we also examined whether they differed from each other—doing so for each dimension of engagement. Table 4 presents these results, which show that each source of variation differed in magnitude from the other source of variation, apart from situational and classroom sources for cognitive engagement. This affirms what an interpretation of Fig. 1 suggests: within each dimension, the sources of variation differ from each other.

Differences across the dimensions of engagement (RQ #2)

For this question, we compared ICCs for the sources of variation in engagement across the three dimensions of engagement. In this way, while RQ #1 informed us about how the sources—individual, situational, and classroom—differed in magnitude from the others *within* each dimension of engagement, this analysis for RQ #2 informed us as to whether the magnitude of each *ICC* differed *between* each dimension of engagement.

As Table 5 demonstrates, this analysis shows us that—for example—while 52% of the variability in cognitive engagement is attributable to individual sources, these individual sources account for only 31% and 35% of the variability in behavioral and cognitive engagement, respectively. In other words, there is greater variability in cognitive engagement that can be explained by individual factors relative to the other engagement dimensions. Furthermore, the situational factors account for a greater proportion of the variability in affective engagement relative to cognitive and behavioral engagement, and classroom factors explain less variation in behavioral engagement than for the other dimensions. In short, cognitive engagement seems to be particularly associated with differences between individuals, and affective engagement is relatively situational in nature; behavioral engagement lies between these in what explains its variation.

In addition to the parameters interpreted using the ROPE strategy to answer RQs #1 and #2, we also provide the estimated correlations among the three dimensions of engagement and the intercepts for each dimension of engagement in Additional file 1: Supplementary Materials C and D.

Discussion

We drew on three data sets using intensive longitudinal methods (one of which used end-of-class reports, and two of which used the experience sampling method), which were uniquely suited to the aim of quantifying the three sources of variation in science learners' situational engagement. In the two subsections that follow, we discuss the substantive and methodological implications of these findings.

Key findings

Across all three engagement dimensions, we found that appreciable variability in situational engagement was attributable to individual learners, specific situations, and broader classroom context sources. These findings provide empirical support for theoretical conjectures that factors at multiple levels impact situational engagement (e.g., Reschly & Christenson, 2012; Skinner & Pitzer, 2012). Data aggregated across three distinct studies suggested that learners' situational engagement in the behavioral, cognitive, and affective domains is a function of who they are (i.e., how inclined they are to be engaged in a science class), what is happening in particular moments in the classroom (i.e., what investigation-related activities they are engaged in), and the nature of the classroom (i.e., classroom norms and the classroom community).

Individual sources of variability were largest, followed by situational sources

While our analyses and overall findings demonstrated that situational, individual, and classroom factors all contribute substantially to students' engagement, these levels of influence did not contribute in equal measure to the variation we observed within or across the engagement dimensions. Within each of the three engagement dimensions, individual sources of variation explained the greatest share of situational engagement, accounting for 30–50% of the total variation observed—a significantly greater proportion than situational or classroom sources of variation. The second greatest share of variation within each of the engagement dimensions was attributable to the level of the situation—or what was going on at the time the engagement collection took place. The smallest share of variation was observed at the level of the classroom, though it is important to note that for the cognitive dimension of engagement, the proportion of variance explained by the situational vs. classroom levels did not differ significantly from one another.

That the largest share of variation in situational engagement was at the individual learner level affirms claims made by both practitioners and researchers that learners bring with them a host of abilities, beliefs, proclivities, and orientations that together shape how they engage in class across a variety of learning situations. This finding also accords with research demonstrating the appreciable effects of specific learner-related factors on their situational engagement in science (e.g., Beymer et al., 2018; Schmidt et al., 2020). Most recent research that uses intensive longitudinal designs to study situational engagement attempts to model this dependency in the data, accounting for the grouping of student responses within individuals (Mo et al., 2013; Schmidt et al., 2015). Our finding that across multiple studies, such significant portions of variance are observed at the person level reaffirms the merits of this approach. We note that this variation was greatest for individuals, which accounted for just over 50% of the variation in cognitive engagement, whereas 30–35% of the variation in affective and behavioral engagement is attributable to the individual student. This suggests that cognitive engagement is *more* of a trait than a state for students than is affective and behavioral engagement, which vary less between individuals and—as we discuss later—are more variable at the situational level.

Situational sources of variability were smaller than individual sources but were still notable

Situational variation has been much less commonly modeled in intensive longitudinal studies of classroom engagement, and our results suggest that modeling this source of variation may be important. This finding accords with how many recent studies on situational engagement in science have documented substantial situational effects—especially the effects of engaging in particular science and engineering practices (Inkinen et al., 2019, 2020; Schmidt et al., 2018). While less variance in engagement was explained at this level than at the person level, our results show that the instructional choices that teachers make on a given day or at each moment may have a measurable and systematic influence on how learners in their classroom engage at that moment. In addition to reporting *that* situational sources of variation are notable, we also found evidence of variation in the dimension of engagement that varies at the situational level. Cognitive engagement was the least dynamic, varying little from situation-to-situation relative to the other two dimensions, while affective engagement and behavioral engagement were more dynamic in that they varied more across situations. This suggests that affective and behavioral engagement may be more malleable by the decisions teachers make or the practices they deploy in their classrooms, while cognitive engagement may be shaped over a more extended period.

Classroom-level sources of variability were the smallest

Finally, the classroom—more distal than the individual or moment (and modeled as such in this and other studies, as level three)—still had a detectable bearing on learners' engagement, explaining between roughly 2–5% of the variation in learners' engagement, with the least associated with behavioral engagement. Behavioral engagement was associated with the least classroom-level variation in engagement compared to affective and cognitive engagement. Across dimensions, these amounts are smaller than teacher effects (comparable to our classroom effects) on achievement, which have been estimated as ranging from around 6%-13% (Nye et al., 2004). Why might variation in engagement at the classroom level be smaller than it is for achievement? As noted earlier, less research has documented instructor or classroom-level differences in engagement. We think that one possibility is that some of the variation at the classroom level could be attributable to the things that teachers do in situations and instructional episodes—things such as providing support to students (Strati et al., 2017).

The role of intensive longitudinal methods

To address the challenge of studying engagement in a more dynamic way and to establish what the sources of variability in learners' engagement are, we used intensive longitudinal methods and advances from Bayesian statistical methods to study many youths' responses to questions about their engagement. A key methodological implication of this study is that situational factors can—and perhaps must for conceptual and statistical reasons (i.e., bias)—be analyzed using suitable techniques. First, it is conceptually important to conceive of and measure engagement at the situational level through a data collection method such as intensive longitudinal methods. In addition, it may be worthwhile for future research to assess the impact of situational factors that are both internal (e.g., how challenging the activity was for learners; Schmidt et al., 2017) and external (e.g., how teachers support learners' autonomy in learning environments; Patall et al., 2019; Strati et al., 2017). Learners' perceptions of their overall challenge or the extent to which teachers supported their autonomy, in general, do not align with how these factors—like engagement—are dynamic and changing from moment-to-moment.

Also, it may be statistically important to account for the cross-classified nature of a great deal of data collected on learners' engagement. As is well-established in the literature on hierarchical linear or multi-level models, the failure to model the dependencies in outcomes that result from the structure of the data can lead to bias (Raudenbush & Bryk, 2002; West et al., 2014). As repeated measures data collected through ESM (Hektner et al., 2007; Zirkel et al., 2015) and comparable techniques (e.g., Bolger & Laurenceau, 2013) have become more common, scholars have also become more familiar with modeling responses "nested" within learners; see Grimm et al. (2016) for numerous examples using multi-level or structural equation models. These are not just statistical issues but also substantive ones. By modeling the moment, one can examine the impact of specific instructional activities on learner engagement, whereas modeling the classroom allows you to test for effects of teacher characteristics like years of experience or broader classroom variables, such as features of the classroom environment.

What is much less established is the use of multilevel models for other sources of dependency, especially those that are situational in nature. This may be because, in some cases, it may not be necessary to model them. A contrast between ESM and longitudinal methods is instructive for establishing when situational factors must be modeled. In the case of longitudinal methods, patterns of change over time within individuals are of interest (Grimm et al., 2016), and when and where participants respond are often not of interest: learners responding to a survey over the course of one week are very likely doing so at different time points and from different contexts (e.g., in a study room and in their residence hall). In the case of intensive longitudinal methods—particularly ESM—responses are often intended to be a random sample of participants' experiences (Hektner et al., 2007), and in classrooms, participants are often signaled to respond to very brief surveys at the same time. The instructor, for example, was likely orchestrating learners' involvement in similar (or the same) activities, and learners' peers were likely to affect their experience in the period immediately before they were asked to respond.

For these reasons and based on our findings showing the substantial variation at the situational level, even after accounting for variability at the individual learner and classroom level, the rarity with which such factors are modeled in suitable ways is likely the result of what Judd et al. (2012) refer to as a *pervasive but largely ignored problem*. Where Judd et al. described the importance of modeling stimuli (in psychological experiments) using random effects in multi-level models, we believe the same problem applies to research using intensive longitudinal methods. We note that Judd et al. are far from the first to identify this problem: it is present whenever cross-classification is present (e.g., Kadengye et al., 2014), and psychometricians regularly model item effects (Levy & Mislevy, 2017). Until recently, much statistical software has struggled to estimate data with cross-classified structures, though this is no longer the case (West et al., 2014).

The data analytic approach we used has utility in the growing educational research using intensive longitudinal methods; not modeling this dependency suggests that the effects of situational factors—like the aforementioned internal and external situational effects that are often of theoretical interest (e.g., Strati et al., 2017) may be estimated with bias, and researchers may be more confident about their significance than were the dependencies modeled using a suitable technique. The larger the source of variation from a grouping factor (e.g., individuals, situations, or classrooms) is, the more important it is to include the group in analyses (Raudenbush & Bryk, 2002; West et al., 2014). With the amount of variation at the situational level being substantial (between 4 and 10%), analyses that do not account for this dependency may be associated with bias in their results.

Limitations and recommendations for future research

There were some differences in the three datasets we used, especially regarding how the data were collected and how comparable items were worded. Our analysis (using all three datasets within one model) meant that we could estimate the variation in situational engagement at the situational, individual, and classroom levels, but this means that specific differences by project and site were not examined. One difference concerns how the data were collected in one study (via daily diaries) relative to the other two (via ESM)—though we note that both are instances of the broader grouping of intensive longitudinal methods (Bolger & Laurenceau, 2013). To account for this, we modeled the effect of the different datasets alone to adjust for any mean-level differences in engagement attributable to these different features. But we think that future studies can explore sources of variation in engagement in other contexts and with other data-generating processes (e.g., learners' interactions with educational technology tools that generate *digital traces* of their interactions; D'Mello et al., 2017; Gobert et al., 2015).

Another notable limitation concerns our reliance on self-report measures. While intensive longitudinal methods have many strengths, they still suffer from some of the issues associated with self-reports—particularly the potential bias associated with what active reports by individual learners about their experiences. Some research has explored other means of studying engagement, including using wearable technology to measure electrodermal activity (e.g., Lee et al., 2019) or using teachers' reports of student engagement (e.g., Lee & Reeve, 2012). More generally, we acknowledge that there is a great deal of variation in how the different dimensions of engagement are defined (Reschly & Christenson, 2012) and that our conclusions about where the variability lies may depend on how we are operationalizing each dimension. For example, we define cognitive engagement as referring to the value a given student places on their science tasks. This operationalization may be less situationally variant than others. Other scholars operationalize cognitive engagement in terms of depth of processing or strategy use (Fredricks et al., 2004).

Lastly, we note that several of the data sets were collected before recent reform efforts (i.e., the Next Generation Science Standards [NGSS]; NGSS Lead States, 2013). Still, the teachers participating in these studies drew upon inquiry-based science teaching practices that bear a semblance to the types of teaching called for in the NGSS, and so we think this data remains relevant for the question of how students engage in typical science classrooms in the United States.

Future research can consider using intensive longitudinal methods to explore variability in learners' situational engagement and other motivation and learning-related constructs, such as situational interest and self-efficacy. Also, we think exploring how differences in the overall data collection method (i.e., ESM versus daily diary) and other differences in data collection choices may impact findings, and so we think that studies that quantify the variability in dynamic constructs at individual, situational, and learning environment levels is merited. Similarly, as out-of-school-time becomes recognized as an important setting for learners to develop capabilities that may be challenging for them to develop in formal, school-based settings (Azevedo & Sherin, 2012), and ESM may be a useful approach for studying the development of such outcomes.

Conclusion

We found that learners' situational engagement in science is associated most with individual sources, followed by situational and then classroom sources. There were also differences between dimensions of engagement: cognitive engagement is explained more (relative to behavioral and affective engagement) by individual sources, whereas affective engagement is explained more by situational sources. This suggests that cognitive engagement, at least as it is defined in this study (in terms of value), may be less malleable at the situational level than affective engagement—with behavioral engagement lying between these two dimensions. This work provides foundational evidence about when and where there are differences in how learners engage—evidence that aligns with calls in the wider field to understand where variation exists using robust analytic methods prior to predicting or attempting to intervene (Yarkoni, 2021). In doing so, this work establishes that theoretical ideas about the dynamic nature of engagement have a strong empirical basis. Also, this study suggests that engagement might not only be an idea that has purchase with teachers, but which—because it is malleable at the situational level—may be a lever through which teachers can engage their students more meaningfully and in ways that impact their experiences in science and STEM classrooms and the other places they learn.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40594-023-00449-0.

Additional file 1. Supplementary Materials.

Acknowledgements

The authors wish to acknowledge N. Naftzger, L. Shumow, A. Durik, M. C. Smith, and D. Moroney for their role in the data collection and to C. Pritchard for assistance with preparing the manuscript for review.

Author contributions

JR analyzed the data and was the primary author of the manuscript. PB wrote sections of the manuscript and assisted with the data analysis. VP wrote sections of the manuscript and assisted with the data analysis. JS collected the data used in the study and wrote sections of the manuscript.

Funding

This work is supported by NSF Grant Nos. to JS (DRL-1421198, HRD-0827526, HRD-1136143, ECR-1661064) and to JR (BCSER-1937700).

Availability of data and materials

The data for this study are available from the authors on request.

Declarations

Competing interests
The authors have no competing interests to disclose.

Received: 7 February 2023 Accepted: 4 September 2023
Published online: 30 November 2023

References
Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools, 45*(5), 369–386.
Azevedo, F. S. (2018). An inquiry into the structure of situational interests. *Science Education, 102*(1), 108–127.
Azevedo, F. S., & Sherin, B. L. (2012). An evolving framework for describing student engagement in classroom activities. *The Journal of Mathematical Behavior, 31*(2), 270–289.
Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist, 50*(1), 84–94.
Beymer, P. N., Ferland, M., & Flake, J. K. (2022). Validity evidence for a short scale of college students' perceptions of cost. *Current Psychology, 41*(11), 7937–7956.
Beymer, P. N., Rosenberg, J. M., & Schmidt, J. A. (2020). Does choice matter or is it all about interest? An investigation using an experience sampling approach in high school science classrooms. *Learning and Individual Differences, 78*, 101812.
Beymer, P. N., Rosenberg, J. M., Schmidt, J. A., & Naftzger, N. J. (2018). Examining relationships among cfhoice, affect, and engagement in summer STEM programs. *Journal of Youth and Adolescence, 47*, 1178–1191.
Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford Press.
Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*, 1–28.
Cabrera, L., Bae, C., & DeBusk-Lane, M. (2023). A mixed methods study of middle students' science motivation and engagement profiles. *Learning and Individual Differences, 103*, 102281.
Carlone, H. B. (2004). The cultural production of science in reform-based physics: Girls' access, participation, and resistance. *Journal of Research in Science Teaching, 41*(4), 392–414.
Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, 76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01
Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *The handbook of research on student engagement*. Springer Science.
Conner, J. O., & Pope, D. C. (2013). Not just robo-students: Why full engagement matters and how schools can promote it. *Journal of Youth and Adolescence, 42*(9), 1426–1442.
Csikszentmihalyi, M. (1997). Flow and the psychology of discovery and invention. *HarperPerennial*, New York.
D'Mello, S., Dieterle, E., & Duckworth, A. (2017). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist, 52*(2), 104–123.
Eccles, J. S. (2016). Engagement: Where to next? *Learning and Instruction, 43*, 71–75.
Eccles, J., & Wang, M. T. (2012). Part I commentary: So what is student engagement anyway? *Handbook of research on student engagement* (pp. 133–145). Springer, US.
Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059
Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. *Handbook of research on student engagement* (pp. 763–782). Springer, US.
Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. *Educational Psychologist, 50*(1), 43–57.
Greene, B. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist, 50*(1), 14–30.
Grimm, K. J., Ram, N., & Estabrook, R. (2016). *Growth modeling: Structural equation and multilevel modeling approaches*. Guilford Publications.
Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software, 33*(2), 1–22.
Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Sage Publications.
Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.
Inkinen, J., Christopher, K., Barbara, S., Kalle, J., Joseph, K., Jari, L., & Katariina, S. A. (2019). Science classroom activities and student situational engagement. *International Journal of Science Education, 41*(3), 316–329.
Inkinen, J., Klager, C., Juuti, K., Schneider, B., Salmela-Aro, K., Krajcik, J., & Lavonen, J. (2020). High school students' situational engagement associated with scientific practices in designed science learning situations. *Science Education., 104*, 667.
Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*(1), 54.
Kadengye, D. T., Ceulemans, E., & Van den Noortgate, W. (2014). A generalized longitudinal mixture IRT model for measuring differential growth in learning environments. *Behavior Research Methods, 46*, 823–840.
Kruschke, J. (2015). Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan.
Lee, V. R., Fischback, L., & Cain, R. (2019). A wearables-based approach to detect and identify momentary engagement in afterschool Makerspace programs. *Contemporary Educational Psychology, 59*, 101789.
Lee, W., & Reeve, J. (2012). Teachers' estimates of their students' motivation and engagement: Being in synch with students. *Educational Psychology, 32*(6), 727–747.
Levy, R. (2016). Advances in Bayesian modeling in educational research. *Educational Psychologist, 51*(3–4), 368–380.
Levy, R., & Mislevy, R. J. (2017). *Bayesian psychometric modeling*. CRC Press.
Lishinski, A., Narvaiz, S., & Rosenberg, J. M. (2022). Self-efficacy, Interest, and Belongingness–URM Students' Momentary Experiences in CS1. In *Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1* (pp. 44–60)
Makowski, D., Ben-Shachar, M. S., Chen, S. H., & Lüdecke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology, 10*, 2767.
Mehl, M., & Conner, T. S. (2013). *Handbook of research methods for studying daily life*. Guilford Press.
Mo, Y., Singh, K., & Chang, M. (2013). Opportunity to learn and student engagement: A HLM study on eighth grade science achievement. *Educational Research for Policy and Practice, 12*(1), 3–19.
National Academies of Sciences, Engineering, and Medicine. (2018). *How people learn II: Learners, contexts, and cultures*. The National Academies Press. https://doi.org/10.17226/24783
Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*(3), 237–257.
Patall, E. A., Pituch, K. A., Steingut, R. R., Vasquez, A. C., Yates, N., & Kennedy, A. A. (2019). Agency and high school science students' motivation, engagement, and classroom support experiences. *Journal of Applied Developmental Psychology, 62*, 77–92.
Patall, E. A., Steingut, R. R., Vasquez, A. C., Trimble, S. S., Pituch, K. A., & Freeman, J. L. (2018). Daily autonomy supporting or thwarting and students' motivation and engagement in the high school science classroom. *Journal of Educational Psychology, 110*(2), 269.
Patall, E. A., Vasquez, A. C., Steingut, R. R., Trimble, S. S., & Pituch, K. A. (2016). Daily interest, engagement, and autonomy support in the high school science classroom. *Contemporary Educational Psychology, 46*, 180–194.

Patall, E. A., Vasquez, A. C., Steingut, R. R., Trimble, S. S., & Pituch, K. A. (2017). Supporting and thwarting autonomy in the high school science classroom. *Cognition and Instruction, 35*(4), 337–362.
Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. *Handbook of research on student engagement* (pp. 259–282). Springer, US.
R Core Team (2021). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.
Reschly, A. L., & Christenson, S. L. (2012). Moving from "context matters" to engaged partnerships with families. *Journal of Educational and Psychological Consultation, 22*(1–2), 62–78.
Rosenberg, J. M. (2018). *Understanding work with data in summer STEM programs: An experience sampling method approach* (Doctoral dissertation). Retrieved from Proquest Dissertations and Theses. (Proquest No. 10747232)
Schmidt, J. A., Beymer, P. N., Rosenberg, J. M., Naftzger, N. N., & Shumow, L. (2020). Experiences, activities, and personal characteristics as predictors of engagement in STEM-focused summer programs. *Journal of Research in Science Teaching, 57*(8), 1281–1309.
Schmidt, J. A., Kafkas, S. S., Maier, K., Shumow, L., & Kackar-Cam, H. Z. (2019). Why are we learning this? Using mixed methods to understand how teachers' statements shape students' perceptions of science utility. *Contemporary Educational Psychology, 56*, 9–31. https://doi.org/10.1016/j.cedpsych.2018.08.005
Schmidt, J. A., Rosenberg, J. M., & Beymer, P. N. (2018). A person-in-context approach to student engagement in science: Examining learning activities and choice. *Journal of Research in Science Teaching, 55*(1), 19–43.
Schmidt, J. A., & Shumow,. (2012). Change in self-efficacy in high school science classrooms: An analysis by gender. In S. L. Britner (Ed.), *Self-efficacy in school and community settings* (pp. 54–73). Nova Science Publishers.
Schmidt, J. A., Shumow, L. & Durik, A. Incremental Mindset and Utility for Science Learning and Engagement (I-MUScLE): A Quasi-Experimental Study of Targeted Classroom Treatments. National Science Foundation grant HRD-1136143.
Schmidt, J. A., Shumow, L., & Kackar-Cam, H. Z. (2015). Teacher effects in outcomes of a growth mindset intervention in seventh grade science classes. *Middle Grades Research Journal, 10*(2), 17–32.
Schmidt, J. A., Shumow, L., & Kackar-Cam, H. Z. (2017). Does mindset intervention predict students' daily experience in classrooms? A comparison of seventh and ninth graders' trajectories. *Journal of Youth and Adolescence, 46*, 582–602.
Schneider, B., Krajcik, J., Lavonen, J., Salmela-Aro, K., Broda, M., Spicer, J., Bruner, J., Moeller, J., Linnansaari, J., Juuti, K., & Viljaranta, J. (2016). Investigating optimal learning moments in US and Finnish science classes. *Journal of Research in Science Teaching, 53*(3), 400–421.
Schwarz, N. (2012). Why researchers should think "real-time": A cognitive rationale. Handbook of research methods for studying daily life, 22.
Shernoff, D. J., Csikszentmihalyi, M., Shneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly, 18*(2), 158.
Shernoff, D. J., Kelly, S., Tonks, S. M., Anderson, B., Cavanagh, R. F., Sinha, S., & Abdi, B. (2016). Student engagement as a function of environmental complexity in high school classrooms. *Learning and Instruction, 43*, 52–60.
Shernoff, D. J., Knauth, S., & Makris, E. (2000). The quality of classroom experiences. In M. Csikszentmihalyi & B. Schneider (Eds.), *Becoming adult: How teenagers prepare for the world of work* (pp. 141–164). Basic Books.
Shumow, L., & Schmidt, J. A. (2014). Enhancing adolescents' motivation for science: Research-based strategies for teaching male and female students. In D. Meyer & L. H. Anderman (Eds.), *APA's Classroom insights from educational psychology series*. Corwin Press.
Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924
Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*(4), 765.
Skinner, E. A., & Pitzer, J. (2012). Developmental dynamics of engagement, coping, and everyday resilience. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *The handbook of research on student engagement* (pp. 21–45). Springer Science.
Strati, A. D., Schmidt, J. A., & Maier, K. S. (2017). Perceived challenge, teacher support, and teacher obstruction as predictors of student engagement in high school science classrooms. *Journal of Educational Psychology, 109*(1), 131–147. https://doi.org/10.1037/edu0000108
Turner, J. C., & Meyer, D. K. (2000). Studying and understanding the instructional contexts of classrooms: Using our past to forge our future. *Educational Psychologist, 35*(2), 69–85.
Upadyaya, K., Cumsille, P., Avalos, B., Araneda, S., Lavonen, J., & Salmela-Aro, K. (2021). Patterns of situational engagement and task values in science lessons. *The Journal of Educational Research, 114*(4), 394–403.
van Braak, M., van de Pol, J., Poorthuis, A. M., & Mainhard, T. (2021). A microperspective on students' behavioral engagement in the context of teachers' instructional support during seatwork: Sources of variability and the role of teacher adaptive support. *Contemporary Educational Psychology, 64*, 1–14.
West, B. T., Welch, K. B., & Galecki, A. T. (2014). *Linear mixed models: A practical guide using statistical software*. CRC Press.
Xie, K., Vongkulluksn, V. W., Heddy, B. C., & Jiang, Z. (2023). Experience sampling methodology and technology: an approach for examining situational, longitudinal, and multi-dimensional characteristics of engagement. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-023-10259-4
Yarkoni, T. (2021). The generalizability crisis. *Behavioral and Brain Sciences, 45*, e1.
Zirkel, S., Garcia, J. A., & Murphy, M. C. (2015). Experience-sampling research methods and their potential for education research. *Educational Researcher, 44*(1), 7–16.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.