Received: 21 September 2017 | Revised: 30 May 2018 | Accepted: 31 May 2018
DOI: 10.1002/sce.21461

WILEY

RESEARCH ARTICLE

Gender disparities in students' motivational experiences in high school science classrooms

Erika A. Patall[1] | **Rebecca R. Steingut**[2] | **Jen L. Freeman**[3] | **Keenan A. Pituch**[3] | **Ariana C. Vasquez**[3]

[1]Rossier School of Education, University of Southern California, Los Angeles, California

[2]Teacher's College, Columbia University, New York, New York

[3]Department of Educational Psychology, The University of Texas at Austin, Austin, Texas

Correspondence
Erika A. Patall, Rossier School of Education, University of Southern California, 3470 Trousdale Parkway, Waite Phillips Hall, Los Angeles, CA 90089-4036.
Email: patall@rossier.usc.edu

This research was supported by a grant from the William T. Grant Foundation (Project #180042).

Abstract

Women remain underrepresented in physical science, engineering, and computer science college majors and careers. Prior research has suggested that motivational beliefs and experiences in high school play a critical role in girls' persistence in science fields. We hypothesized that compared to male students, female high school students may experience less daily autonomy support from teachers and lower psychological need satisfaction in physics, chemistry, and engineering courses. In turn, we expected that these differences would explain girls' lower daily engagement in these courses compared to boys. In line with current trends indicating gender parity in biology and biomedical fields, we did not expect to find gender differences in biology courses. Results from a six-week intensive longitudinal study in which high school students reported their daily experiences of teacher autonomy support, need satisfaction, and engagement during science class supported our hypotheses. The implications of the results for theory and practice are discussed.

KEYWORDS

autonomy support, engagement, high school, intensive longitudinal design, motivation, need satisfaction, science, teaching practice

Engaging girls in science, technology, engineering, and mathematics (STEM) classes remains a critical challenge for educators. Though they have achieved parity in biological and biomedical degrees and occupations (Snyder, de Brey, & Dillow, 2016), women remain underrepresented in non-biological and biomedical STEM college majors and careers, including physics, engineering, and computer sciences (National Science Foundation, 2017; Riegle-Crumb, & King, 2010). These gender disparities persist even though girls obtain higher course grades in math and science than boys and are just as likely to be enrolled in advanced STEM courses in high school (e.g., Hyde, Lindberg, Linn, Ellis, & Williams, 2008; National Science Foundation, 2017; Riegle-Crumb, King, Grodsky, and Muller, 2012; Snyder et al., 2016).

This underrepresentation of women in the physical sciences (e.g., physics and chemistry) and engineering areas of STEM is troubling as increasing female participation is paramount to meeting increasing demands for STEM professionals equipped to meet the challenges of the 21st century in the current global marketplace (National Academy of Sciences, 2007a,b). Further, discoveries in STEM are likely to be enhanced by diversity of people and ideas (Allen-Ramdial & Campbell, 2014). Finally, this underrepresentation of women is troubling from a social equity perspective, as STEM careers, particularly in engineering and computer science, are linked with greater economic prosperity and social status (Cataldi, Siegel, Shepherd, & Cooney, 2014).

Undoubtedly, girls' classroom experiences and motivational beliefs during high school play a critical role in predicting whether they persist in STEM fields into college and for their careers (e.g., Legewie & DiPrete, 2012; Tai, Liu, Maltese, & Fan, 2006). However, thus far, little research has focused on gender differences in daily perceptions of teachers' motivating practices and subsequent motivation and engagement during class. In particular, a substantial body of research based on self-determination theory (Ryan & Deci, 2000) has suggested that when teachers are perceived to engage in practices that encourage students to see themselves as the initiator of their actions, known as autonomy-supportive practices, students' experience numerous benefits. These benefits include greater psychological need satisfaction, that is, a greater sense of autonomy, competence, and relatedness, as well as enhanced motivation and engagement (e.g., Assor, Kaplan, & Roth, 2002; Patall, Dent, Oyer, & Wynn, 2013; Reeve & Jang, 2006; Reeve, Jang, Carrell, Jeon, & Barch, 2004). With this in mind, we asked the following research questions.

1. Do high school girls and boys differ in their experiences of teachers' autonomy support, need satisfaction, and engagement in science courses, particularly in physical science courses?
2. Do gender differences in students' experiences of teachers' autonomy support and need satisfaction explain gender differences in students' engagement in science courses, particularly in physical science courses?

This research investigated the intriguing novel hypothesis that students' daily experiences of autonomy support and psychological need satisfaction is lower for girls than for boys in physics, chemistry, and engineering courses, given pervasive negative stereotypes about girls' suitability particularly for physical and engineering science fields (e.g., Leslie, Cimpian, Meyer, & Freeland, 2015; Wang & Degol, 2017). Moreover, we hypothesized that these gender differences in daily experiences of teachers' autonomy support and need satisfaction would explain girls' lower daily engagement in these courses compared to boys. Finally, given that women are well represented in biology and biomedical fields and that social perceptions of those fields are less inconsistent with social roles traditionally ascribed to women (e.g., Leslie et al., 2015), we hypothesized that such gender differences would not be found in biological courses. To examine these hypotheses, we conducted a six-week intensive longitudinal study with high school students in which students provided daily reports of their experiences of teacher autonomy support, psychological need satisfaction (for autonomy, competence, and relatedness), and engagement in science class each class day. Repeatedly assessing students' perceptions of teacher practice and students' psychological need satisfaction and engagement over class days provided an opportunity to collect strong evidence regarding gender differences in students' daily experiences in science class.

1 | EXPLAINING FEMALE UNDERREPRESENTATION IN STEM

Explanations for gender disparities in physical and engineering science fields, fields that emphasize the study of inanimate objects and have a strong inclusion of mathematics, abound. Though differences in ability, particularly mathematical and spatial ability, have often been implicated to explain disparities (e.g., Baron-Cohen, 2003; Halpern et al., 2007), most researchers agree that such ability differences are small (if they exist) and do not seem to fully account for female underrepresentation in STEM careers (Ceci, Williams, & Thompson, 2011; Wai, Cacchio, Putallaz, & Makel, 2010). Rather, the literature on gender differences suggests that despite comparable levels of achievement and STEM course-taking among male and female students prior to college, lower female persistence in physical and engineering science fields is likely explained by a complex set of motivation-related beliefs shaped by experiences in the home, school, and society. Moreover, students' experiences during high school are particularly critical in shaping beliefs and determining whether students pursue a STEM college major and career (Legewie & DiPrete, 2012; Tai et al., 2006).

Girls decide in high school to abandon the pursuit of a science or engineering education at a higher rate than boys, even if they previously expressed interest. Such decisions made in high school are critical, because the highly prescribed nature of the coursework for STEM majors makes it difficult for students to switch into a STEM field following the start of college (e.g., Wang & Degol, 2013).

Along these lines, research has routinely documented gender differences in adolescents' motivational beliefs and attitudes that promote persistence in STEM. In particular, girls experience lower perceived competence and interest in STEM during adolescence and perceive greater costs of pursuing careers in STEM fields than boys (e.g., for review see Wang & Degol, 2013, 2017). For example, high school boys tend to rate their math competence higher than girls with similar grades and test scores in math, both globally (e.g., Correll, 2001; Nagy et al., 2008), and when they are asked to rate their skill on a particular task on any given day in the science classroom (e.g., Shumow & Schmidt, 2013). Research suggests that high school girls' value and interest toward science are lower than of boys (e.g., Frenzel, Goetz, Pekrun, & Watt, 2010; Shumow & Schmidt, 2013; Watt, 2004), and some prior research suggests that unlike boys, girls' interest in math and science decreases in adolescence (Eccles & Harold, 1992; Köller, Baumert, & Schnabel, 2001; Sadler, Sonnert, Hazari, & Tai, 2012). These gender differences in perceived competence or self-efficacy, value, and interest are important given the well-established links of these variables to achievement-related behaviors and performance, as well as academic and career decision-making (e.g., for review, see Wang & Degol, 2013, 2017). Indeed, these factors are likely to be equally or more important as ability and achievement in explaining students' STEM persistence. For example, adolescents with average math achievement and high interest in pursuing a science occupation are far more likely to complete a science degree in college than adolescents with high math achievement but little interest in pursuing a STEM occupation (Tai et al., 2006).

While the importance of adolescents' motivational beliefs has clearly been recognized, few studies have explored gender differences in students' *daily experiences* in high school STEM classrooms, despite that ongoing daily experiences in the classroom and other social contexts help to mold the enduring motivational beliefs predictive of gender disparities (e.g., for a review of sociocultural influences on motivational beliefs, see Wang & Degol, 2013, 2017). Thus, in this study, we focus on discrepancies in students' daily engagement during science class: we examine the extent to which students are actively involved in tasks or activities in class each day, as well as the environmental and psychological experiences during class that shape that engagement. Engagement is a multidimensional construct that plays a critical role in students' learning (Archambault, Janosz, Fallu, & Pagani, 2009; Fredricks, Blumenfeld, & Paris, 2004; Reeve & Tseng, 2011; Skinner, Kindermann, Connell, & Wellborn, 2009). *Behavioral* engagement describes students' involvement in class in terms of their effort, attention, persistence, and participation. *Emotional* engagement describes students' involvement in class in terms of interest, enjoyment, and other positive emotions. *Cognitive* engagement describes students' involvement in class in terms of their attempts to regulate the process of learning. *Agentic* engagement describes students' involvement in terms of their proactive attempts to influence the instruction so that it better supports their own motivation and learning. Although classroom engagement has yet to be the focus of much research focused on gender disparities in STEM, we focus on engagement (across these dimensions) because it is a *malleable* critical mediator between the school environment and student learning, school success, and persistence in STEM fields (Croninger & Lee, 2001; Fredricks et al., 2004; Jang, Kim, & Reeve, 2012; Ladd & Dinella, 2009; Sinatra, Heddy, & Lombardi, 2015).

2 | TEACHER AUTONOMY SUPPORT

Students' engagement is facilitated by high-quality interactions with teachers who consider students' motivation in their practice (Allen, Pianta, Gregory, Mikami, & Lun, 2011). According to self-determination theory, people have fundamental psychological needs for autonomy, competence, and relatedness that underlie their motivation and well-being (Ryan & Deci, 2000). Students thrive when they experience their behavior as emanating from or consistent with their understandings of self (i.e., autonomy), success in interacting with the environment (i.e., competence),[1] and connections with other individuals (i.e., relatedness). However, teachers can also support students' experiences of need satisfaction with their practice.

Autonomy support in the classroom context reflects a motivational approach in which teachers identify, nurture, and develop students' inner motivational resources so that students perceive themselves as the initiator of their actions (Reeve, 2009). Autonomy-supportive teachers offer choices, encourage students to work in their own way or at their own pace, are open to students' opinions and negative feelings, and respond to students' progress with encouraging, informational feedback. Such teachers attempt to structure course activities around students' interests, but when that is impossible, they provide meaningful rationales to explain the usefulness or importance of "boring" course activities (see Patall, Vasquez, Steingut, Trimble, & Pituch, 2017; Reeve, 2009; Reeve & Jang, 2006; Su & Reeve, 2011 for a review of autonomy-supportive practices). Extensive research has indicated that teachers' autonomy-supportive practice and, more proximally, perceiving teachers to engage in autonomy-supportive practices predict students' need satisfaction and, in turn, need satisfaction predicts students' engagement in class (Hafen et al., 2012; Jang et al., 2012; Patall et al., 2017, 2018). For example, Jang and colleagues (2016) found that Korean high school students' perceptions that teachers supported their autonomy predicted changes in need satisfaction, which, in turn, predicted changes in engagement over the course of a school semester.

Moreover, although autonomy support and the experience of need satisfaction, and particularly autonomy need satisfaction, has been found to have widespread benefits across people and contexts, several factors suggest it may be particularly important in high school science classes. First, high school students' perceptions that teachers are supporting their autonomy and the subsequent experience of autonomy need satisfaction may be particularly important because adolescence coincides with an increased need for autonomy and independence (Eccles, Wigfield, Harold, & Blumenfeld, 1993; Erikson, 1968). Second, given that innovation and discovery, appreciation for ambiguity, and building on past discoveries and failures are all fundamental values in science (e.g., Allchin, 1999; Bartos, and Lederman, 2014; Kuhn, 1962), support for personal autonomy may be particularly important to science education.

Taken together, it seems that students' perceptions of teachers' autonomy support and the need satisfaction that such support facilitates are important predictors of engagement in STEM courses, and ultimately persistence in STEM. Given women's underrepresentation in physical and engineering science fields, such relationships beg the question: do high school girls experience the same level of daily teacher autonomy support, need satisfaction, and ultimately, engagement, as boys in their high school physics, chemistry, and engineering courses?

3 | GENDER DIFFERENCES IN STUDENTS' EXPERIENCES IN THE SCIENCE CLASSROOM

Despite the benefits of autonomy-supportive teaching across content areas including STEM (e.g., Black, and Deci, 2000; Reeve et al., 2004), pervasive stereotypes associating the physical and engineering sciences with innate talent that men, and not women, possess (e.g., Leslie et al., 2015; Nosek et al., 2009; Tiedemann, 2002) may influence the extent to which teachers equitably support male and female students' motivation in physical and engineering science classes. Research suggests that teachers stereotype math and science as domains most suitable for boys, believing success can be attributed to ability for boys but to effort for girls (Fennema, Peterson, Carpenter, & Lubinski, 1990; Keller, 2010; see Li, 1999 for a review; Tiedemann, 2002). These stereotypes predict teachers' disparate expectations for male and female students (e.g., Chalabaev, Sarrazin, Trouilloud, & Jussim, 2009) and, in turn, differential treatment (e.g., Hattie, 2009; Jussim & Harber, 2005; Turner & Patrick, 2004). Moreover, although the effects of differential expectations and treatment are typically small (Jussim & Harber, 2005), among students from stigmatized groups, such as girls within physical and engineering science fields, the cumulative negative effects on students' motivation and achievement are more substantial (Green, 2002; Jussim, Eccles, & Madon, 1996).

Today, even though few U.S. adults openly endorse gender stereotypes in STEM (e.g., Schmader, Johns, & Barquissau, 2004; Shumow & Schmidt, 2013), teachers' behavior may nonetheless unintentionally reflect implicit bias and internalized stereotypes (e.g., Milford & Tippett, 2013). In fact, classic and contemporary research have documented discriminatory treatment of girls in science, suggesting that teachers' interaction patterns differ across student genders. Often without explicit awareness, teachers communicate higher expectations of, ask more questions of, and spend more time interacting with male than female students in science classes (e.g., Altermatt, Jovanovic, & Perry, 1998; Becker, 1981; Jones & Wheatley, 1990; Shumow & Schmidt, 2013).

Furthermore, students' own socialization and endorsement of widespread stereotypes about girls' suitability for physical and engineering science fields may negatively skew their perceptions of supportive teacher practices and experiences of need satisfaction in STEM classrooms. Repeated direct and indirect (e.g., mass media) observations of men and women lead children to associate men and women with particular roles and activities (e.g., Eagly, and Wood, 2012; Wood & Eagly, 2012). In addition, gender-disparate ability and value beliefs of parents and teachers (and even peers) not only influence these socializers' behaviors to create opportunities for students to engage with science, but also shape students' own science ability and value beliefs to be consistent with socializers over time (e.g., Eccles-Parsons, Adler, & Kaczala, 1982; Simpkins, Fredricks, & Eccles, 2012; Wang & Degol, 2013). These stereotypes of science as a male domain emerge at a young age (Chambers, 1983; Steffens, Jelenec, & Noack, 2010) and are pervasive across cultures (Nosek et al., 2009). Within this decade, only 35% of American children depicted a woman when asked to a draw a picture of a scientist (Farland-Smith, 2009). Given these pervasive stereotypes, teachers may be less likely to provide autonomy support to girls in physical and engineering science courses, given explicit or implicit expectations that girls will benefit less from such support, and additionally, girls may expect and perceive their teachers to provide greater support to boys compared to girls in such courses and, likewise, experience less need satisfaction. Given these pervasive stereotypes, it is also possible that girls might need more autonomy support from teachers, or even different kinds of support compared to their male peers, in order to experience the same resulting sense of need satisfaction and engagement.

It is important to emphasize that science-related gender stereotypes are particularly pronounced in physical and engineering science. Recent research (e.g., Leslie et al., 2015) suggests that women are underrepresented particularly in fields in which practitioners believe that innate talent is a prerequisite for success, because women are stereotyped as having less innate talent compared to men. In contrast to biological and biomedical areas (as well as fields like education, psychology, and anthropology), math-intensive fields like physics, engineering, and computer science (along with some humanities and arts such philosophy and music composition) are among the fields whose practitioners most strongly endorse such beliefs. These trends suggest that teachers' practices, students' perceptions of teachers' practices, and, in turn, students' subsequent need satisfaction, motivation, and engagement are likely to vary by student gender, particularly in physical science courses, including engineering, rather than in biological science courses.

4 | THE PRESENT INVESTIGATION

In an effort to better understand persistent underrepresentation of women in physical and engineering science fields, the purpose of the current study was to explore gender differences in high school students' daily experiences of teachers' autonomy support and psychological need satisfaction in science class, as well as whether such differences might explain any observed gender differences in daily engagement during science class. Given self-determination theory's assumption that students' subjective experiences of teachers' practice, not the objective reality of teacher practice, ultimately influences their motivation and engagement, we focused on students' perceptions of autonomy-supportive teaching and experience of need satisfaction. We hypothesized that female students would perceive less daily autonomy support from their teachers and experience lower need satisfaction compared to male students in physics, chemistry, and engineering courses. Moreover, we expected girls' weaker experiences of autonomy support and need satisfaction would explain their lower daily engagement relative to boys in such courses.[2] In contrast, and in line with the equal representation of women in biological and biomedical fields and less salient gender stereotypes (e.g., Leslie et al., 2015), we did not expect to observe any gender differences in biological courses.

In order to strengthen confidence in the findings, we explored these hypotheses after controlling for a variety of student and classroom characteristics (e.g., students' ethnicity, free or reduced price lunch eligibility, age, and prior course grade, as well as classroom content difficulty and teacher years of experience). We controlled for these covariates given prior research, which has suggested that these student and classroom factors may influence students' engagement and perceptions of the environment (e.g., Eccles et al., 1993; Murdock, 1999; Solomon, Battistich, & Hom, 1996), particularly within the science domain (e.g., Sinatra et al., 2015). Given some evidence that race, ethnicity, and gender are linked to attrition from STEM fields (for example, see Griffith, 2010, for evidence that racial minorities may have higher rates of attrition due to disparate educational experiences but Riegle-Crumb, and King, 2010, for evidence that disparities in choice of college major may be limited to gender), we thought it was particularly important to control for these student characteristics in our analyses.

Overall, we expected the current study to extend evidence related to engaging girls in science by providing the first test of whether gender differences in students' daily perceptions of teachers' need supportive practices and psychological need satisfaction explain gender differences in daily engagement in science classes. Moreover, this study has the particular strength of distinguishing these relationships by characteristics of subdomain of science. Finally, it used a rigorous intensive longitudinal design to provide accurate and reliable evidence of students' daily experiences in science classrooms.

5 | METHOD

5.1 | Participants

Two-hundred and eight students (54% female; 13 to 18 years of age; 41% eligible for free or reduced lunch) from 41 science classrooms across eight urban and suburban public high schools in the southwestern United States participated (see Table 1 for sample descriptive information). Students responded to a questionnaire related to their classroom experiences after every science class during a six-week instructional unit (2,176 total reports). The study was conducted between January 2013 and May 2014.

Every classroom was led by a different science teacher. Students were approximately evenly split between biology courses and physical science (physics, chemistry, and engineering) courses. The gender composition of biology and physical science courses was similar; slightly over 50% of participants in both biology and physical science courses were female (52% female in biology and 56% female in physical science). Forty-eight percent of students were enrolled in a grade-level biology, physics, or chemistry course, and the rest were enrolled in an advanced biology, physics, or chemistry course or a specialty topic science course (anatomy, environmental systems, engineering, or aquatic science). Forty-two percent of the students across these classes were Hispanic/Latino, 32% were white, 10% were black, 2% were Asian, and 14% were of mixed ethnicities or another ethnicity. Forty-two percent of students were 9th graders, 24% were 10th graders, 17% were 11th graders, and 17% were 12th graders. The mean grade point average (GPA) was 2.92 (SD = 0.96; minimum = 0.82, maximum = 4.0) on a 4-point scale at the start of the study. A comparison of our sample demographics to demographics of the urban and suburban school districts from which students were drawn indicated that our sample was representative of the student population being served in these school districts.

Participation was voluntary. Parent permission was required for students under the age 18. Three to six students in each class participated in the study. In 35 classes, five or six students were randomly selected to participate among volunteers from the class. In the remaining six classes, fewer than five students volunteered. In five of these classes, four students participated and in one class, three students participated. Despite randomly selecting among volunteers in classes in which we were able, given that participation was contingent on volunteering and a limited number of students in each class volunteered, this sample should not be mistaken for a true random sample and should be considered a convenience sample. Students were paid $5 for every complete survey and received a $50 bonus for completing all reports for which they were present in class.

TABLE 1 List of science courses and sample descriptive information

	Classes, *n*	Students *n* (%)	
Grade-level biology	7	36 (17)	
Grade-level physics	6	28 (13)	
Grade-level chemistry	4	21 (10)	
Integrated grade-level physics and chemistry	3	14 (7)	
Pre-AP Biology	8	42 (20)	
Pre-AP chemistry	4	21 (10)	
Pre-AP physics	2	10 (5)	
AP physics	1	5 (2)	
Anatomy	2	11 (5)	
Aquatic science	2	9 (4)	
Environmental systems	1	5 (2)	
Engineering	1	6 (3)	
	Biology (*n* = 103)	**Physics, chemistry, and engineering (*n* = 105)**	**Total (*n* = 208)**
Class characteristics			
Students in advanced class, *n* (%)	48 (46)	43 (42)	91 (44)
Mean teacher years of experience (SD)	9.62 (7.77)	11.31 (11.05)	10.45 (9.53)
Student characteristics			
Mean age	14.96 (1.20)	16.11 (1.05)	15.54 (1.26)
Female, *n* (%)	58 (56)	54 (52)	112 (54)
Latino/Hispanic, black, or other, *n* (%)	60 (58)	71 (68)	131 (63)
Free/reduced lunch, *n* (%)	32 (31)	54 (51)	86 (41)
Prior unit course grade, M (SD)	82.84 (18.40)	81.59 (17.86)	82.21 (18.10)
Autonomy support, M (SD)	2.97 (0.65)	2.93 (0.73)	2.54 (0.66)
Need satisfaction, M (SD)	2.51 (0.61)	2.57 (0.71)	2.95 (0.69)
Engagement, M (SD)	2.96 (0.63)	2.86 (0.73)	2.91 (0.68)

Note. M = mean. SD = Standard deviation.

Teachers were 25 to 66 years of age (M = 38.12; SD = 12.49) and had between 0 and 40 years of experience (M = 10.40, SD = 9.85). Most teachers were white (30) and female (30), with more male teachers teaching physical science courses.[3] One teacher was black, three were Asian, three were Hispanic/Latino, and four were of mixed ethnicities or another ethnicity. Teachers received $50 for participation. Schools received $100 for every teacher who participated in the study.

5.2 | Procedure

The recruitment process included multiple steps. First permission was obtained from the two school districts, as well as individual high school principals, vice principals, and science chairs at each of the eight schools. Next, group information sessions were held to recruit teachers. Teachers were informed that the purpose of the study was to examine the relationship between students' experiences in the classroom and their motivation and engagement. The methods involved in the study were also explained to teachers. Teachers were encouraged to view participation in the study as an educational experience, since they would be provided information about the relationships between teacher practices and students' motivation and engagement at the end of the study and all the information collected as part of the study was confidential. With that in mind, the research team encouraged teachers to select their most typical course for participation that suited the study best for scheduling reasons and contained a diverse group of students. The research team discouraged teachers from selecting a course because they felt it was the one in which they or their students would perform best (or worst). Across all schools, approximately 50% of recruited science teachers expressed willingness to participate and approximately 40% actually participated in the study.

Student participants were recruited via in-person classroom visits in which the study was described. During these visits, a parent information letter and consent documents in both English and Spanish were distributed. A box was set up in the main office of each school so students could return signed consent documents in a sealed envelope without teachers knowing which students opted to participate.

Each participant met individually with a member of the research team for a study orientation. During this meeting, participating students received an Apple iPod touch for completing surveys for the duration of the study. Student participants also practiced using the iPod during this meeting and completed a background survey regarding their age, grade level, sex, ethnicity, eligibility for free or reduced lunch at school based on U.S. government policy, school GPA, and science course grade for the most recent instructional unit.

On every class day of the six-week instructional unit, students were emailed during their first non-instructional free period following the science class session with a survey about their teachers' practices, their experiences of need satisfaction, and their engagement in class. All surveys were programmed using Qualtrics and completed by students online. All classes met approximately every other school day. The number of report opportunities varied across courses and six-week instructional units. The number of scheduled class sessions ranged between 11 and 17, with classes having between 8 and 17 opportunities to report on experiences as a result of various disruptions to class sessions (median = 14). Students had until the next class day to complete their daily report. Student participants completed between 1 and 17 reports across the instructional unit (mean (SD) = 10.46 (3.77); mode = 10), with girls and boys completing a similar number of reports (girls mean(SD) = 10.64 (3.56); boys mean(SD) = 10.14 (4.02)). Only one student completed just one report, and thus this student's responses were excluded because change in outcomes was the focus of our analyses. Figure 1 provides an illustration of the procedures for this study.

5.3 | Measures

5.3.1 | Engagement

Engagement is conceptualized as a multidimensional construct, including four related but distinct behavioral, emotional, cognitive, and agentic components (e.g., Fredricks et al., 2004; Reeve, 2013; Skinner et al., 2009). In line with this multidimensional conceptualization, students' daily engagement in science class was assessed with 15 items we adapted from the Engagement versus Disaffection with Learning Student Report (Furrer & Skinner, 2003; Skinner & Belmont, 1993; Skinner et al., 2009), the Metacognitive Strategies Questionnaire (Wolters, 2004) and the Agentic Engagement Scale (Reeve & Tseng, 2011). The Engagement versus Disaffection with Learning Student Report contains scales for behavioral and emotional engagement from which we adapted seven items (three behavioral and four emotional engagement) for the daily context. We used items from these particular scales because prior research has found that scores on these scales are associated with important student outcomes including achievement, and they are regularly used in school-based self-determination theory research to assess behavioral and emotional engagement (e.g., Reeve, 2013; Skinner et al., 2009). Four items measuring learning strategies adapted from the Metacognitive Strategies Questionnaire were used to assess students' cognitive engagement in science class. We used items from this particular scale because its conceptualization of cognitive engagement focused on the use of cognitive learning strategies (distinguishing it from behavioral engagement), it has been linked with students' achievement (e.g., Wolters, 2004), and it has been used in prior school-based self-determination theory research (e.g., Reeve, 2013). Students' agentic engagement in science class was assessed with four items adapted from the Agentic Engagement Scale (Reeve & Tseng, 2011), currently the only scale for assessing agentic engagement. For all engagement items, students rated the extent to which

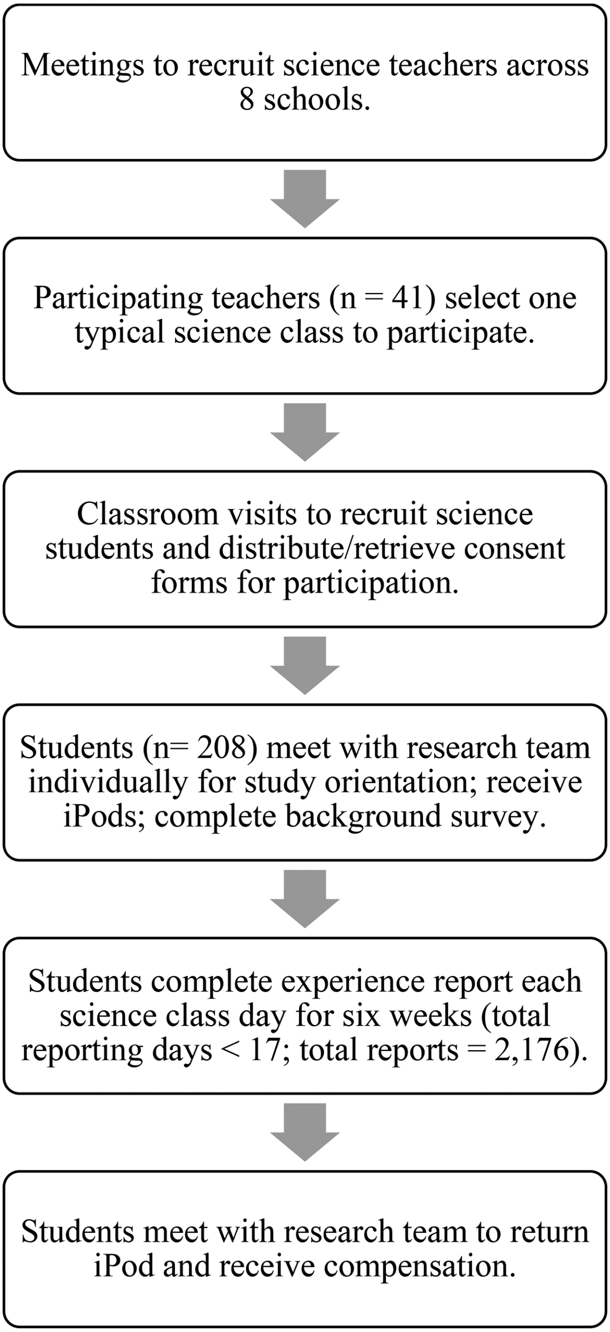

FIGURE 1 Illustration of study procedures

they agreed with each item on a 5-point Likert scale ranging from not at all true (1) to extremely true (5). The validity and reliability of all engagement scales for cross-sectional research have been established in previous studies (Furrer & Skinner, 2003; Reeve, 2013; Reeve & Tseng, 2011; Wolters, 2004). Given that we adapted and shortened the measures in order to use them in an intensive longitudinal design, we conducted factor and reliability analyses to confirm that these adapted measures were appropriate for the daily context.

To assess the factorial validity of daily measures of engagement, we conducted a multilevel confirmatory factor analysis (ML-CFA) with six factors (includes two additional disengagement scales not the focus of this investigation) at both the day and student levels in Mplus 6.12. Parameters were estimated using a maximum likelihood estimation procedure (i.e., MLR) that is robust to violations of both the assumptions of normality and independence of observations and provides for optimal parameter estimates when data are missing at random. We examined both day- and student-level (by computing the mean across class days for each student) factor structures. Given the complexity of modeling a three-level exploratory factor structure with the small number of classrooms, we used the TYPE = COMPLEX TWO LEVEL command in Mplus to adjust standard errors and chi-square tests of model fit, accounting for the clustering at the classroom level (level 3). To obtain proper estimates at each level, we followed standard multilevel modeling practices and used group-mean centering for items at both the day and student levels using the student as the group for the lowest level and the class as the group for the student level. A well-fitting model was defined by a comparative fit index (CFI) of approximately .95, a root mean square of approximation (RMSEA) less than .05, a standardized root mean square residual (SRMR) less than .08, and factor loadings >.40 (Kline, 2015). Items were allowed to load only on their target factor (i.e., behavioral, emotional, cognitive, or agentic) and factors were allowed to correlate. Inspection of fit indices for the model (CFI = .92, RMSEA = .03, and SRMR <.04 both for day and student levels) indicated that the model fit the data adequately (Kline, 2015). Factor loadings (i.e., standardized regression coefficients) at both levels suggested that items loaded sufficiently (>.40) onto their respective factors. Items and the mean daily reliabilities of each subscale are reported in Table 2.

Given that our hypotheses did not differ by type of engagement, we created a composite engagement variable by averaging daily behavioral, emotional, cognitive, and agentic engagement scales (mean daily $\alpha = .91$). Sizeable correlations between factors supported this approach: the correlation between behavioral, emotional, cognitive, and agentic engagement factors ranged between .42 and .66. Moreover, using an aggregated measure of engagement was an appealing approach given that it limited the number of statistical tests conducted and yielded excellent reliability.

5.3.2 | Need satisfaction

Students' daily need satisfaction in science class was assessed with 12 items adapted for use in our daily intensive longitudinal design from the Perceived Self-Determination Scale (Reeve, Nix, & Hamm, 2003), the Perceived Competence Scale (Williams & Deci, 1996), and perceived relatedness subscale of the Activity-Feeling States (Reeve & Sickenius, 1994). Autonomy need satisfaction was measured with six of nine items from the Perceived Self-Determination Scale. Reversed items were dropped because they typically load most weakly in factor analyses. Competence need satisfaction was assessed with three of four items from the Perceived Competence Scale. Relatedness need satisfaction was assessed with the three items from the Activity-Feeling States for the relatedness subscale. Students rated the extent to which they agreed with each item on a 5-point Likert scale ranging from not at all true (1) to extremely true (5). We selected items from these particular scales to assess need satisfaction on the three needs because they are relatively brief, and are regularly used in school-based self-determination theory research (e.g., Patall & Leach, 2015; Patall et al., 2013; Reeve & Lee, 2014), and because previous studies have established the validity and reliability of all scales for cross-sectional research (Reeve & Jang, 2006; Reeve & Sickenius, 1994; Reeve et al., 2003). We conducted multilevel factor analyses and reliability analyses to confirm that our abbreviated measures were appropriate for our daily context.

To examine the structure of the daily need satisfaction measure, we conducted a multilevel confirmatory factor analysis (ML-CFA) using MLR to examine a three-factor structure at both day and student levels using the TYPE = COMPLEX TWO LEVEL syntax in Mplus to account for clustering at the classroom level. Items were group-mean centered for both the day and student levels using the student as the group for the lowest level and the class as the group for the student level. Items were allowed to load only on their target factor and factors were allowed to correlate. Inspection of model fit indices (CFI = .96 RMSEA = .03; and SRMR = .03 for the day level and .09 for the student level) indicated that the model fit the data well (Kline, 2015). Factor loadings at both levels suggested that items loaded sufficiently (> .44) onto their respective factors. Items and the mean daily reliabilities of each subscale are reported in Table 2.

TABLE 2 Daily report items and reliabilities

Construct (mean daily Cronbach's alpha)	Subconstruct (mean daily Cronbach's alpha)	Items
Daily perceived autonomy support (.91)	Provision of choice (.83)	My teacher allowed me to choose which questions or parts of an assignment to work on today.
		My teacher provided options for the kinds of assignments or activities I could do today.
		My teacher allowed me to choose how to do my work in the classroom today.
		My teacher allowed me to choose how to use my time for studying and classwork today.
		My teacher encouraged me to work in my own way today.
	Consideration for student interests and preferences (.87)	My teacher structured class activities today around my interests.
		My teacher took my preferences into consideration for assignments today.
		My teacher worked my interests into his/her lesson(s) today.
	Rationales identifying usefulness, importance, and relevance of activities (.86)	My teacher explained how what we were learning today is important.
		My teacher demonstrated how what we were learning today is useful.
		My teacher explained how the course assignments today were important.
		My teacher talked about the connection between what we are studying in school today and real life.
	Consideration for negative affect (.79)	My teacher was open to hearing criticism or complaints about activities today.
		My teacher was accepting of any negative feelings about course material today.
	Informational/ encouraging feedback (.73)	My teacher provided encouragement when I struggled with the course material today.
		My teacher gave suggestions when I struggled with course work today.
		My teacher told me when I made progress in learning course material today.
Daily psychological need satisfaction (.90)	Autonomy (.87)	I felt I did what I wanted to be doing in my science class today.
		I felt I pursued goals that were my own in science class today.
		I felt a sense of personal freedom in science class today.
		In science class today, I felt free.
		I believe I generally had choices in science class today.
		I feel like learning was my own choice in science class today.
	Competence (.82)	I felt competent while working on assignments for my science class today.
		I felt confident about my ability to learn the material during science class today.
		I felt able to perform well in science class today.

(Continues)

TABLE 2 (Continued)

Construct (mean daily Cronbach's alpha)	Subconstruct (mean daily Cronbach's alpha)	Items
	Relatedness (.88)	Today in science class, I felt I belonged and the people in class care about
		Today in science class, I felt involved with close friends
		Today in science class, I felt emotionally close to the people around me.
Daily engagement (.91)	Behavioral (.68)	I worked as hard as I can in science class today.
		I participated today in science class discussions.
		I paid attention today in science class.
	Emotional (.89)	When I was in science class today, I felt good.
		I felt interested today in science class.
		I enjoyed science class today.
		I felt happy today in science class.
	Cognitive (.86)	I tried to connect what I was learning in science class today with my own experiences.
		I tried to make different ideas fit together and make sense in science class today.
		When doing work for science class today, I tried to relate what I'm learning to what I already know.
		I made up my own examples to help me understand the important concepts in science class today.
	Agentic (.80)	I let my science teacher know what I needed and wanted today.
		During science class today, I expressed my preferences and opinions.
		During science class today, I asked questions to help me learn.
		I let my science teacher know what I am interested in today.

Given that our hypotheses did not differ across the different components of psychological need satisfaction, we created a composite need satisfaction variable by averaging daily perceived autonomy, perceived competence, and perceived relatedness scales (mean daily $\alpha = .90$). Correlations between factors supported this approach, ranging between .47 and .62. Moreover, using an aggregated measure of need satisfaction limited the number of statistical tests conducted and yielded excellent reliability characteristics. Finally, this approach is consistent with prior research in which needs are aggregated in a single measure (e.g., Jang, Kim, & Reeve, 2016).

5.3.3 | Perceived autonomy-supportive teacher practices

Students' perceptions of the extent to which their teachers used practices theorized to support autonomy on a given class day were assessed with a measure designed explicitly for use in this study (see Table 2 for the final set of items). We attempted to capture a broad set of practices that have been conceptualized as autonomy-supportive, drawing on a variety of prior measures used in cross-sectional research (Patall et al., 2013; as well as Assor et al., 2002; Assor, Kaplan, Kanat-Maymon, & Roth, 2005; Belmont, Skinner, Wellborn, & Connell, 1992; Connell, 1990; Katz, Kaplan, & Gueta, 2010; Reeve & Jang, 2006; Wellborn & Connell, 1987). Twenty items assessed perceptions of six supportive daily practices hypothesized to be related to autonomy need satisfaction and motivation based on prior research. Supportive practices included (a) provision of choices, (b) opportunities for students to work in their own way, (c) consideration for student opinions, preferences, and interests, (d) rationales regarding the usefulness and importance of course material, (e) consideration for students' negative affect, and (f) encouraging, informational feedback. Students rated the extent to which they agreed with each item on a 5-point Likert scale ranging from not at all true (1) to extremely true (5).

To assess the factorial validity of daily measures of perceived autonomy-supportive teacher practices, we conducted multilevel exploratory factor analyses (Roesch et al., 2010) using the oblique geomin rotation and MLR in Mplus 6.12 to examine both day- and student-level factor structures. A range of factors at each level of the nested data structure (from 1 to 7 factors) were tested. Again, we used the TYPE = COMPLEX TWO LEVEL command in Mplus to account for the clustering at the classroom-level and group-mean-centered items for both the day and student levels using the student as the group for the lowest level and the class as the group for the student level. We used a change of .01 or greater in the CFI as our model selection criterion when attempting to identify the best fitting model (Cheung & Rensvold, 2002).

The results of ML-EFAs of these 20 items (plus three additional items not included in this investigation focused on question opportunities) supported a six-factor structure (CFI = .98, RMSEA = .018, SRMR (day/student) = .007/.009). All items loaded sufficiently (>.40) on the intended factor as expected with minimal cross-loadings, with the caveat that perceptions of provision of choice items and opportunities for students to work in their own way items loading on a single factor rather than two separate factors. Several items were retained only at the student level. These included one item assessing the provision of choice and two items assessing consideration for student interests and preferences.

Scale scores for each perceived teacher practice were calculated by taking the mean of all items loading above .40 on the factor. When factor analyses suggested that a slightly different version of a scale should be used at day versus student levels, we computed multiple versions of the scale to be used at the appropriate level. However, for the purposes of this investigation, we used only day-level scales (final items and mean daily reliabilities of each subscale are reported in Table 2), though results were nearly identical using either version of the scales. In addition, given that our hypotheses did not differ across the various perceived autonomy-supportive practices and our prior research indicated that all scales were significant predictors of students' experience of autonomy (Patall et al., 2017), we created a composite perceived autonomy support variable by averaging daily perceptions across the five subscales (mean daily $\alpha = .91$). Again, using aggregated measures across types of perceived autonomy-supportive practices was an appealing approach given that it limited the number of statistical tests conducted and yielded excellent reliability. Finally, this approach is consistent with a great deal of other research in which various practices are aggregated in a single measure (e.g., Jang et al., 2012; Katz et al., 2010).

5.4 | Multilevel analyses

We tested our main hypotheses about the relationships among gender, daily perceptions of teacher autonomy-supportive practices, daily experiences of need satisfaction, and daily engagement with a series of three-level (day, student, and class) regressions where the intercept was allowed to vary randomly using the Mixed procedure in SPSS 21. In line with recommendations from experts on conducting intensive longitudinal data analysis (e.g., Bolger & Laurenceau, 2013), we used hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) for our primary tests because it appropriately addressed non-independence of observations and the hierarchically nested design of our data set in which lower-level units (i.e., days) were nested within a second higher-level unit (i.e., students), and students were nested within a third higher-level unit (i.e., classrooms). HLM supports generalizations of the findings to a wider population by treating students and classrooms as a random rather than a fixed effect.

For all multilevel models, at level 1 (day level) we included the class day (time) and the outcome reported on the previous day. Some models also included daily perceived autonomy support or daily need satisfaction as predictors when the purpose was to examine the extent to which daily autonomy support or need satisfaction explained the link between gender and engagement. We constructed the class day (time) variable by consecutively numbering each class session during the unit starting with zero. We opted to use class session as the time metric, as opposed to calendar days or school days elapsed, given Kim-Spoon and Grimm's (2016) recommendation to consider the dominant reasons for why changes in the outcome might occur when selecting a time metric. In our investigation, the dominant reason student perceptions, need satisfaction, and engagement in science class were expected to vary is because of their experiences during science class sessions. The prior class session's outcome was entered to control for possible carryover effects from one class day to the next (e.g., see Sheldon, Gable, Roscoe, & Ryan, 2000, for an example of this strategy). To minimize missing data, reports were carried forward to the next available day of reporting when creating lagged variables. Including the prior class session's outcome value as a predictor allowed us to predict day-to-day change in the outcome rather than sheer level (Cohen & Cohen, 1983) in models examining the predictive power of students' perceptions of teacher autonomy support or need satisfaction reported on the same class day as the engagement outcome.

At level 2 (student level), we included student gender as a focal variable (0 = male, 1 = female), as well as student ethnicity (0 = white or Asian, 1 = black, Hispanic/Latino, or other ethnic minority), students' free or reduced price lunch eligibility (0 = not eligible, 1 = eligible), students' age, and students' course grade for the prior unit in all models. At level 3 (class level), we included variables representing whether the class was advanced or grade typical (0 = grade typical, 1 = advanced) and teacher years of experience in all models. Some models also included the mean of perceived autonomy support or need satisfaction across days for each student in order to examine the extent to which cumulative autonomy support or need satisfaction explained the link between gender and engagement.

To decompose within-student (day) effects from between-student and between-class effects, daily perceived autonomy support and need satisfaction were student-mean-centered (around each student's own average score) and student-level (between-student) autonomy support and need satisfaction predictors were class-mean centered (around the average score for the class). Likewise, student gender, as well as ethnicity, age, free/reduced lunch eligibility, and prior course grade were class-mean centered (around the average score for the class).[4] Class day (time) and the value of the outcome variable from the prior class session were grand-mean centered since they were simply control variables in these models, as were the class-level covariates. To treat missing data, we used a maximum likelihood estimation procedure with robust estimates of standard errors (REML). We also specified an AR(1) correlated error structure since adjacent residuals may be correlated across measurements in repeated measures data (Bolger & Laurenceau, 2013).

6 | RESULTS

6.1 | Preliminary analyses

First, we computed correlations among the main study variables, aggregating daily variables across days in the unit. We computed correlations separately for biology courses and physics, chemistry, and engineering courses. As expected, significant negative correlations were found between gender and perceived autonomy support, need satisfaction, and engagement in physics, chemistry, and engineering courses (r ranged between −.25 and −.38), though corresponding correlations were not statistically significant for biology courses. In physics, chemistry, and engineering courses, teachers' years of experience was negatively correlated with both perceived autonomy support and need satisfaction. In these physical science courses, students' eligibility for free/reduced lunch was positively correlated with perceived autonomy support. In biology courses, students' eligibility for free/reduced lunch was negatively correlated and prior course grade was positively correlated with need satisfaction. Also, as expected, there were strong correlations among perceived autonomy support, need satisfaction, and engagement, regardless of domain (r ranged between .70 and .87). Table 3 provides a correlation matrix for all study variables.

6.2 | Gender differences in students' daily engagement, perceived autonomy support, and need satisfaction

First, the extent to which students' gender predicted their daily engagement was examined with two random intercept three-level (day, student, and class) regressions, one model for physics, chemistry, and engineering courses and one model for biology courses, that included gender, and all other covariates as previously described. Results (see Table 4)

TABLE 3 Correlations among student demographic variables and aggregated daily variables

	Advanced class	Teacher experience	Ethnicity	Age	Free lunch	Prior grade	Gender	Autonomy support	Need satisfaction	Engagement
Advanced class	–	–0.28**	0.15	–0.13	0.03	0.07	0.15	0.01	0.01	0.01
Teacher experience	–0.01	–	–0.16	0.39***	–0.03	0.05	–0.06	0.08	–0.05	–0.01
Ethnicity	–0.13	–0.30**	–	0.04	0.37***	–0.11	–0.07	0.07	0.04	0.11
Age	–0.07	–0.03	0.10	–	–0.23*	0.23*	0.05	0.17	0.18	0.18
Free Lunch	–0.21*	–0.40***	0.55***	0.05	–	–0.23*	0.03	–0.06	–0.24*	–0.13
Prior grade	0.05	–0.08	–0.20*	–0.06	–0.13	–	–0.07	0.11	0.22*	0.19
Gender	–0.24*	0.06	0.16	–0.14	–0.02	–0.04	–	0.02	–0.07	–0.05
Autonomy support	0.003	–0.27**	0.15	0.04	0.31***	–0.02	–0.25*	–	0.70***	0.77***
Need Satisfaction	0.12	–0.22*	0.01	0.12	0.12	0.07	–0.38***	0.75***	–	0.84***
Engagement	0.11	–0.17	0.06	0.08	0.16	0.07	–0.29**	0.80***	0.87***	–

Note. Sample size (n) = 96 to 105 students for physics, chemistry, and engineering courses and 97 to 103 students for biology courses. Correlations below the diagonal are for physics, chemistry, and engineering courses. Correlations above the diagonal are for biology courses. Perceived teacher autonomy support, need satisfaction, and engagement variables were aggregated across class sessions for individual students. For student sex, 0 = male and 1 = female. For ethnicity, 0 = white or Asian and 1 = black, Hispanic/Latino, or other ethnic minority. For free lunch, 0 = not eligible for free/reduced price lunch and 1 = eligible for free/reduced price lunch. For class type, 0 = grade typical class and 1 = advanced class. Students' age and prior course grade were measured continuously. Teacher experience was measured continuously as the number of years teachers had been professionally teaching. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

TABLE 4 Multilevel regressions for students in physics, chemistry, and engineering courses

	Perceived autonomy support		Need satisfaction		Engagement	
Fixed effects	***b*(SE)**	***β***	***b*(SE)**	***β***	***b*(SE)**	***β***
Class level						
Intercept	2.54 (0.05)		2.92 (0.05)		2.85 (0.05)	
Advanced class	0.01 (0.10)	0.01	0.14 (0.11)	0.08	0.13 (0.10)	0.07
Teacher experience	−0.01 (0.004)	−0.10	−0.01 (0.005)	−0.10	−0.01 (0.004)	−0.07
Student level						
Age	−0.04 (0.06)	−0.04	−0.03 (0.07)	−0.02	−0.03 (0.07)	−0.03
Free/reduced lunch	0.22 (0.11)	0.10	−0.02 (0.13)	−0.01	0.02 (0.13)	0.01
Ethnicity	−0.04 (0.11)	−0.02	−0.11 (0.13)	−0.05	−0.03 (0.13)	−0.01
Prior unit course grade	−0.002 (0.003)	−0.03	−0.002 (0.003)	−0.03	−0.0001 (0.003)	−0.002
Gender	**−0.27 (0.09)**	**−0.13****	**−0.35 (0.10)**	**−0.16****	**−0.25 (0.10)**	**−0.12***
Day Level						
Class day (time)	0.007 (0.003)	0.04*	−0.003 (0.004)	−0.01	−0.003 (0.003)	−0.01
Lagged outcome	0.44 (0.03)	0.44***	0.37 (0.03)	0.37***	0.42 (0.03)	0.42***
Random effects	**Variance**	**SE**	**Variance**	**SE**	**Variance**	**SE**
Class (level 3) intercept	0.02	0.02	0.02	0.02	0.01	0.02
Student (level 2) intercept	0.11**	0.03	0.15***	0.05	0.16***	0.04
Day (level 1)						
Residual	0.22***	0.01	0.27***	0.01	0.24***	0.01
Autocorrelation	−0.19**	0.06	−0.13*	0.07	−0.16**	0.06

Note. Level 1 (daily reports) $n = 939$ to 982 reports. Level 2 (students) $n = 100$. Level 3 (classes) $n = 21$. The class day ("time") variable reflects the day of reporting across the 6-week instructional unit. The "lagged outcome" variable reflects the prior class session's value for the outcome. For student sex, 0 = male and 1 = female. For student ethnicity, 0 = white or Asian and 1 = black, Hispanic/Latino, or other ethnic minority. For free and reduced lunch status, 0 = not eligible for free/reduced lunch and 1 = eligible for free/reduced lunch. For advanced class, 0 = grade typical class and 1 = advanced class. Students' age and prior course grade were measured continuously. Teacher experience was measured continuously as the number of years teachers had been professionally teaching. b = unstandardized regression coefficient. β = standardized regression coefficient. Standardized estimates were computed using the following formula (Hox, Moerbeek, & van de Schoot, 2010): $\beta = (b \times sdx)/sdy$. SE = standard error. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

suggested that girls' daily engagement in physics, chemistry, and engineering classes was lower than that of boys, controlling for student and classroom characteristics, the lagged outcome, and class day (time). In contrast, there was no gender difference in engagement in biology courses (see Table 5). In both models, the only other variable that significantly predicted students' daily engagement was their level of engagement on the prior class day.

Next, the extent to which students' gender predicted their daily perceptions that teachers engaged in practices supportive of their autonomy was examined with two random intercept three-level regressions with the same set of predictors, one regression for physics, chemistry, and engineering courses and a separate regression for biology courses. As expected, results (see Table 4) suggested that girls perceived less autonomy support from their teachers in physics, chemistry, and engineering classes compared to boys, controlling for student and classroom characteristics, the lagged outcome, and class day (time). Again, there was no gender difference in perceived autonomy support in biology courses (see Table 5). The perceived autonomy support on the prior class day positively predicted students' daily perceptions of autonomy support in each model. Class day also positively predicted daily perceptions of autonomy support in physics, chemistry, and engineering courses, suggesting that students' perceived autonomy support increased over the six weeks.

TABLE 5 Multilevel regressions for students in biology courses

Fixed effects	Perceived autonomy support b(SE)	β	Need satisfaction b(SE)	β	Engagement b(SE)	β
Class level						
Intercept	2.52 (0.07)		3.00 (0.05)		2.97 (0.05)	
Advanced class	0.06 (0.14)	0.04	−0.01 (0.11)	−0.01	0.001 (0.11)	0.001
Teacher experience	0.01 (0.01)	0.07	−0.003 (0.01)	−0.03	0.0002 (0.007)	0.002
Student level						
Age	−0.04 (0.08)	−0.03	0.03 (0.07)	0.02	0.005 (0.07)	0.004
Free/reduced lunch	0.10 (0.12)	0.05	−0.09 (0.10)	−0.05	0.01 (0.10)	0.01
Ethnicity	−0.01 (0.10)	−0.01	0.05 (0.09)	0.03	0.06 (0.09)	0.04
Prior unit course grade	−0.002 (0.003)	−0.03	0.002 (0.002)	0.03	0.001 (0.002)	0.03
Gender	**0.04 (0.10)**	**0.02**	**−0.01 (0.08)**	**−0.003**	**−0.03 (0.08)**	**−0.02**
Day level						
Class day (time)	0.001 (0.003)	0.004	−0.001 (0.003)	−0.004	−0.002 (0.003)	−0.01
Lagged outcome	0.35 (0.03)	0.35***	0.46 (0.03)	0.46***	0.47 (0.03)	0.47***
Random effects	**Variance**	**SE**	**Variance**	**SE**	**Variance**	**SE**
Class (level 3) intercept	0.05	0.03	0.03	0.02	0.02	0.02
Student (level 2) intercept	0.12***	0.04	0.09*	0.04	0.08**	0.03
Day (level 1)						
Residual	0.17***	0.01	0.22***	0.02	0.18***	0.01
Autocorrelation	−0.22***	0.06	−0.29***	0.07	−0.21**	0.07

Note. Level 1 (daily reports) $n = 773$ to 840 reports. Level 2 (students) $n = 91$. Level 3 (classes) $n = 20$. The class day ("time") variable reflects the day of reporting across the 6-week instructional unit. The "lagged outcome" variable reflects the prior class session's value for the outcome. For student sex, 0 = male and 1 = female. For student ethnicity, 0 = white or Asian and 1 = black, Hispanic/Latino, or other ethnic minority. For free and reduced lunch status, 0 = not eligible for free/reduced lunch and 1 = eligible for free/reduced lunch. For advanced class, 0 = grade typical class and 1 = advanced class. Students' age and prior course grade were measured continuously. Teacher experience was measured continuously as the number of years teachers had been professionally teaching. b = unstandardized regression coefficient. β = standardized regression coefficient. Standardized estimates were computed using the following formula (Hox et al., 2010): $\beta = (b \times sdx)/sdy$. SE = standard error. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Third, the extent to which students' gender predicted their daily need satisfaction was examined with two identical random intercept three-level regressions. Results (see Table 4) suggested that girls' daily need satisfaction in physics, chemistry, and engineering classes was lower than that of boys, controlling for student and classroom characteristics, the lagged outcome, and class day (time). However, there was no gender difference in need satisfaction in biology courses (see Table 5). In both models, the only other variable that significantly predicted students' daily need satisfaction was their level of need satisfaction on the prior class day.

6.3 | Mediational analyses

Given that we found gender differences in physics, chemistry, and engineering classes, but not biology courses, we examined the extent to which students' daily and cumulative perceptions of their science teachers' autonomy-supportive practices and need satisfaction explained gender differences in daily engagement in physics, chemistry, and engineering courses (see Table 6). To examine the extent to which daily perceived autonomy support and need satisfaction mediated the association between gender and daily engagement within physics, chemistry, and engineering courses, we estimated two three-level models similar to those described above for engagement (one for perceived

TABLE 6 Multilevel regressions to test mediation of relationship between gender and engagement for students in physics, chemistry, and engineering courses

	Model 1		Model 2	
Fixed effects	*b*(SE)	β	*b*(SE)	β
Class level				
Intercept	2.87 (0.08)		2.89 (0.09)	
Advanced class	0.13 (0.16)	0.07	0.19 (0.18)	0.10
Teacher experience	−0.01 (0.01)	−0.10	−0.01 (0.01)	−0.10
Student level				
Age	0.01 (0.06)	0.01	0.01 (0.05)	0.01
Free/reduced lunch	−0.23 (0.12)	−0.10	−0.11 (0.10)	−0.05
Ethnicity	0.06 (0.12)	0.03	0.11 (0.10)	0.05
Prior unit course grade	0.002 (0.003)	0.04	0.003 (0.002)	0.05
Gender	−0.03 (0.10)	−0.02	0.09 (0.08)	0.05
Autonomy support$_{mean}$	**0.72 (0.08)**	**0.46*****	–	–
Need satisfaction$_{mean}$	–	–	**0.91 (0.06)**	**0.63*****
Day Level				
Autonomy support	**0.48 (0.03)**	**0.25*****	–	–
Need satisfaction	–	–	**0.47 (0.03)**	**0.35*****
Class day (time)	−0.01 (0.003)	−0.04*	−0.01 (0.003)	−0.04
Lagged outcome	0.18 (0.03)	0.18***	0.15 (0.03)	0.15***
Random effects	**Variance**	**SE**	**Variance**	**SE**
Class (level 3) intercept	0.09*	0.04	0.13*	0.06
Student (level 2) intercept	0.12***	0.03	0.08***	0.02
Day (level 1)				
Residual	0.18***	0.01	0.18***	0.01
Autocorrelation	0.04	0.07	0.07	0.08

Note. Level 1 (daily reports) $n = 951$ reports. Level 2 (students) $n = 100$. Level 3 (classes) $n = 21$. The class day ("time") variable reflects the day of reporting across the 6-week instructional unit. The "lagged outcome" variable reflects the prior class session's value for the outcome. For student sex, 0 = male and 1 = female. For student ethnicity, 0 = white or Asian and 1 = black, Hispanic/Latino, or other ethnic minority. For free and reduced lunch status, 0 = not eligible for free/reduced lunch and 1 = eligible for free/reduced lunch. For advanced class, 0 = grade typical class and 1 = advanced class. Students' age and prior course grade were measured continuously. Teacher experience was measured continuously as the number of years teachers had been professionally teaching. *b* = unstandardized regression coefficient. β = standardized regression coefficient. Standardized estimates were computed using the following formula (Hox et al., 2010): $\beta = (b \times sdx)/sdy$. SE = standard error. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

autonomy support and one for need satisfaction as the mediator). However, in these models we added students' daily and aggregate (mean across all days in the six-week unit) perceived autonomy support or need satisfaction as predictors (see Table 4). Then, we used the Monte Carlo method for assessing mediation (MCMAM; Selig & Preacher, 2008) using 20,000 replications. MCMAM allowed us to specify the aforementioned level-1 residual structure in our three-level multilevel models and simultaneously estimate and test indirect effects at the day and student levels. Statistical methods traditionally used to test for mediation assume that an indirect effect is normally distributed, which it generally is not. The MCMAM method allows for a test of mediation without assuming a particular distributional shape for the indirect effect and yields more accurate inferences.

As expected, in physics, chemistry, and engineering courses, perceptions of teachers' daily autonomy support predicted an increase in daily engagement over above the prior class day and controlling for student and classroom characteristics and class day (time). Likewise, greater cumulative perceptions of teachers' autonomy support predicted greater engagement in physics, chemistry, and engineering courses. More importantly, tests of the indirect effects suggested that girls' lower daily and cumulative experiences of autonomy support from teachers explained the discrepancy in daily engagement between girls and boys in physics, chemistry, and engineering courses (daily-level indirect path $b = -.13$, 95% confidence interval [CI −.21, −.05] and student-level indirect path $b = -.19$ 95% CI [−.33, −.07]).

Likewise, using the same strategy in a separate model with need satisfaction as the mediator, students' need satisfaction predicted an increase in daily engagement over and above the prior class day and controlling for student and classroom characteristics and class day (time). As expected, cumulative need satisfaction also predicted greater engagement. Again, tests of the indirect effects suggested that girls' lower daily and cumulative need satisfaction explained the discrepancy in engagement between girls and boys in physics, chemistry, and engineering courses (daily-level indirect path $b = -.16$, 95% CI [−.26, −.07] and student-level indirect path $b = -.32$ 95% CI [−.51, −.14]).

7 | DISCUSSION

The present investigation examined whether girls experienced lower daily engagement during their high school science classes, and the extent to which gender differences in daily engagement could be explained by differences in students' daily perceptions of autonomy support and need satisfaction during class. We used a diary method to track students' daily perceptions of teacher practices and experiences during science class over a six-week instructional unit.

The findings were consistent with our hypotheses and fit with predictions derived from self-determination theory and the research on gender stereotypes and gender socialization. We found that in physics, chemistry, and engineering courses where gender stereotypes disadvantaging women are particularly prominent, girls perceived their teachers to provide about 1/8 of a standard deviation less daily support for their autonomy and experienced about 1/6 of a standard deviation less daily psychological need satisfaction in class than boys. While the size of these effects might seem rather modest, it is important to note that daily effects are likely to accumulate over time to create meaningful differences. In turn, these differences, both on a daily basis and cumulatively across an instructional unit, explained girls' lower daily engagement in these courses. For every standard deviation decrease in daily autonomy support or need satisfaction, there was 1/4 to 1/3 of a standard deviation decrease in daily engagement. At the student level, for every standard deviation decrease in students' aggregated experience of autonomy support or need satisfaction across the six weeks, there was about 1/2 to 2/3 of a standard deviation decrease in daily engagement.

In line with our predictions and current trends suggesting that women are underrepresented in physical and engineering science fields, but well represented in biological and biomedical fields, we did not find gender differences in biology courses. Notably, little else predicted students' daily engagement, though we included a variety of covariates in our models. In particular, we did not observe any differences in daily engagement, perceptions of autonomy support, or need satisfaction in high school physics, chemistry, and engineering courses among black and Hispanic/Latino students compared to white or Asian students, which is consistent with at least some evidence suggesting that STEM disparities upon entering college are mostly centered on gender rather than race/ethnicity (e.g., Riegle-Crumb & King, 2010).

The findings of this study lead us to speculate that girls' underrepresentation in physical and engineering science college majors and careers may be partly explained by routinely lower levels of engagement during relevant high school classes that emerge as a function of experiencing less autonomy support from teachers and lower need satisfaction compared to boys. As this study did not examine students' intentions to pursue STEM majors and careers, it falls on future research to examine the extent to which differences in students' in-class engagement experiences during high school predict gender differences in persistence in STEM fields in college and beyond. The existing research and theory certainly implicate high school engagement in gender discrepancies in STEM. Students' engagement has been linked with school success at all educational levels and persistence in STEM (e.g., Nelson Laird, Chen, & Kuh, 2008; Sinatra et al., 2015). Moreover, evidence suggests that girls are particularly likely to decide not to pursue a STEM educational path before they even get to college (e.g., Legewie & DiPrete, 2012). As such, it is not hard to imagine that girls may fail to develop or maintain intentions to persist in physical and engineering science fields in reflection of their unremarkable daily experiences of engagement during relevant high school classes. However, regardless of the extent to which girls' in-class engagement predicts long-term persistence in physical and engineering science, the fact that girls demonstrate lower levels of daily engagement in high school physics, chemistry, and engineering classes presents a challenge to the goal of producing equitable educational experiences and outcomes in high school.

Our evidence suggests girls may be less engaged during physics, chemistry, and engineering classes because they do not feel the context is supporting their psychological needs for autonomy, competence, and relatedness. Given widespread criticism of the teaching practices in college-level STEM coursework that lead to low levels of engagement (e.g., Gasiewski, Eagan, Garcia, Hurtado, & Chang, 2012), our observed findings at the high school level are likely to replicate at the college level and further contribute to girls' underrepresentation in physical and engineering science fields. Based on our findings, we can conclude that girls *perceive* teachers to engage in autonomy-supportive practices (e.g., provision of choices, provision of rationales about the importance of course activities, use of activities that incorporate students' existing interests, consideration for negative affect, and provision of encouraging informational feedback) and *experience* need satisfaction to a lesser extent than boys in these science courses. Do high school teachers in physics, chemistry, and engineering science courses really support girls' motivation to a lesser extent than boys' motivation? This study cannot answer this question, though as previously reviewed, research on gender stereotypes in science and teachers' differential practice suggests that it is certainly possible. An important next step for research is to explore the extent to which observation methods replicate or extend findings based on students' perceptions.

Even if teachers do not objectively treat students differently, girls' less positive perceptions of teacher practices and poorer need satisfaction still predict lower levels of engagement compared to boys. On the one hand, it is possible that girls own gender stereotypes guide their expectations for how teachers support their motivation and what their motivational experiences should be like in physics, chemistry, and engineering classes. Consequently, they perceive teachers to support needs to a lesser extent and experience less need satisfaction in those classes compared to boys, independent of teachers' objective practice. However, both quantity and quality are key to determining the effects of teacher practices. In line with self-determination theory, students' interpretation of teacher practices and determination of how they relate to personal values, interests, and preferences is key to determining their effects. As such, an alternative explanation of the findings is that teachers infrequently engage in autonomy-supportive practices in a way that is meaningful to girls and, thus, readily perceived by girls and experienced as need supportive. That is, for example, the nature of the choices, rationales, or "interesting" activities that teachers provide may not be personally relevant to girls' and fail to tap into their values, interests, and preferences. To summarize, three potentially complementary, non-mutually exclusive explanations for the findings seem possible: (1) teachers may objectively engage in differential practices that students more or less accurately perceive, (2) girls' own gender socialization leads them to experience the classroom environment in ways that are consistent with gender stereotypes, or (3) girls fail to perceive teachers' motivation support and experience its benefits because the quality of support is not well aligned with girls' interest, values, and preferences. An important direction of future research will be to uncover which of these mechanisms explains gender differences in students' classroom experiences in physical and engineering science courses.

7.1 | Limitations and implications for future research

Given the importance of increasing women's representation in physical and engineering science, replicating and extending the findings of the current investigation in future research is crucial. Strengths of the current investigation include that it was the first to examine gender differences in students' experiences of autonomy support, need satisfaction, and engagement in science classrooms separately for various science fields and using an intensive longitudinal design that allowed us to assess daily experiences. Despite the strengths of the current design, there are also a number of limitations that need to be addressed in future research.

The exclusive reliance on student reports in the current investigation presents a limitation that should be addressed in future research. Although the focus on student perceptions of teachers' practice is reasonable given self-determination theory's assumption that it is students' subjective experiences that are the most powerful predictor of their motivation and engagement, response-bias and shared-method variance may have influenced the results. Accordingly, using independent observations of the classroom to explore the extent to which teachers are objectively engaging in differential levels of autonomy support across male and female students, and the extent to which such observations explain gender differences in students' engagement during class is an important next step in this scholarship, though we acknowledge that observations present their own set of limitations and biases. This is an important step in disentangling the extent to which various possible mechanisms (e.g., differences in teachers' objective treatment, differences in students' perceptions of treatment, or differences in students' experiences of treatment) explain the current findings.

Also related to our measures of teacher autonomy support, it is important to note that we focused on relatively general practices (e.g., choice provision, informational feedback, etc.) in science classrooms that are also routinely used to support students' autonomy, motivation, and learning across educational domains. While this study provided a good first step for understanding how experiences of those practices and their correlates may vary across students in a science classroom, it may be useful for future research to explore the use and correlates of autonomy-supportive practices that are more unique to the science classroom.

Moreover, students' intentions to take STEM coursework or pursue STEM educational and career paths was not measured in this study, nor was the extent to which students or teachers implicitly or explicitly endorsed gender stereotypes in science. However, information about each would be important to further explain the underlying mechanisms of the findings observed in this investigation. Qualitative assessments of students' experiences of need support in the classroom would also be a useful strategy for uncovering the extent to which the autonomy support teachers provide is or is not perceived as personally meaningful to girls. Along these lines of understanding mechanisms and the context of these findings, future research might consider the extent to which the gender of the teacher explains variation in the relations between students' perceptions of autonomy support, experiences of need satisfaction, and engagement. Although the small teacher sample and uneven distribution of teacher genders across science domains prevented us from exploring this possibility, research pointing to the importance of role models in girls' identities and persistence in STEM (e.g., Ramsey, Betz, & Sekaquaptewa, 2013) suggests that gender and other characteristics of the teacher (e.g., femininity, masculinity) may influence these relations.

We also note that the intense nature of students' reporting across six weeks in the current design required that we rely on a small sample of volunteers from each class and teachers who were willing to participate. Although we randomly selected student participants among volunteers and successfully recruited approximately 40% of teachers across participating schools, the voluntary and selective nature of the sample leaves open the possibility of biased results that are idiosyncratic to the current sample. Future research should attempt to address this limitation by randomly sampling classes and students.

Finally, the correlational nature of the design means that the present findings cannot be taken to imply causation. Consequently, findings of this investigation should be corroborated with experimental designs in authentic classroom contexts that intentionally target girls' need satisfaction and engagement via teachers' motivating practices. Although current reform efforts have primarily focused on increasing girls' STEM opportunities, achievement, and exposure to advanced curriculums, this and other evidence suggests that such efforts alone will not be sufficient for addressing the complexity of motivational factors that contribute to girls' underrepresentation in in physical and engineering science fields. The findings of this research suggest that interventions should include professional development to help high school teachers effectively use autonomy-supportive practices to boost motivation and engagement, particularly for female students. Targeting teachers' daily motivating practice stands out as a potentially powerful lever for reducing gender gaps, as teachers may be more accessible and amenable to change than families or media, who also play critical roles.

8 | CONCLUSION

In conclusion, this investigation adds to the vital body of research exploring girls' motivation in science in an effort to explain their underrepresentation in physical and engineering science. This study goes beyond those previously conducted by using an intensive longitudinal design to examine students' perceptions of the autonomy support they received from teachers, their experience of need satisfaction, and their engagement in science classrooms on a daily basis. Results suggested that girls perceived their teachers to provide less daily autonomy support and experienced less daily need satisfaction during physics, chemistry, and engineering classes. In turn, results suggested that these experiences explained girls' lower daily engagement in these courses compared to boys. We encourage researchers and educators to test the effectiveness of interventions that focus on supporting adolescents', and especially girls', critical psychological needs in science classes. We expect that such efforts will have significant payoffs for increasing gender diversity in STEM.

ENDNOTES

[1] Competence need satisfaction shares similarities with related concepts, including self-efficacy, self-concepts of ability, expectations for success (e.g., see Bandura, 1997; Valentine, Dubois, & Cooper, 2004; Bong & Clark, 1999) and other competence beliefs that prior evidence has linked with gender differences in STEM (see Wang & Degol, 2013, 2017 for review).

[2] Given extensive existing evidence linking autonomy-supportive practices and need satisfaction (e.g., Jang, Kim, & Reeve, 2012; 2016; Patall, Dent, Oyer, & Wynn, 2013; Patall, Vasquez, Steingut, Trimble, & Pituch, 2017; Reeve & Jang, 2006), we focus on the mediating role of each need in the link between gender and daily engagement, rather than looking at the link between perceived support and need satisfaction in mediation analyses.

[3] Of the 30 female teachers, 16 taught a biology course and 14 taught a physical science (e.g., physics, chemistry) course. Of the 10 male teachers, 3 taught a biology course and 8 taught a physical science course. Given the small sample size at the teacher level and the uneven distribution of teacher genders across science domains, we did not conduct any analyses to examine whether teacher gender moderated relationships between students' perceptions of teachers' autonomy support, need satisfaction, and engagement. Such analyses would have limited statistical power and would not be reliable.

[4] Group-mean centering of continuous and dichotomous predictors is often recommended by multilevel modeling experts (see Raudenbush & Bryk, 2002, p. 34; Enders, & Tofighi, 2007). The primary reason for such centering is to ensure that slope estimates of focal variables at lower levels of the analysis (e.g., day and student) are unbiased. Use of grand-mean centering (or no centering) can produce a slope value that is an "uninterpretable blend" of within-cluster and between-cluster slopes for predictors at lower levels of the analysis (Raudenbush & Bryk, 2002). Since grand-mean centering can produce biased slope estimates for predictor variables at the lower levels of the analysis, we used group-mean centering, which avoids this bias, for our focal (i.e., non-control) variables of interest.

ORCID

Erika A. Patall http://orcid.org/0000-0003-3448-8932

REFERENCES

Allchin, D. (1999). Values in science: An educational perspective. *Science & Education*, *8*(1)1–12. https://doi.org/10.1023/A:1008600230536

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, *333*(6045), 1034–1037. https://doi.org/10.1126/science.1207998

Allen-Ramdial, S.-A. A., & Campbell, A. G. (2014). Reimagining the pipeline: Advancing STEM diversity, persistence, and success. *BioScience*, *64*(7), 612–618. https://doi.org/10.1093/biosci/biu076

Altermatt, E. R., Jovanovic, J., & Perry, M. (1998). Bias or responsivity? Sex and achievement-level effects on teachers' classroom questioning practices. *Journal of Educational Psychology*, *90*(3), 516–527. https://doi.org/10.1037//0022-0663.90.3.516

Archambault, I., Janosz, M., Fallu, J.-S., & Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *Journal of Adolescence*, *32*(3), 651–670. https://doi.org/10.1016/j.adolescence.2008.06.007

Assor, A., Kaplan, H., Kanat-Maymon, Y., & Roth, G. (2005). Directly controlling teacher behaviors as predictors of poor motivation and engagement in girls and boys: The role of anger and anxiety. *Learning and Instruction*, *15*(5), 397–413. https://doi.org/10.1016/j.learninstruc.2005.07.008

Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *The British Journal of Educational Psychology*, *72*(2), 261–278. https://doi.org/10.1348/000709902158883

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

Baron-Cohen, S. (2003). *The essential difference: The truth about the male and female brain*. New York, NY: Basic Books.

Bartos, S. A., & Lederman, N. G. (2014). Teachers' knowledge structures for nature of science and scientific inquiry: Conceptions and classroom practice. *Journal of Research in Science Teaching*, *51*(9), 1150–1184. https://doi.org/10.1002/tea.21168

Becker, J. R. (1981). Differential treatment of females and males in mathematics classes. *Journal for Research in Mathematics Education*, *12*(1), 40–53. https://doi.org/10.2307/748657

Belmont, M., Skinner, E., Wellborn, J. G., & Connell, J. P. (1992). *Teacher as social context (TASC): Two measures of teacher provision of involvement, structure, and autonomy support*. Rochester, NY: Graduate School of Education and Human Development, Department of Psychology, University of Rochester.

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, *84*(6), 740–756. https://doi.org/10.1002/1098-237X(200011)84:6<740::AID-SCE4>3.0.CO;2-3.

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.

Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, *34*(3), 139–153. https://doi.org/10.1207/s15326985ep3403_1

Cataldi, E. F., Siegel, P., Shepherd, B., & Cooney, J. (2014). *Baccalaureate and beyond: A first look at the employment experiences and lives of college graduates, 4 years On* (No. NCES 2014–141). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from https://nces.ed.gov/pubsearch

Ceci, S. J., Williams, W. M., & Thompson, R. F. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(8), 3157–3162.

Chalabaev, A., Sarrazin, P., Trouilloud, D., & Jussim, L. (2009). Can sex-undifferentiated teacher expectations mask an influence of sex stereotypes? Alternative forms of sex bias in teacher expectations. *Journal of Applied Social Psychology*, *39*(10), 2469–2498. https://doi.org/10.1111/j.1559-1816.2009.00534.x

Chambers, D. W. (1983). Stereotypic images of the scientist: The draw-a-scientist test. *Science Education*, *67*(2), 255–265. https://doi.org/10.1002/sce.3730670213

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Connell, J. P. (1990). Context, self and action: A motivational analysis of self-system processes across the life-span. In D. Cichhetti (Ed.), *The self in transition: From infancy to childhood* (pp. 61–97). Chicago, IL: University of Chicago Press.

Correll, P. (2001). Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology*, *106*(6), 1691–1730.

Croninger, R., & Lee, V. (2001). Social capital and dropping out of high school: Benefits to at-risk students of teachers' support and guidance. *Teachers College Record*, *103*(4), 548–581. https://doi.org/10.1111/0161-4681.00127

Eagly, A. H., & Wood, W. (2012). Social role theory. In Paul A. M. Van Lange, Arie W. Kruglanski, & E. Tory Higgins (Eds.), *Handbook of theories of social psychology: Volume 2* (Vols. 1–2, pp. 458–476). London: SAGE Publications Ltd. https://doi.org/10.4135/9781446249222

Eccles, J. S., & Harold, R. D. (1992). Gender differences in educational and occupational patterns among the gifted. In N. Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), *Talent development: Proceedings from the 1991 Henry B. and Jocelyn Wallace National Research symposium on talent development* (pp. 3–29). Unionville, NY: Trillium Press.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*(3), 830–847. https://doi.org/10.2307/1131221

Eccles-Parsons, J., Adler, T. F., & Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: Parental influences. *Child Development*, *53*, 322–339. https://doi.org/10.2307/1128973

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 131–148.

Erikson, E. H. (1968). *Identity: Youth and crisis*. Oxford, England: Norton & Co.

Farland-Smith, D. (2009). How does culture shape students' perceptions of scientists? Cross-national comparative study of American and Chinese elementary students. *Journal of Elementary Science Education*, *21*(4), 23–42. https://doi.org/https://doi.org/10.1007/BF03182355

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, *21*(1), 55–69. https://doi.org/https://doi.org/10.1007/BF00311015

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

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, *20*(2), 507–537. https://doi.orghttps://doi.org/10.1111/j.1532-7795.2010.00645.x

Furrer, C., & Skinner, E. A. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, *95*(1), 148–162. https://doi.org/10.1037/0022-0663.95.1.148

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education*, *53*(2), 229–261. https://doi.org/10.1007/s11162-011-9247-y

Green, S. K. (2002). Using an expectancy-value approach to examine teachers' motivational strategies. *Teaching and Teacher Education*, *18*(8), 989–1005. https://doi.org/10.1016/S0742-051X(02)00055-0

Griffith, A. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, *29*(6), 911–922. https://doi.org/10.1016/j.econedurev.2010.06.010

Hafen, C. A., Allen, J. P., Mikami, A. Y., Gregory, A., Hamre, B., & Pianta, R. C. (2012). The pivotal role of adolescent autonomy in secondary school classrooms. *Journal of Youth and Adolescence*, *41*(3), 245–255. https://doi.org/10.1007/s10964-011-9739-2

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, *8*(1), 1–51. https://doi.org/10.1111/j.1529-1006.2007.00032.x

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

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. New York: Routledge.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*(5888), 494–495. https://doi.org/10.1126/science.1160364

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

Jang, H., Kim, E. J., & Reeve, J. (2016). Why students become more engaged or more disengaged during the semester: A self-determination theory dual-process model. *Learning and Instruction*, *43*, 27–38. https://doi.org/10.1016/j.learninstruc.2016.01.002

Jones, M. G., & Wheatley, J. (1990). Gender differences in teacher-student interactions in science classrooms. *Journal of Research in Science Teaching*, *27*(9), 861–874. https://doi.org/10.1002/tea.3660270906

Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. *Advances in Experimental Social Psychology*, *28*(c), 281–388. https://doi.org/10.1016/S0065-2601(08)60240-3

Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality & Social Psychology Review*, *9*(2), 131–155. https://doi.org/10.1207/s15327957pspr0902_3

Katz, I., Kaplan, A., & Gueta, G. (2010). Students' needs, teachers' support, and motivation for doing homework: A cross-sectional study. *Journal of Experimental Education*, *78*(2), 246–267. https://doi.org/10.1080/00220970903292868

Keller, C. (2010). Effect of teachers' stereotyping of students' stereotyping of mathematics as a male domain. *The Journal of Social Psychology*, *141*, 165–173. https://doi.org/10.1080/00224540109600544

Kim-Spoon, J., & Grimm, K. J. (2016). Latent growth modeling and developmental psychopathology. In D. Cicchetti (Ed.), *Developmental psychopathology* (3rd ed., Vol. 1, pp. 986–1041). Theory and Method. Hoboken, NJ: John Wiley & Sons.

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). New York: Guilford Publications.

Köller, O., Baumert, J., & Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *Journal for Research in Mathematics Education*, *32*(5), 448–470. https://doi.orhttps://doi.org/10.2307/749801

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Ladd, G. W., & Dinella, L. M. (2009). Continuity and change in early school engagement: Predictive of children's achievement trajectories from first to eighth grade? *Journal of Educational Psychology*, *101*(1), 190–206. https://doi.org/10.1037/a0013153

Legewie, J., & DiPrete, T. A. (2012). School context and the gender gap in educational achievement. *American Sociological Review*, *77*(3), 463–485. https://doi.org/10.1177/0003122412440802

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science (New York)*, *347*(6219), 262–265. https://doi.org/10.1126/science.1261375

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research*, *41*(1), 63–76. https://doi.org/10.1080/0013188990410106

Milford, T. M., & Tippett, C. D. (2013). Preservice teachers' images of scientists: Do prior science experiences make a difference? *Journal of Science Teacher Education*, *24*(4), 745–762. https://doi.org/10.1007/s10972-012-9304-1

Murdock, T. B. (1999). The social context of risk: Status and motivational predictors of alienation in middle school. *Journal of Educational Psychology*, *91*(1), 62–75. https://doi.org/10.1037/0022-0663.91.1.62

Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., & Eccles, J. (2008). Gendered high school course selection as a precursor of gendered occupational careers: The mediating role of self-concept and intrinsic value. In Helen M. G. Watt, PhD, & Jacquelynne S. Eccles, PhD (Eds.), *Gender and occupational outcomes: Longitudinal assessment of individual, social, and cultural influences* (pp. 115–143). APA Books.

National Academy of Sciences. (2007a). *Beyond bias and barriers: Fulfilling the potential of women in academic science and engineering*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11741

National Academy of Sciences. (2007b). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11463

National Science Foundation. (2017). *Women, minorities, and persons with disabilities in science and engineering: 2017* (Special Report NSF 17–310). Arlington, VA: National Science Foundation. Retrieved from www.nsf.gov/statistics/wmpd/

Nelson Laird, T. F., Chen, D., & Kuh, G. D. (2008). Classroom practices at institutions with higher-than-expected persistence rates: What student engagement data tell us. *New Directions for Teaching & Learning*, *2008*(115), 85–99. https://doi.org/10.1002/tl.327

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., … Steele, C. M. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10593–10597.

Patall, E. A., Dent, A. L., Oyer, M., & Wynn, S. R. (2013). Student autonomy and course value: The unique and cumulative roles of various teacher practices. *Motivation and Emotion*, *37*(1), 14–32. https://doi.org/10.1007/s11031-012-9305-6

Patall, E. A., & Leach, J. (2015). The role of choice provision in academic dishonesty. *Contemporary Educational Psychology*, *42*, 97–110. https://doi.org/10.1016/j.cedpsych.2015.06.004

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), 269–288. https://doi.org/10.1037/edu0000214

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 & Instruction*, *35*(4), 337–362. https://doi.org/10.1080/07370008.2017.1358722

Ramsey, L. R., Betz, D. E., & Sekaquaptewa, D. (2013). The effects of an academic environment intervention on science identification among women in STEM. *Social Psychology of Education*, *16*(3), 377–397.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Thousand Oaks, CA: Sage.

Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, *44*(3), 159–175. https://doi.org/10.1080/00461520903028990

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology*, *105*(3), 579–595. https://doi.org/10.1037/a0032690

Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, *98*(1), 209–218. https://doi.org/10.1037/0022-0663.98.1.209

Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion*, *28*(2), 147–169. https://doi.org/10.1023/B:MOEM.0000032312.95499.6f

Reeve, J., & Lee, W. (2014). Students' classroom engagement produces longitudinal changes in classroom motivation. *Journal of Educational Psychology*, *106*(2), 527–540. https://doi.org/10.1037/a0034934

Reeve, J., Nix, G., & Hamm, D. (2003). Testing models of the experience of self-determination in intrinsic motivation and the conundrum of choice. *Journal of Educational Psychology*, *95*(2), 375–392.

Reeve, J., & Sickenius, B. (1994). Development and validation of a brief measure of the three psychological needs underlying intrinsic motivation: The AFS scales. *Educational and Psychological Measurement*, *54*(2), 506–515. https://doi.org/10.1177/0013164494054002025

Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, *36*(4), 257–267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Riegle-Crumb, C., & King, B. (2010). Questioning a white male advantage in STEM: Examining disparities in college major by gender and race/ethnicity. *Educational Researcher*, *39*(9), 656–664. https://doi.org/10.3102/0013189X10391657

Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2012). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM college majors over time. *American Educational Research Journal*, *49*(6), 1048–1073. https://doi.org/10.3102/0002831211435229

Roesch, S. C., Aldridge, A. A., Stocking, S. N., Villodas, F., Leung, Q., Bartley, C. E., & Black, L. J. (2010). Multilevel factor analysis and structural equation modeling of daily diary coping data: Modeling trait and state variation. *Multivariate Behavioral Research*, *45*(5), 767–789. https://doi.org/10.1080/00273171.2010.519276

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68. https://doi.org/10.1037/0003-066X.55.1.68

Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and volatility of STEM career interest in high school: A gender study. *Science Education*, *96*(3), 411–427. Htpps://doi.orghttps://doi.org/10.1002/sce.21007

Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles*, *50*(11), 835–850. https://doi.org/10.1023/B:SERS.0000029101.74557.a0

Selig, J. P., & Preacher, K. J. (2008). Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects [Computer software].

Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin*, *26*(4), 419–435. https://doi.org/10.1177/0146167200266002

Shumow, L., & Schmidt, J. A. (2013). Academic grades and motivation in high school science classrooms among male and female students: Associations with teachers' characteristics, beliefs, and practices. In Rolf Haumann, & George Zimmer (Eds.), *Handbook of academic performance: Predictors, learning strategies and influences of gender* (pp. 53–72). Nova Science Publishers, Inc.

Simpkins, S. D., Fredricks, J., & Eccles, J. S. (2012). Charting the Eccles' expectancy–value model from mothers' beliefs in childhood to youths' activities in adolescence. *Developmental Psychology*, *48*, 1019–1032. https://doi.org/10.1037/a0027468

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. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, *85*(4), 571–581. https://doi.org/10.1037/0022-0663.85.4.571

Skinner, E. A., Kindermann, T. A., Connell, J. P., & Wellborn, J. G. (2009). Engagement and disaffection as organizational constructs in the dynamics of motivational development. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 223–245). New York, NY: Routeledge.

Snyder, T. D., de Brey, C., & Dillow, S. A. (2016). *Digest of education statistics, 2014* (No. NCES 2016-006). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from https://nces.ed.gov/pubs2016/2016006.pdf

Solomon, D., Battistich, V., & Hom, A. (1996). Teacher beliefs and practices in schools serving communities that differ in socioeconomic level. *The Journal of Experimental Education*, *64*(4), 327–347. https://doi.org/10.1080/00220973.1996.10806602

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology*, *102*(4), 947–963. https://doi.org/10.1037/a0019920

Su, Y.-L., & Reeve, J. (2011). A meta-analysis of the effectiveness of intervention programs designed to support autonomy. *Educational Psychology Review*, *23*(1), 159–188. https://doi.org/10.1007/s10648-010-9142-7

Tai, R. H., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Science*, *312*(5777), 1143–1144.

Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics*, *50*(1), 49–62.

Turner, J., & Patrick, H. (2004). Motivational influences on student participation in classroom learning activities. *Teachers College Record*, *106*(9), 1759–1785. https://doi.org/10.1111/j.1467-9620.2004.00404.x

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational psychologist*, *39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence*, *38*(4), 412–423. https://doi.org/10.1016/j.intell.2010.04.006

Wang, M.-T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*(4), 304–340. https://doi.org/10.1016/j.dr.2013.08.001

Wang, M.-T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review*, *29*(1), 119–140. https://doi.org/10.1007/s10648-015-9355-x

Watt, H. M. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th-through 11th-grade Australian students. *Child Development*, *75*(5), 1556–1574. https://doi.org/10.1111/j.1467-8624.2004.00757.x

Wellborn, J. G., & Connell, J. P. (1987). *Manual for the Rochester assessment package for schools*. Rochester, NY: University of Rochester.

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. *Journal of Personality and Social Psychology*, *70*, 767–779. https://doi.org/10.1037/0022-3514.70.4.767

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology*, *96*(2), 236–250. https://doi.org/10.1037/0022-0663.96.2.236

Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. *Advances in Experimental Social Psychology*, *46*, 55–123. https://doi.org/10.1016/B978-0-12-394281-4.00002-7

How to cite this article: Patall EA, Steingut RR, Freeman JL, Pituch KA, Vasquez AC. Gender disparities in students' motivational experiences in high school science classrooms. *Sci Ed*. 2018;1–27. https://doi.org/10.1002/sce.21461