International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

The psychological characteristics of experiences that influence science motivation and content knowledge

Meghan Bathgate & Christian Schunn

To cite this article: Meghan Bathgate & Christian Schunn (2017): The psychological characteristics of experiences that influence science motivation and content knowledge, International Journal of Science Education, DOI: 10.1080/09500693.2017.1386807

To link to this article: http://dx.doi.org/10.1080/09500693.2017.1386807

Published online: 26 Oct 2017.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2017
https://doi.org/10.1080/09500693.2017.1386807

Routledge
Taylor & Francis Group

Check for updates

The psychological characteristics of experiences that influence science motivation and content knowledge

Meghan Bathgate[a] and Christian Schunn[b]

[a]Center for Teaching and Learning, Yale University, New Haven, CT, USA; [b]Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, USA

ABSTRACT

While motivational changes towards science are common during adolescence, our work asks which perceived classroom experiences are most strongly related to these changes. Additionally, we examine which experiences are most strongly associated with learning classroom content. In particular, using self-reports from a sample of approximately 3000 middle school students, this study investigates the influence of perceived science classroom experiences, namely student engagement and perceived success, on motivational change (fascination, values, competency belief) and content knowledge. Controlling for demographic information, school effects, and initial levels of motivation and content knowledge, we find that dimensions of engagement (affect, behavioural/cognitive) and perceived success are differentially associated with changes in particular motivational constructs and learning. Affective engagement is positively associated with motivational outcomes and negatively associated with learning outcomes, behavioural–cognitive engagement is associated only with learning, and perceived success is related only to motivational outcomes. Theoretical and practical implications are discussed.

ARTICLE HISTORY
Received 14 October 2016
Accepted 27 September 2017

KEYWORDS
Motivation; learning outcomes; K-12; engagement

Introduction

The importance of educating upcoming generations in the sciences has been extensively argued (National Research Council [NRC], 2008, 2009; PCAST, 2012), but many students are performing poorly and are losing a desire to persist in science (NRC, 2009; Osborne, Simon, & Collins, 2003; Shen & Tam, 2008). Yet, there is growing evidence that this decline is not uniform, but rather is influenced by what kinds of instruction students experience (Gottfried, Fleming, & Gottfried, 2001; Maltese, Melki, & Wiebke, 2014; Vedder-Weiss & Fortus, 2011, 2012). As such, rather than focusing on motivational loss during middle school, we are interested in understanding how students' perceptions of classroom science experiences (e.g. levels and nature of engagement during a science class activity) contribute to motivational changes. Further, we wish to understand whether the same experiences that improve one's learning of science content are associated with growth in motivation.

CONTACT Meghan Bathgate meghan.bathgate@yale.edu Center for Teaching and Learning, Yale University, 320 York St., New Haven, CT 06511, USA

Identifying if, and to what extent, the same classroom experiences relate to motivational changes and content knowledge in science would inform the development of interventions that specifically target one, or both, of these outcomes. Understanding these relationships provides the opportunity for better alignment between interventions and particular outcomes. For example, if an intervention is specifically targeted at building competency beliefs, there may be a specific element related to confidence that an activity emphasizes, as opposed to encouraging general engagement. Further, by measuring a variety of outcomes in the same sample (motivational changes and content knowledge outcomes), our findings tease apart separate effects, providing theoretical contribution as to how experiences in middle school science courses relate to specific motivational change and content learning within the same individuals.

What are motivational pillars in science?

Before investigating factors that may attenuate science motivation during middle school, it is important to identify critical motivational outcomes that early experiences should seek to support. We focus on motivational factors based on expectancy-value theory (Wigfield & Eccles, 2000), but also have support from work on interest (Hidi & Renninger, 2006), self-efficacy (Bandura, 1997), and self-determination theory (Ryan & Deci, 2000). Three motivation constructs are broadly recognised as important: learners need to be (1) intrigued by scientific phenomena and concepts (*intrinsic motivation*), (2) see both the daily and long-term value of knowing science and scientific practices (*utility motivation*), and (3) feel capable of engaging in scientific practices and discussions (*competency beliefs*). These concepts are captured in the following motivational constructs.

Intrinsic motivation: fascination

Intrinsic motivation towards a task arises from a learner's internal desire to engage with the task and underlying topic itself (Bathgate, Schunn, & Correnti, 2013; Bryan, Glynn, & Kittleson, 2011; Deci & Ryan, 2000; Ryan & Deci, 2000). Learners who are intrinsically motivated towards an activity, for example, are driven by their enjoyment of the activity itself, as opposed to being motivated by the consequences of that activity (e.g. getting a good grade on an activity). Interest is one example of this type of motivation. The role interest plays in driving behaviour has gained increasing attention, as researchers and educators continue to demonstrate its association with persistence in science (Simpkins, Davis-Kean, & Eccles, 2006), future science choices (Sha, Schunn, & Bathgate, 2015), and science learning (Glynn, Brickman, Armstrong, & Taasoobshirazi, 2011; Zusho, Pintrich, & Coppola, 2003). Defined as both an emotional and cognitive factor, interest is thought to develop over time as learners reengage positively with particular content (Hidi & Renninger, 2006), such as science. When learners are interested in an area, particularly when that interest is at least somewhat developed and stable over time, they tend to persist and enjoy future activities related to that same content (Hidi & Renninger, 2006). For example, if a learner positively engages with science material that sparks her interest, the more likely she will seek out future related science activities and have positive experiences with them.

In science, intrinsic motivation also involves curiosity in the workings of the natural world (Baram-Tsabari & Yarden, 2008; Jenkins & Pell, 2006; Prokop, Prokop, &

Tunnicliffe, 2007; Zimmerman, 2012). Both interest and curiosity are usually found to coincide with a mastery approach (Cho & Summers, 2012; Hilpert, Stempien, van der Hoeven Kraft, & Husman, 2013), in which a learner desires to deeply understand a topic or discipline (Ames, 1992; Elliot, 1999; Elliot & McGregor, 2001). In general, a mastery approach is often thought to be intrinsically driven, where a learner is pursuing a knowledge or skill for its own reward as opposed to proving or demonstrating knowledge (referred to as a performance approach; Elliot & McGregor, 2001).

Interest, curiosity, and mastery goals each has been associated with deeper learning and persistence (Elliot, 1999; Elliot & McGregor, 2001; Hidi & Renninger, 2006; Loewenstein, 1994; Richey, Nokes-Malach, & Wallace, 2014; Silvia, 2006). Because of the similar effects on learning behaviours, their regular co-occurrence (Hilpert et al., 2013) to the point of loading on a single underlying factor, interest, curiosity, and mastery are conceptualised as components of the overarching construct of *Fascination* (Activation Lab Fascination Technical Report, 2014; http://www.activationlab.org/tools/). Collectively, Fascination then involves the interest and positive affect one has towards science, curiosity towards the natural world, and goals towards acquiring and mastering scientific skills and ideas.

Extrinsic motivation: values

Often held in contrast to intrinsic motivation, extrinsic motivations arise from a desire to meet a secondary goal. In other words, an extrinsically motivated learner participates in an activity for the consequences of that activity (e.g. receiving a good grade or payment for completing a task), irrespective of the enjoyment he receives from participating. For example, a student may choose to enrol in a science camp because they need additional experience to apply for college, not because they enjoy the processes of science. The value students have towards science in middle school is not only predictive of later science value, but also longer-term successes in science, such as selecting future science courses (Simpkins et al., 2006).

Our conceptualisation of value stems from expectancy-value theory (Wigfield & Eccles, 2000) – specifically the concept of utility value – and represents both the personal and societal value one places on science towards meeting another goal. In the area of science specifically, extrinsic motivation reflects the value the learner places on science because of its utility in meeting personal goals (e.g. doing well in classes, understanding how the world works) or its utility to society (e.g. science is helpful to solving environmental problems), standing in contrast to the intrinsic motivation of simply being curious or enjoying the content. Both intrinsic and extrinsic motivations can drive student participation in science, but intrinsic motivations tend to produce deeper learning processes and thus better learning outcomes (Pintrich, 2004; Belenky & Nokes-Malach, 2012).

The relationship between intrinsic and extrinsic is not always opposing. In fact, students can be highly intrinsically interested in science while simultaneously seeing the extrinsic value of science (e.g. a student deeply enjoys doing science research and sees how that work is meeting her longer-term goals). As such, these seemingly discrete concepts are often positively correlated within an individual.

Competency beliefs

Motivational theories such as Eccles and Wigfield's expectancy-value theory (Eccles & Wigfield, 2002) and Bandura's work in efficacy beliefs (Bandura, 1993, 1997), as well as reviews of motivational theories (Koballa & Glynn, 2007; Pintrich, 2003), make an important delineation between the (intrinsic or extrinsic) valuation of an area and learner perceptions of ability to perform well in that area (i.e. their competency beliefs). For example, a learner may be interested in a topic but feels he or she does not have the skills needed to do well within an activity about that topic (e.g. a student is interested in robotics but does not think he or she would do well participating in a robotics club). These competency beliefs have been studied specifically in the context of science and have repercussions for whether a leaner chooses to participate in science activities, as well as how well they perform within them (Beghetto, 2007; Britner & Pajares, 2006; Glynn et al., 2011).

The psychological experiences related to motivational change

Learner's experiences within a science activity can influence these three motivational factors. For example, a learner given too little support to complete a complex science activity may not be able to complete the activity well and will leave that experience feeling discouraged. Perhaps with repeated experiences like this one, they conclude they are not good at science and become disinterested. Conversely, a learner given more support may have positive, successful experiences that then build a stronger sense of competence to develop a stable interest in science and expand their interest and knowledge.

Learners interact with science across a range of settings. We focus on classroom activities for two reasons. First, while not every student has access to quality out-of-school science activities (e.g. camps, museums), all students take part in school science from middle school onwards. As such, understanding the influence of these experiences allows us to explore the mechanisms at work within them and maximise the positive influences of these activities (and suppress the negative) to impact many students. Finally, these experiences are malleable through a number of education policies (e.g. new curriculum guidelines or teacher professional development efforts).

Within classroom activities, we focus on two well-researched constructs, engagement and perceived success, each of which has been studied independently in relation to motivation or content knowledge in science, the outcomes of the current study. Perceived science classroom experiences, that is one's self-reported emotional, behavioural, and cognitive experiences within a classroom, provide learners explicit feedback on achievement and repeated exposure to science content. We focus on perceived experiences (rather than objective characteristics such as presence of hands-on materials or amount of teacher support) because measuring learners' perceptions (as opposed to measuring type of activity, for example) provides insight into how learners experience a situation rather than assuming the type of activity is indicative of the experience itself. That is, it is the perception of these experiences that are expected to influence motivation, rather than the type of activity itself. We recognise some activities may be more generally engaging, but the mechanism for motivational change lies within learners' perceptions of this experience.

Engagement

Engagement refers to the way an individual interacts with a particular task (or repeatedly in a domain across many such tasks) and is generally thought to consist of three dimensions: (1) affective engagement, how one feels (i.e. the emotional experience) during an activity or task, (2) behavioural engagement, what one actually does during an activity or task, and (3) cognitive engagement, the way one thinks (i.e. the type of cognitive processing utilised, such as degree of attention and making connections among ideas) during an activity or task (Fredricks et al., 2011; Fredricks, Blumenfeld, & Paris, 2004). Overall, engagement has a robust connection with achievement, participation, and motivation in many educational areas and specifically in the sciences (Ainley & Ainley, 2011; Fredricks et al., 2004; Pekrun & Linnenbrink-Garcia, 2012; Bathgate & Schunn, in press; Tytler & Osborne, 2012; Wang & Eccles, 2012). Specifically, affective (Bathgate & Schunn, in press), behavioural and cognitive (Connell & Wellborn, 1991; Marks, 2000; Wang & Eccles, 2012) engagement each has empirical relationships with achievement in adolescence. Engagement has also been shown to relate to high school students' academic motivation (e.g. self-efficacy) over time (e.g. Reeve & Lee, 2014).

Therefore, we expect there to be a relationship between the degree of engagement a learner experiences within an activity and larger changes in their motivation and learning over time in the current similarly aged sample. However, as the field learns more about engagement and how it functions, more questions are raised regarding the relationship and influence of the individual engagement dimensions on various outcomes, and the best way to measure engagement across diverse science learning activities. For example, each dimension may be differentially associated with a given scenario and even the empirical distinctions among the dimensions are not robustly established (Bathgate et al., 2013; Wang & Degol, 2014). Depending on the measure and context, behavioural and cognitive engagement, in particular, may be related. In highly constrained tasks, behavioural data can be thought of representative of cognition, such as the context of eye-tracking studies where delayed movement is associated with increased cognition or processing. We address these open questions by examining the structure of engagement relative relationship of engagement dimensions on multiple motivational and learning outcomes using a measure targeted at students' reflection immediately following their experience.

Perceived success

In addition to a learner's immediate engagement in an activity, a learner also has a sense of how well they performed on a given task, often based on implicit or explicit feedback they receive during or following the activity. These perceived success experiences are hypothesised to relate to one's seeking similar experiences and, subsequently, more challenging experiences to develop their skills and also have a history of being associated with achievement (e.g. Bandura, 1993; Britner & Pajares, 2006; Ryan & Deci, 2000; Schiefele, 2009). For example, a learner who feels as if they did well on an activity shows greater interest in similar activities in the future and may feel more capable of performing well on them (Pajares, 1997).

Both engagement and perceived success are malleable factors that can be supported by teacher, peer, or tool scaffolds (Jang, Reeve, & Deci, 2010; Marks, 2000; Wang & Degol, 2014; Wang & Holcombe, 2010) and each of these elements (engagement and perceived success) are theorised to impact not only the immediate experience of a learner, but also build towards their broader motivations and learning within a domain (i.e. science) (Vedder-Weiss & Fortus, 2011, 2012). Therefore, we focus on these constructs as the main potential predictors of motivational change and content knowledge gains in middle school science.

Current study

It is clear from the existing literature that the concepts described above have a strong history of empirical research; however, there remain critical open questions our current work will address. Specifically, while we know engagement and perceived success are associated with motivation and learning outcomes, from a theory building perspective and a pragmatic perspective, we do not know their influence *across* these outcomes (i.e. do they have a similar impact for fascination vs. value vs. content learning?). From a scientific rigour perspective, we also do not know their relative effects when examined concurrently (i.e. is each still influential when controlling for other?). Prior studies with a narrower focus may have false attributed connections by failing to include appropriate controls given that the experience factors are likely correlated (e.g. affective engagement and perceived success are likely correlated). By using a single data set with four outcomes (fascination, values, competency belief, and content knowledge) and measuring engagement and perceived success at the same time, we are able to describe the pattern of results across three types of motivation change[1] and content learning as well as control for any shared variance between engagement and perceived success. Figure 1 shows our overall framework and the connections among our variables tested in the study. The literature contains at least some support for each of the tested connections. Here we examine which connections are robust when correlations among experiential factors and important confounds are appropriately controlled.

Specifically, we ask the following research questions (RQs):

RQ 1: Which aspects of motivational change are associated with perceived classroom experiences of engagement and perceived success?

RQ 2: Are engagement and perceived success associated with content knowledge?

As part of addressing these questions, we also examine the internal structure and co-variation of engagement and perceived success in middle school classroom science. To

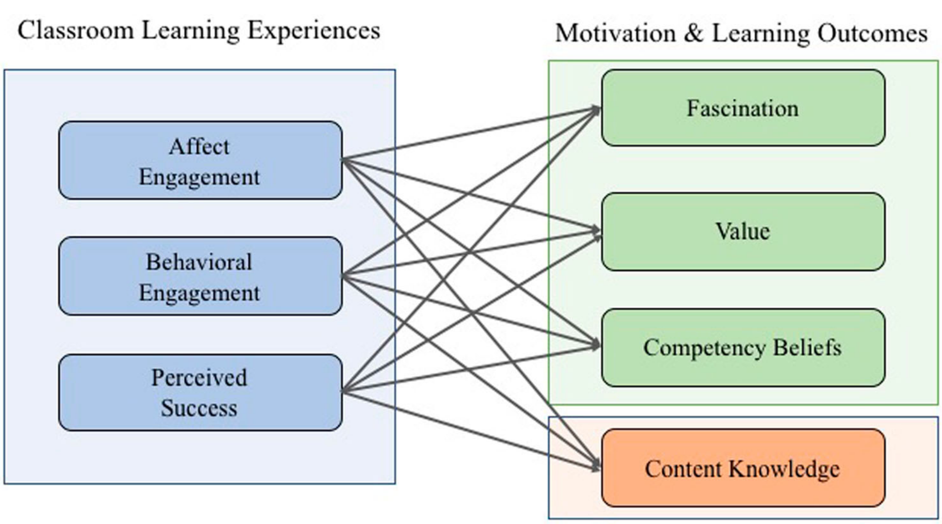

Figure 1. Tested model of how classroom learning experiences to motivation and learning outcomes.

foreshadow the results, it is possible that some aspects co-occur so highly in this kind of learning context that separation of all four aspects is not possible, which is itself an important finding.

Method

Participants

These data were taken from the Activation Lab: Enables Success 2014 study whose focus is on urban middle school students in the United States. Approximately 3000 middle school students (49% sixth grade, 51% eighth grade) provided at least some data in this study, although the sample size varies by analyses depending on the completion of particular instruments, as commonly occurs in longitudinal data collection in urban school contexts. ||Specifically, 2707 students completed all time-1 measures, 2312 completed all time-2 measures, 2233 completed all classroom experience measures (both days of engagement and perceived success surveys; 793 completed the classroom experience measures on four days to test temporal stability of the experience sampling method), and 1180 completed all demographic data. Together, 901 students provided every piece of data discussed in the main study analyses; however, many of our analyses have upwards of 1600 students since they do not require full data from every student. Therefore, sample size varies by analyses and these differences are noted where appropriate. These differences are largely due to absences across multiple data collection administrations (administration of all instruments took place across six days in total; see Procedure) and omission of demographic control variables (specially, highest parental education).

Six urban middle schools in Western Pennsylvania and five urban middle schools from the Northwest region of United States participated. The schools were recruited to represent a diverse range of types of science learning and socio-economic environments. Specifically, public-available records for these schools show there was a wide range across these schools in the proportion of students eligible for free/reduced lunch (24–92% receiving free/reduced lunch; M = 56%, SD = 24%) or from under-represented minorities (36–99% minority population; M = 56%, SD = 22%). The overall sample was evenly split on gender (50% female) with the following breakdown of ethnicity based on those students providing this information: 44% Caucasian, 29% African-American, 18% Hispanic/Latino, 10% Asian, 7% Native American/Pacific Islander, and 6% Indian/Middle-Eastern. These gender and ethnicity variables were described as part of the study (see Demographic control variables below).

In terms of the type of science learning, schools varied in the extent to which they used a hands-on inquiry science curriculum or a textbook-focused science curriculum. Since teachers also have control over the use of learning resources, the schools varied in a more continuous way (based on teacher-log self-report) from primarily using hands-on inquiry to primarily using textbook-focused science learning. Also, the particular science content being learned varied across schools, although the sixth graders most commonly studied topics related to weather or Earth Science and the eighth graders more commonly studied topics related to Biology or Ecology.

These schools were recruited by contacting the sixth- and eighth-grade science teachers, who were compensated based on a number of participating classes; almost all science teachers in these schools participated.

Measures

Instrument development

These scales were all created as part of the Activation Lab project (http://activationlab.org/). The purpose of the project required revision and creation of instruments that met our needs. More specifically, where we did not use existing tools, it was because: (1) we are aimed at a broad range of middle school students and therefore aimed our items for a third-grade reading level (many existing measures are developed for and used in high school and college level), (2) we wanted science-specific measures for our motivation measures that reflect each construct (many existing measures were developed for different content), and (3) our engagement and perceived success measures were carefully designed to reflect a range of experiences and activities occurring across classes while still representing each of the constructs.

All measures (except the basic demographic variables) were developed through literature review and the use of student input. Specifically, cognitive interviews were conducted in which middle school students met 1:1 with a trained researcher (each scale was reviewed by three to six students using this process). Cognitive interviewing is a validation procedure that allows researchers to vet their scales by having a small sample of their target population, in this case, middle school students complete their items while discussing their thought process (Desimone & Le Floch, 2004; Ouimet, Bunnage, Carini, Kuh, & Kennedy, 2004; Willis, 2004). Specifically, this procedure ensures that (1) students are able to read and interpret the item, (2) and that this interpretation is representative of the theory behind that item. Students were asked to read the item, reword it in their own words, respond to the item, and then provide reasons for their answer. Responses were audio-recorded and then carefully analysed for the match to the researcher intentions of each item. Using this method, we were able to validate that the students' perceptions corresponded with the construct being measured and, when necessary, make edits to item wording.

In addition to establishing content validity, statistical validity was also assessed, and all scales were iteratively improved through principal components factor analyses (exploratory and confirmatory) and item response theory analyses to insure a single-factor structure for each construct, adequate discrimination across the scale, and no differential discriminability by gender, age, or ethnicity (see http://activationlab.org/tools for detailed technical reports).

Perceived learning experiences: engagement

Conceptually, engagement reflects the degree of positive vs. negative behavioural, cognitive, and affective participation in a science learning activity. It involves the experience itself and reflects influences of learner characteristics (e.g. abilities and attitudes), the activity (e.g. its difficulty and novelty), and various contextual aspects of the activity (e.g. interactions with other learners or support from adults). We developed these scales by reviewing a number of existing engagement surveys (such as those reviewed by Fredricks et al., 2011) and, in addition to adapting them to be: (1) science specific, (2) at an appropriate reading level for lower ability middle school students, (3) clearly related to a specific form of engagement (affect, behavioural, or cognitive) within a particular experience, and (4) relating to a broad number of possible science experiences.

Table 1. Engagement items.

	Empirical subcategory	Item	Loadings by day
E1	Affect	I felt bored (*r*)	0.81, 0.81
E2	Affect	I felt happy	0.87, 0.87
E3	Affect	I felt excited	0.86, 0.86
E4	Affect	Time went by quickly	0.65, 0.61
E5	Behavioural-cognitive	I was daydreaming a lot (*r*)	0.77, 0.75
E6	Behavioural-cognitive	I was focused on the things we were learning most of the time	0.70, 0.68
E7	Behavioural-cognitive	I was busy doing other tasks (*r*)	0.73, 0.74
E8	Behavioural-cognitive	I talked to others about stuff not related to what we were learning (*r*)	0.75, 0.74

Notes: The engagement and perceived success items all shared the prompt: 'During this activity.' Due to scheduling logistics, two teachers did not participate in the second engagement administration. Sample size each day is between 2234 and 2563. Factor loadings are from the final EFA run for affective and behavioural-cognitive items, respectively.

An important feature of this measure, and a major reason for its development in place of the use of an existing measure, is the reference to a single activity a student just completed. Many scales focus on a larger or more general scope (e.g. academic or school engagement without a set timeframe) (e.g. High School Survey of Student Engagement, Motivation and Engagement Scale; School Engagement Measure). However, students experience a range of activities within a given context, each of which may be differentially engaging. A scale is needed that is broad enough to capture variation in the feelings, behaviours, and thoughts across a variety of activities while still representing specific aspects of engagement (see Wang & Degol, 2014 for a discussion on the multilevel conception of engagement).

Further, reflections about extended time periods are at risk of becoming measures of general attitudes and beliefs rather than direct summaries of experiences. Therefore, we designed a scale to be used following single activities. The eight-item engagement measure was developed through the use of cognitive interviews described above. Conceptually, the scale includes three affective, two behavioural, and three cognitive items (See Table 1 for items). Exploratory factor analyses (EFAs) below describe the properties of the empirically determined factors used in subsequent analyses. To foreshadow the results, these items empirically separated into two different dimensions (affect and behavioural/cognitive) and Table 1 reflects the final empirical distinction among these items (see the Instrument Structure of Engagement and Perceived Success analyses in the Results for more details).

Engagement can be measured through behavioural observation or self-report survey. For our purposes, we selected a self-report measure of engagement; a common approach in the engagement literature (e.g. Fredricks et al., 2011; Wang & Eccles, 2012). We preferred this approach over observational measures for three reasons. First, with observational measures, each child would need to be observed, prohibiting large-scale work that teases apart correlational variables as measured here. Second, observational measures are often coarsely assessed (i.e. does not assess mental work) or too activity-specific (e.g. captures only certain kinds of behaviours) for the breadth of experiences across the course of a semester. Finally, the current work is theoretically focused on what students perceive their experience to be rather than what external observers perceive they are taking away, in terms of drivers of motivational change.

Perceived learning experiences: perceived success

Conceptually, perceived success captures the learner's sense of success in a particular science learning activity. Like engagement, the construct is focused on the experience

Table 2. Perceived success items.

	Empirical subcategory	Item	Loadings by day
PS1	Absolute	I did a good job	0.77, 0.77
PS2	Absolute	It was easy for me	0.72, 0.72
PS3	Absolute	I felt I was very successful	0.80, 0.79
PS4	Absolute	I did everything well	0.70, 0.70
PS5	Relative	I did a better job than the others	0.64, 0.67
PS6	Relative	I was more successful than everyone else	0.78, 0.80

Note: Factor loadings are from the final EFA run for perceived success items.

itself, with possible contributions from the learner, the activity, and the contextual aspects of the activity. Perceived success items were also developed using cognitive interviews and pilot data collection. The final scale consists of six items (See Table 2) that ask about students' beliefs in how well they did on an activity they just completed. Since learners may be influenced by how they perform relative to their peers as well as how they perform relative to their own prior achievement (e.g. Elliot & McGregor, 2001; Senko, 2016), our measure taps into both relative and absolute perceptions of success. Four items asked about students' *absolute* perceived success; that is, whether they felt successful by their own standards (e.g. I felt I was very successful). Two items asked about students' perceptions of *relative* success; that is, the perception of success in relationship to peers' performance (e.g. I was more successful than everyone else). EFAs below describe the psychometric properties of this measure.

Motivational variables

Each of our three motivational variables (fascination, values, competency beliefs) described below have undergone extensive empirical validation (using cognitive interviews, exploratory and confirmatory factor analyses, as well as item response theory analyses) and each scale has been shown to make up its own single-factor structure (e.g. single eigenvalue > 1, comparative fit index & Tucker-Lewis index all >0.90) with good item fit across student ability/levels of motivation (e.g. expected a posteriori all greater than 0.83, Wright map showed good spread across ability, scale responses in correct ranked order), and good internal reliability (Cronbach's alphas all >0.8). Response options are purposely varied to encourage respondents to process each item carefully, and IRT analyses validate the treatment of the scales as interval scales (i.e. similar distances between lower and upper levels of each item). For specific item information please see the technical reports (http://www.activationlab.org/tools/) and published works (Bathgate & Schunn, 2016; Bathgate & Schunn,2017; Dorph, Cannady, & Schunn,2016). Conceptually, the scales represent more stable self-characterisations of the learner across time and place, in contrast to engagement and perceived success, which represent subjective experiences during a specific activity in a specific moment in time. Each of the three motivation scales consists of eight items on a four-point Likert scale, each of which is averaged into a mean score in our analyses.

The *Fascination* measure (eight items) captures the intrinsic attachment to science content and activities (Ames, 1992; Elliot & McGregor, 2001; Hidi & Renninger, 2006; Jirout & Klahr 2012; Loewenstein, 1994; Ryan & Deci, 2000; Wigfield & Eccles, 2000). Conceptually, it involves subdimensions of emotional attachment (e.g. In general, when I work on science I: love it, like it, do not like it, hate it), mastery goals (e.g. I want to know everything about science: YES!, yes, no, NO!), and persistent curiosity (e.g. After a really interesting science experience is over, I look for more information about it: YES!, yes, no, NO!).

The *Values* measure (eight items) captures the extrinsic drive towards science learning (Ryan & Deci, 2000; Wigfield & Eccles, 2000). Conceptually, it includes valuing science for personal benefits (e.g. Knowing science helps me understand how the world works: All the time, most of the time, sometimes, never) and for the benefit of society (e.g. Science makes the world a better place to live: YES!, yes, no, NO!).

Finally, the *Competency Belief* measure (eight items) captures student expectations for successful participation in diverse forms of science learning (Bandura, 1993, 1997; Beghetto, 2007; Britner & Pajares, 2006). It includes students' beliefs about their ability to do well on both in and out-of-school-specific science activities (e.g. I can do the activities I get in class: all the time, most of the time, half the time, rarely; If I went to a science museum, I could figure out what is being shown in: all areas, most areas, a few areas, none of it) and perceived mastery of skills involved in completing science learning activities (e.g. I think I'm very good at coming up with questions about science).

Science content knowledge

To match the experimental context, our content knowledge measures were developed to test content taught over the course of the semester. Because different teachers covered different content (especially across grades and regions, but also somewhat across schools within regions), items were selected to correspond to the content taught by the teacher. These measures were developed by selecting items from released state test items and research-based item banks (e.g. American Association for the Advancement of Science, Trends in International Mathematics and Science Study, Misconceptions-Oriented Standards-Based Assessment Resources for Teachers) and were all multiple-choice items measuring students' learning of facts and conceptual knowledge from the content covered in their science classes for that semester. *Z*-scores are used to address differential difficulty across test forms.

Demographic control variables

Students were asked their *gender* and *ethnicity* (asking participants to check all that apply from a longer list). Ethnicity was then recoded to a binary variable (minority, non-minority), with minority being coded if any of the checked options included one of the traditionally under-represented minorities in the STEM (i.e. all but Caucasian and Asian). A binary variable of *grade* (sixth/eighth) was also recorded.

We also include two variables relating to students' home experiences that are expected to influence results in a meaningful way: *Home resources* and *highest parental education*. Since home environments vary in their access to learning materials (e.g. dictionaries, science books) and parental education has a long-established relationship to achievement outcomes, both are included in our analyses. However, since these variables are relatively stable and not as readily influenced by classroom learning experiences, we include them as control variables.

Home resources were measured via survey by asking students to select the frequency of availability of seven resources located in their home (e.g. Are these things available for use in your home? Study or homework area: always, most of the time, rarely, never). The scale has an acceptable Cronbach's alpha of 0.73. Highest parental education was collected by asking students each of their parent's highest education history (the coding and options were as follows: 1 = did not graduate from high school, 2 = graduated from high school, 3 = went to college but did not graduate, 4 = graduated from college, 5 = went to more school after college [master's degree, Ph.D., M.D., etc.]). We then selected the highest education of either parent to use in our analyses. For example, if one parent graduated high school (2) and the other graduated college (4), the coding for the parent graduating college would be used.

Procedure

All measures were collected during the students' science class at various time points during the year (as described below; see Figure 2) using paper surveys distributed by the researchers, with students bubbling responses in pencil directly onto the surveys. Students were told their individual responses would not be shown to their teacher and would not affect their grade.

Motivational variables

Fascination, values, and competency belief measures were collected at the start of the fall semester to control for students' starting levels of motivation (time-1) and were collected again at the end of the fall semester (time-2), approximately four months later.

Science content knowledge

Students were given the content knowledge test at the start of the school year. This initial knowledge measure was used to control for any entering knowledge from prior classwork and informal learning sources. Students were then given the same measure at the close of the semester to capture their content knowledge over the course of the semester. As a reminder, the included items varied across teachers because different teachers taught very different content. Test-specific *Z*-scores were used to equate difficulty across the different test items included on each test form.

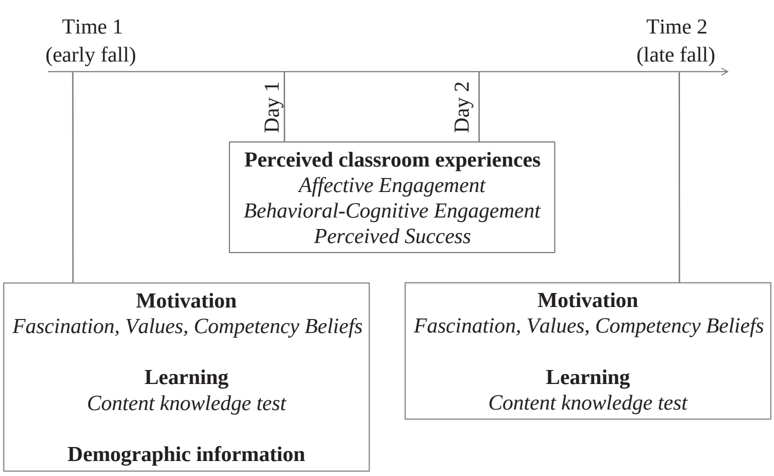

Figure 2. Procedure for data collection.

Demographic variables

Gender and ethnicity were collected at the beginning of the school year, but following the administration of the motivational surveys to avoid any stereotype threat on these other measures that may occur by first answering demographic questions that invoke identities with negative science stereotypes (i.e. female or under-represented minorities).

Engagement

Students completed both the engagement and perceived success measure immediately following two different science lessons separated from each other by at least a month but also separated by at least two weeks from the time-1 and time-2 data collections. A subset of the participants completed these measures on four different days to test the generalisability of the estimates based on only two days. The sampled days purposely avoided testing days and focused on typical class activities, but were also influenced by complexities of scheduling so many classes for data collection. Particular activities varied greatly and included both hands-on and lecture structures as well as variation in teacher-directed versus student- or group-driven work. The purpose of such variety is to improve the generalisability of the findings and to insure variation in engagement and perceived success.

Data analysis

We first briefly describe EFAs used to test the structure of the engagement and perceived success scales. Next, we explore the changes in students' means for our outcomes and predictors across time-1 and time-2, as well as describe the correlations among our variables. We subsequently answer each of our RQs using mixed-level modelling. This approach is selected to account for the nested structure of the data (i.e. students are nested within schools/teachers) that may vary systematically. Details for these analyses are included in each subsection for clarity. Additionally, these analyses are not intended to make strong statements of causality, being correlational in nature. Our predictor variables come temporally prior to our outcome variables, but are not tested in a way to confirm causal direction.

Results

Instrument structure for engagement and perceived success

Given our central focus on perceived classroom learning experiences, we examined the structure of the engagement and perceived success measures to explore dimensionality within each and discriminant validity between them. Since the two different days of administration came from the same range of science activity types, splitting the results this way represents a simple replication test of the analysis patterns as students become familiar with the scale.

EFAs with a Varimax rotation were run on data from each day of administration independently to ask whether engagement is separable from perceptions of success. In other words, are students able to separate engagement from whether or not they perceive themselves successful (and vice versa)? This analysis also provides evidence of the dimensionality of engagement; namely, whether the three forms of engagement are empirically distinct in this kind of learning context.

The EFAs yields a three-factor solution (based on Eigen-values above 1) across both days of administration. Most items load on only one factor at above 0.3, and no items load on a secondary factor at or above 0.4. All but one item (discussed below) load on the primary factor at above 0.6. All perceived success variables – and no engagement variables – load on factor 1 consistently (36% and 35% of the variance were explained by day, respectively).

Factor 2 includes affect items E1–E4. While E4 was initially conceptualised by the research team as a cognitive item, students responded to it more like an affective engagement item. In retrospect, E4 could be conceptualised as one's sense of flow, which has been previously related to the emotional experience of an activity (Csikszentmihalyi, 1990; Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003).

The third factor includes both behavioural and cognitive engagement items together (E5–E8). This same 2-factor structure of engagement was replicated through an EFA with perceived success items removed. The two cognitive variables (E5 and E6) have some evidence of double-loading across factors, suggesting that cognitive engagement may regularly co-occur to some degree with affect, at least as measured here in this context. Nonetheless, at both time points, each variable fits best on the behavioural–cognitive factor, and thus was kept within its *a priori* conceptual category. Applying EFAs to items within each engagement factor to insure no further divisions existed, a single underlying factor was found to provide a good fit for each engagement factor across the two days of assessment (see Table 1 for final factor loadings). The single-factor structure for each construct explained a good amount of variance (affective engagement: Day 1 = 64%, Day 2 = 63%; behavioural–cognitive engagement: Day 1 = 54%, Day 2 = 53%).

Additionally, perceived success items were included in a follow-up EFA with engagement items removed and showed a single-factor solution with factor loadings from 0.64 to 0.78 for the first administration and a two-factor solution for the second administration (absolute items on one dimension and relative items on the second dimension). However, when items from the second administration were constrained to a single-factor EFA, items were appropriately fitting, with loadings from 0.67 to 0.80. This final single-factor structure EFA for perceived success items explained 54% and 55% of the variance, respectively, across days. Table 2 shows the final factor loadings.

Overall, the EFAs revealed a consistent and coherent set of results: one perceived success factor and two engagement factors (affective and behavioural–cognitive). While behavioural and cognitive engagement is conceptually distinct, students' responses to behavioural and cognitive items are sufficiently correlated to load on a single factor, making them not distinct from each other empirically, at least not in the kinds of in-class science learning activities that were examined in this study. Based on these results, we move forward with a one-factor characterisation of perceived success (as an average of those items) and a two-factor characterisation of engagement (affective engagement as the average of items E1–E4 and behavioural–cognitive engagement as the average of E5–E8).

Overall variable means and reliabilities for engagement (affective and behavioural–cognitive) and perceived success were similar across the two days (shown in Table 3). Cronbach alphas for each were also good across both days (ranging from 0.71 to 0.83). A

Table 3. Means, standard deviations, and Cronbach alphas for predictor and outcome variables across administrations.

	Day 1			Day 2			Combined		
	M	SD	*α*	*M*	SD	*α*	*M*	SD	*α*
Affect Eng.	2.7	0.72	0.80	2.7	0.70	0.79	2.7	0.66	0.84
Beh-Cog Eng.	3.0	0.64	0.72	3.0	0.61	0.71	3.0	0.58	0.80
Perc. Success	2.9	0.57	0.83	2.9	0.58	0.83	2.9	0.53	0.86

Note: Day 1 $N = \sim 2700$; Day 2 $N = \sim 2600$; Combined = ∼3000.

combined average across the two days of administration was computed for affective engagement, behavioural–cognitive engagement, and perceived success, respectively, as a (noisy) estimate of typical science classroom experiences over the semester for each student for use in the regression results. The correlations in scores between days are moderate ($r = 0.52$, 0.56, and 0.49 for affective engagement, behavioural-cognitive engagement, and perceived success, respectively). Further, for the subset of participants who completed four experience surveys, engagement levels between this two-day mean and means based on four days was $r = 0.90$, 0.91, and 0.88 for affective engagement, cognitive–behavioural engagement, and perceived success, respectively. These results support the use of the mean across two days as an estimate of typical engagement and perceived success levels for each learner over the semester.

Overall changes in the motivational and content knowledge variables

At the mean level, students are moderately motivated towards science at both time-1 and time-2 measures, with means around 2.7 on the 4-point scales (See Table 4). There are also statistically significant differences from time-1 to time-2 testing across all three types of motivation (Fascination: $t(2{,}356) = -10.18$, $p < .001$, Cohen's $d = 0.16$; Values: $t(2{,}353) = -2.26$, $p = .024$, Cohen's $d = 0.06$; Competency Beliefs: $t(2{,}340) = 3.80$, $p < .001$, Cohen's $d = 0.06$). However, practically speaking, these mean differences are very small, nudging the mean score approximately 0.1 on a four-point scale with negligible effect sizes. Furthermore, while the average shift is small, there are significant individual variations occurring with some students showing meaningful growth and other showing meaningful declines. Table 4 shows the interquartile range of students' time 1–time 2 changes in each type of motivation, revealing important variations of change occurring across students. Many students are declining by over 0.5 standard deviations on each scale in just

Table 4. Descriptive data and change data for motivation and learning outcomes.

	Time 1			Time 2			Change
	M	SD	*α*	*M*	SD	*α*	Interquartile range
Motivation outcomes							
Fascination	2.7	0.6	0.86	2.6	0.6	0.86	[−0.4, 0.3]
Values	2.6	0.5	0.83	2.6	0.5	0.83	[−0.3, 0.3]
Competency belief	2.8	0.6	0.84	2.8	0.6	0.83	[−0.3, 0.3]
Learning outcome							
Content knowledge	7.3	3.2	0.60[a]	49.1	20.8	0.70[a]	[28.9–56.7]

Note: Based on students with both time-1 and time-2 data ($N = \sim 2300$).
[a]A weighted theta was used as a measure of reliability for content knowledge, as it is a better measure for dichotomous scales and can be weighted for the sample size associated with each content test.

four months; at the same time, many other students are increasing by over 0.5 standard deviation units in that same short period.

Therefore, the purpose of our primary analyses for the first research question is not to examine what leads to the small mean decline in motivation (or, in the case of competency belief, a small, negligible mean increase), but rather to examine whether differences in engagement and perceived success during science class account for some of this large individual variation in motivational changes. For example, does student engagement and perceived success in class predict whether some students grow and other students decline in motivation?

Gains in content knowledge were strong between the two time points ($t(2{,}348) = 108.31$, $p < .001$, Cohen's $d = 2.75$), as would be expected from a test tied to the content being covered over the course of the semester. However, there was a large range in the amount different students learned, as indicated by the interquartile range. Our second research question is aimed at understanding what may account for some of this difference in relative amounts of learning.

Before addressing our central RQs, we first examined the correlation structure among all our predictor variables to show first-order correlational relationships between predictors and outcomes, and also to screen for potential multicollinearity problems in the predictors or redundancies in the outcomes. In terms of the relationships among the predictor variables (Table 5, upper-left grey box), the correlations among engagement (both affective and behavioural–cognitive) and perceived success are moderate and show discriminant validity. This relationship is theoretically expected (i.e. low perceived success can come from a low behavioural–cognitive engagement or cause low affective engagement). However, the correlations are sufficiently low that the multiple regressions should not suffer from severe multicollinearity (additionally, VIF statistics in the models were all well within acceptable limits, i.e. below 2.0). At the same time, the predictors are sufficiently correlated that first-order correlations between predictors and outcomes may simply reflect indirect rather than direct connections.

Moving to the relationship among the outcome variables (lower-right grey box), we see the strongest correlations are among fascination, values, and competency beliefs. The strength and direction of these relationships are typical of previous motivational research examining these variables. Despite the moderate to moderately high correlation, there remains enough variation for each outcome to be potentially driven by a unique set of factors (or, perhaps, driven by the same factors to varying degrees). Prior knowledge of the class content has relatively low overlap with the motivational variables.

Table 5. Pearson correlations among predictor variables, among outcomes variables, and between predictors and outcomes.

	Beh-Cog. Eng.	Perc. Succ.	Fasc.	Values	Comp. Bel.	Time-2 Content Knowl.
Affective Eng	0.54***	0.41***	0.42***	0.33***	0.27***	0.08***
Behavioural-cognitive Eng		0.40***	0.26***	0.22***	0.24***	0.13***
Perceived success			0.21***	0.20***	0.33***	0.10***
Fascination				0.70***	0.56***	0.15***
Values					0.55***	0.20***
Competence beliefs						0.27***

***$p \le .001$.

Finally, the relationship among the predictor variables and the outcome variables (upper-right white box) tend to be low to moderate and have varying strength depending on the particular predictor-outcome pairings. For example, fascination is most strongly associated with affective engagement and has a lower and roughly equal relationship with behavioural–cognitive engagement and perceived success. Values show a similar pattern in relationship to the predictors as fascination, but with a slightly lower correlation with affective engagement. Competency beliefs are most strongly associated with perceived success, and content knowledge is roughly equally related with each predictor, showing the lowest correlations. Multiple regressions are needed to examine whether indeed every predictor is actually associated with every outcome, or whether many of the connections are actually mediated.

RQ 1: Which aspects of motivational change are associated with perceived classroom experiences of engagement and perceived success?

Changes in fascination

Three linear mixed models were run with time-2 fascination average as the outcome (see Table 6) using RStudio software. First, a fully unconditional model was run to assess the amount of school-level (level 2) variance in fascination. The second model (Baseline Control) adds in first-level fixed effects variables and includes time-2 fascination averages to control for students' initial motivational levels and demographic control variables (gender [male/female], ethnicity [minority/non-minority], grade [sixth/eighth]). The third model (final) adds affective engagement, behavioural–cognitive engagement, and perceived success first-level fixed effects variables. It is important to note that over this four-month period, fascination levels (and indeed all three examined motivations) are relatively stable, even though some students are showing significant growth or decline. Thus, the various measured experiences can only have moderate predictive power for time-1 fascination.

Table 6. Fixed variable standardised coefficients in models predicting time-2 fascination.

	Model 1, unconditional	Model 2, baseline control	Model 3, final
(Intercept, unstandardised)	2.57***	2.48***	2.46***
Level-2 control variables			
Time 1-Fascination		0.62***	0.54***
Male		0.05**	0.05**
Minority		−0.01	<0.01
Eighth grade		0.08***	0.09***
Home resources		0.05*	0.03
Highest parent education		0.03	0.02
Level-2 perceived classroom experiences			
Affective Eng.			0.19***
Behavioural-cognitive Eng.			<0.01
Perceived success			0.07***
R^2	0.01	0.41	0.46
AIC	4671.31	2065.42	1900.77

Note: $N_{school} = 11$, $N_{student}$ for fully unconditional model: 2676, $N_{student}$ for baseline control model = 1673, $N_{student}$ for final model = 1635.
***$p \leq .001$.
**$p \leq .01$.
*$p \leq .05$.

The fully unconditional model shows that the random effects of school do account for significant amount of variance in time-2 fascination (i.e. there is some systematic variation in students' fascination by school) indicating the need to use mixed-linear modelling for this data: X^2 (1, $N = 2676$) = 7.17, $p = .007$. However, this variance is considerably small at 0.01%. Model 2 and Model 3 consistently improve in fit (see R^2 and Akaike information criteria [AIC] indices in Table 6) as our additional variables are added.

Affective engagement is consistently and most strongly associated with changes in fasciation, followed by perceived success. Behavioural–cognitive engagement shows no relationship with changes in fascination.

Control variables are also associated with changes. Specifically, males and eighth graders each show greater increases in fascination compared to females and sixth graders, respectively. Home resources also show an effect in the baseline model, but this effect becomes non-significant once the perceived classroom experiences are accounted for.

Changes in values

The same three models previously described (fully unconditional, baseline, final) were run with the time-2 values variable as the outcome. Similar to the fascination outcome, school-level variance in the fully unconditional model accounts for only 1% of the variance. This effect is significant for the fully unconditional model (X^2 [1, $N = 2673$] = 6.51, $p = .01$), but becomes non-significant after including the control variables in Models 2 and 3. Models 2 and 3 consistently improve the model fit (see R^2 and AIC indices in Table 7) as our level-two variables are added. Looking at the standardised coefficients in the final model, affective engagement is most predictive of changes in values, followed by perceived success, but no relationship with behavioural–cognitive engagement.

Gender and grade are also predictive of changes in values, with a similar pattern of that seen in fascination (i.e. males and eighth graders each show increases in values compared to females and sixth graders, respectively). Home resources and highest parental education

Table 7. Fixed variable standardised coefficients associated with changes in values.

	Model 1, unconditional	Model 2, baseline control	Model 3, final
(Intercept, unstandardised)	2.62***	2.57***	2.55***
Level-2 control variables			
Time 1-values		0.58***	0.52***
Male		0.08***	0.07***
Minority		−0.01	−0.01
Grade		0.04*	0.07***
Home resources		0.05**	0.04^{+}
Highest parent education		0.06**	0.06**
Level-2 perceived classroom experiences			
Affective Eng.			0.12***
Behavioural-cognitive Eng.		0.04	
Perceived success			0.07***
R^2	0.01	0.37	0.40
AIC	4165.06	1830.44	1733.18

Note: $N_{school} = 11$, $N_{school} = 11$, $N_{student}$ for fully unconditional model: 2673, $N_{student}$ for baseline control model = 1673; $N_{student}$ for final model = 1635.
***$p \leq .001$.
**$p \leq .01$.
*$p \leq .05$.
$^{+}p \leq .10$.

also show a relationship. That is, increases in each are associated with significant increases in values. However, home resources have only a marginal relationship once perceived classroom experiences are included in Model 3.

Changes in competency belief

The same three models were run with time-2 competency beliefs as the outcome. The fully unconditional model shows school association accounts for 6% of the variance (X^2 [1, N = 2662] = 112, $p < .001$), but this effect becomes non-significant with the inclusion of control variables in Models 2 and 3. Model 2 and Model 3 consistently improve the model fit (see R^2 and AIC indices in Table 8) as our level-two variables are added. Looking at the standardised coefficients in the final model in Table 8, perceived success is the strongest predictor of changes in competency beliefs, with some additional role of affective engagement. Behavioural–cognitive engagement shows no relationship to change.

All control variables are associated with changes in competency beliefs. Similar to values, males and eighth graders each show a positive gain compared to their counterparts. Home resources and highest parental education have similarly sized positive relationships with time-2 competency beliefs. However, for competency belief changes, minority status is also a predictive of change: minority students show a larger decrease in competency beliefs compared to non-minority peers.

RQ 2: Are engagement and perceived success associated with content knowledge?

Using the same three models run with time-2 test as the outcome, the fully unconditional model shows school association accounting for about 24% of the variance. (X^2 [1, N = 2673] = 549, $p < .001$), indicating the need for mixed-level modelling. However, this contribution was reduced to 13% once background control variables were included in the model. Model 2 improves the model fit, as does Model 3, although only slightly (see R^2 and AIC indices in Table 9). Table 9 shows the standardised coefficients for the baseline

Table 8. Fixed variable standardised coefficients associated with changes in competency belief.

	Model 1, unconditional	Model 2, baseline control	Model 3, final
(Intercept, unstandardised)	2.83***	2.84***	2.83***
Level-2 control variables			
Time 1-competency beliefs		0.58***	0.50***
Male		0.05**	0.05**
Minority		−0.05*	−0.05**
Grade		0.05*	0.05**
home resources		0.09***	0.08***
Highest parent education		0.08***	0.08***
Level-2 perceived classroom experiences			
Affective Eng.			0.06**
Behavioural-cognitive Eng.		0.02	
Perceived success			0.19***
R^2	0.06	0.44	0.49
AIC	4287.72	1742.52	1599.13

Note: N_{school} = 11; $N_{student}$ for fully unconditional model: 2662, $N_{student}$ for baseline control model = 1667, $N_{student}$ for final model = 1629.
***$p \leq .001$.
**$p \leq .01$.
*$p \leq .05$.

Table 9. Standardised coefficients associated with changes in content knowledge.

	Model 1, unconditional	Model 2, baseline control	Model 3, final
(Intercept, unstandardised)	49.33***	51.91***	51.91***
Control variables			
Time 1-test score		0.47***	0.47***
Male		−0.01	−0.01
Minority		−0.09***	−0.09***
Grade		0.04*	0.04*
Home resources		0.05*	0.04*
Highest parent education		0.12***	0.12***
Perceived classroom experiences			
Affective Eng.			−0.05*
Behavioural-cognitive Eng.		0.05*	
Perceived success			0.03
R^2	0.24	0.44	0.44
AIC	23,240.28	14,413.65	14,130.67

Note: $N_{school} = 11$, $N_{student}$ for fully unconditional model: 2669, $N_{student}$ for background model = 1731, $N_{student}$ for final model = 1697.
***$p \leq .001$.
**$p \leq .01$.
*$p \leq .05$.

and final model. Again, the baseline model consists of time-1 test content scores and demographics. The full model adds engagement and perceived success variables.

The pattern of results in this final model is markedly different than the previous three motivational outcomes. Here, only behavioural–cognitive engagement is positively associated with time-2 test scores. Perceived success shows no significant relationship and affective engagement shows a significant negative relationship with time-2 test scores. This change in effect demonstrates the importance of considering related experiential variables simultaneously.

With the exception of gender, all control variables show some relationship with content knowledge. Minority status is negatively associated with learning gains whereas the remaining variables are positively associated. Parental education has the largest relationship of any variable in the model (with the exception of the time-1 test scores) and this relationship is maintained even once perceived classroom experiences are included in the model.

Summary of results across motivation and content knowledge outcomes

To understand the patterns across our four outcomes, Table 10 shows the strength of the independent contributions each variable had to each respective outcome (based on the final model of each analysis). Figure 3 shows the empirical connections from this work. Thicker lines in the figure indicate stronger connections and the dotted line indicates a negative connection.

Most saliently, changes in motivational variables are predicted by affective engagement and perceived success but not behavioural–cognitive engagement. That is, the emotional experience one has with classroom science activities and one's perception of how well they completed those activities contributes to growth/decline in one's motivation towards science. The largest relationship is found between affective engagement and fascination. Similar (and relatively small) relationships are found among perceived success for both fascination and values. However, competency beliefs are most strongly predicted by

Table 10. Summary of effects across four outcome variables based on final model.

	Fascination	Values	Competency belief	Content knowledge
Control variables				
Male	+	+	+	ns
Underrepresented minority	ns	ns	+	−
Grade	+	+	+	ns
Home resources	ns	+	+	+
Highest parent Ed.	ns	+	+	+
Experience variables				
Affect Eng.	++	+	+	−
Beh-Cog Eng.	ns	ns	ns	+
Perceived success	+	+	++	ns

Note: '++' = Strong positive effect (significant coefficient > 0.15), '+'=positive effect (significant coefficient < 0.15), '−' = negative effect (significant coefficient < 0.15), ns = not statistically significant.

perceived success, much more strongly than perceived success predicts fascination or values. In other words, how behaviourally cognitively engaged a student is does not change their beliefs about their ability in science and even students' affective engagement has a comparatively small impact on competency beliefs.

For science content knowledge, both behavioural–cognitive and affective engagement predicted content knowledge growth in opposing directions (controlling for time-1 test scores) and perceived success had no relationship. In other words, only productive behaviours and thought processes (behavioural–cognitive engagement) have a positive relationship with content knowledge and one's affective engagement in class is associated with decreases in content knowledge. There is no evidence here that the perceptions of success directly impact one's content knowledge.

The pattern of effects of the demographic control variables also adds depth to these results. Gender showed a consistent effect across all motivational outcomes with boys being associated with relatively higher motivation. But there were no gender differences in content knowledge. In other words, boys are more motivated in science, but perform

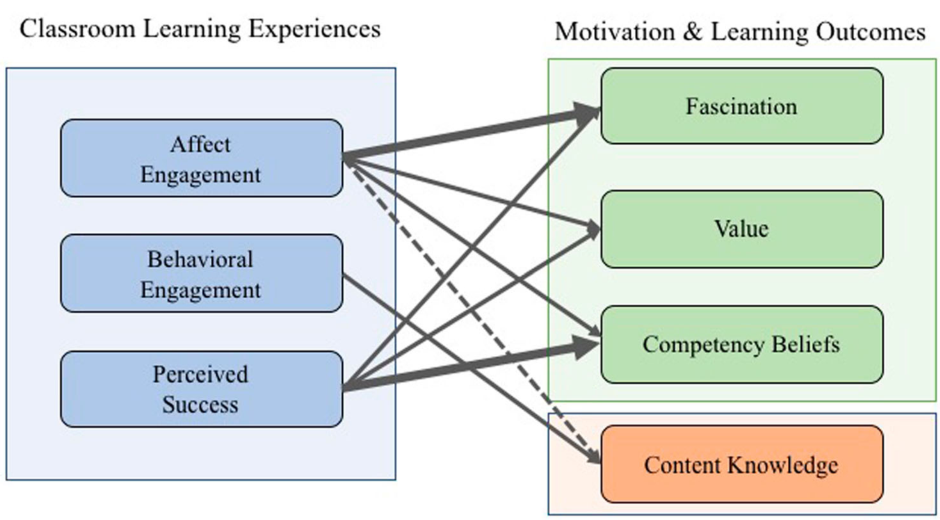

Figure 3. Final model showing the summary of empirical connections across analyses.

no differently than girls. By contrast, having a minority status was associated with reductions in both competency beliefs and amount of content knowledge learned. Eighth graders showed greater growth (or perhaps less loss) in motivation. Home resources showed a small relationship with all outcomes, except fascination. That is, increases in student access to productive home resources are associated with more values, competency beliefs, and content knowledge, even once initial scores are controlled for. However, access to resources at home does not have a relationship with students' fascination. Parental education also shows a relationship to all outcomes except fascination. However, parental education has relationships to competency beliefs and content knowledge that are worth noting. It is the strongest predictor to content knowledge (with the exception of time-1 test) and, in both cases, has a stronger relationship than minority and gender.

It is important to note that the demographic effects were almost entirely stable even when we added engagement and perceived success variables; that is, the demographic effects cannot be explained by the way students perceive their activity experiences during class. Other out-of-school factors, like informal experiences or societal stereotypes, are likely relevant. At the same time, this disconnect highlights that the association of engagement and perceived success with changes in motivation and content knowledge cannot be attributed to third variable correlations through these demographic variables.

Discussion

What are the perceived science classroom experiences that drive changes in motivation for science and science content learning? In simple correlational terms, all aspects of the experience are correlated with all changes, as was previously found in the literature, showing predictive and convergent validity (e.g. Ainley & Ainley, 2011; Connell & Wellborn, 1991; Fredricks et al., 2004; Pekrun & Linnenbrink-Garcia, 2012; Wang & Eccles, 2012; Wang & Holcombe, 2010). However, the current findings demonstrate the importance of considering related experiential variables simultaneously: very different features of the experiences are associated with motivational changes vs. science content knowledge, and different features matter for different motivational changes.

Before discussing each of the observed relationships in terms of prior findings and theories, we begin with a brief comment regarding causality. Clearly, as an observational study using regression techniques, no strong claims regarding causality can be made. However, the cases in which no significant association was found by the regression analyses do rule out some causal connections, or at least suggest they are at best quite small or limited to narrow contexts. Further, the regressions looked at change over time, and thus reverse causal relationships are ruled out (e.g. growth in fascination at the end of the semester cannot have caused high levels of affective engagement earlier in the semester). Finally, the use of multiple regressions that control for a number of plausible confounded factors does reduce concerns about associations caused by third variables.

Additionally, while our models explained a substantial amount of variance across outcomes (40–49%), there is also substantial unmeasured variance. A number of other sources likely also matter, such as the quality of teaching or activities that varied across classrooms, incomplete sampling of engagement across the semester, and individual reasoning ability unmeasured in the current study. Given the complex and broad range of factors that likely also matter, it is notable that the variables measured in this study were strong enough to account for such a large portion of motivational changes and content knowledge gains.

Experiences that change science motivation

Overall, the relative levels of affective engagement and perceived success are associated with changes in all three motivational variables; in no case did behavioural–cognitive engagement significantly account for any motivational outcome. Further, affective engagement is most strongly associated with changes in fascination, and perceived success with changes in competency beliefs.

The two strong connections – affective engagement with fascination and perceived success with competency beliefs – can be thought of in similar ways: as an internalisation or stabilisation of situational experience to a stable set of attitudes and beliefs. For example, the growth in fascination is consistent with the work on interest by Hidi and Renninger (2006), who theorised a developmental sequence from situational interest (more temporal, relying on environmental support) to individual interest (more stable, personally driven). Our work provides quantitative empirical support for their theory, which has previously been assessed primarily through qualitative data. Further, it goes beyond their work to show that the affective elements of the experience (rather than the cognitive or behavioural elements of the experience), along with perceptions of success, are what contribute to the development of a more stable individual interest.

Turning to the growth in competency beliefs, our work builds on seminal theorising by Bandura (1993, 1997), also showing that these beliefs build from attributions of success in various experiences. Interestingly, we also show that growth in competency beliefs is associated with affective engagement, but not behavioural–cognitive engagement. In other words, students become more confident in their competencies when the experience is pleasant rather than involving cognitive or behavioural effort. This may stem from differential attributions regarding the two kinds. On the one hand, the experience of high vs. low affective engagement might change the perceptions of effort required to complete the tasks (i.e. time going quickly with high affective engagement and time going slowly with low affective engagement), which then is internalised as signals of competence. On the other hand, the experience of cognitive–behavioural engagement may not indicate competence at all because off-task cognition and behaviour do not allow the learner to judge competence in the counterfactual case of them having attended.

Changes in values are not strongly associated with any particular experience measure, but two experiential factors are significant predictors. The observed association with perceived success is consistent with self-determination theory, whose supporting research finds that when learners feel more successful in domain activities, they tend to value that area more (Deci & Ryan, 2000; Ryan & Deci, 2000). Although the association of changes in values with affective engagement is theoretically consistent (e.g. Deci & Ryan, 2000; Renninger & Bachrach, 2015), our findings are unique in that we examine learners' affect during a particular science activity and their subsequent values of science (controlling for initial values), whereas many studies have focused on the opposite relationship (values predicting affective engagement), have measured values related to the activity context rather than the activity itself (e.g. value of peer support), or have subsumed value within emotional engagement (e.g. Finn, 1993). By examining affective engagement directly from science experiences and relating them to changes in values of science, we can better understand how learners' more fine-grained experiences relate to larger motivational shifts within the same content area.

Unlike affective engagement, behavioural–cognitive engagement shows no significant association with any motivational change. In other words, engaging (or not) in the thoughts and behaviours productive for science class activities are not associated with increases (or decreases) in motivation. For example, it is possible for a student to be behaviourally and cognitively engaged in class, but not have those experiences influence their overall science interest. This finding has been posited in the interest literature (Renninger & Bacrach, 2015) and may reflect the myriad of reasons one engages beyond the interest in a topic or activity (e.g. desire to get a good grade, performance goals). In other words, students may make other attributions than interest to the cause of their staying on-task. Further, the type of science learning activities may be important here; if the activities are highly scripted or involving very closed-ended tasks, students may not be given opportunities to deepen their interests from completing these activities (Chi, 2009; Marks, 2000).

Experiences that drive science learning

Content knowledge gains show a different pattern from the motivational variables. Unlike for the motivational shifts, learning is associated with behavioural–cognitive engagement. That is, the more behavioural and cognitive engaged a learner is in their science class, the greater their learning. This connection was expected, given that a learner needs to attend appropriately to the content being taught (e.g. think about that content, complete the activities around that content) to effectively learn it. Prior literature and review articles on behavioural and cognitive engagement have also demonstrated the importance of such on-task behaviours and effective attention and effort to academic achievement and learning (Connell & Wellborn, 1991; Finn, 1989, 1993; Fredricks et al., 2004; Marks, 2000; Wang & Eccles, 2012).

More interesting and novel is the finding that one's affective experiences in science classes are negatively associated with learning. Although this negative relationship is counterintuitive, there are multiple plausible explanations for this finding. For example, perhaps students who find the content more affectively engaging are less familiar with it, or high positive emotional experiences in class leads to students attending to the wrong information (i.e. superficial or only what they are most excited by). Our current data cannot directly explain these effects and further investigation and replication across more administrations are needed to better understand this relationship and why it occurs. However, it does appear that multicollinearity among our predictor variables (particularly between affective engagement and behavioural–cognitive engagement) is not an issue here. VIF statistics were all acceptable and the correlation between the two forms of engagement was only moderate. This finding emphasises the need for research using the finer-grained measurement of affective experiences to model their influence on motivations and academic achievement (Linnenbrink-Garcia & Pekrun, 2011; Wang & Degol, 2014).

Another interesting aspect of the current findings is the lack of relationship of perceived success in supporting content knowledge gains. That is, believing one's self as successful in science class activities did not result in higher learning. Explaining this effect is beyond the current data and may be due to multiple factors. Perhaps students in this age and context are relatively inaccurate at their reflections on how they performed, as they are still coming to understand the scope of what they know vs. what they do not know. Or perhaps there is misalignment between the activities completed in class and the content being tested in the exam. That is, if the activities the students did prior to the perceived success survey did not closely align with the content test, their perceptions of success may not be a good indicator of learning. Future research will have to examine this relationship in greater detail to examine what underlies it and, in turn, how to best address it.

Educational implications

The current study highlights the aspects of the experience that are most important to target in order to achieve particular changes. For example, if learning content is the primary goal, behavioural–cognitive engagement is the primary level to target. Both behavioural–cognitive engagement and perceived success can be supported by setting clear objectives with an explicit path as to how to reach them (e.g. Connell & Wellborn, 1991; Fredricks et al., 2004). Having tasks that require clear learning behaviours and directs attention to particular features associated with learning objectives can encourage learners' cognitive–behavioural engagement. Additionally, learners likely feel the most success when an activity meets – and slightly extends – their current ability (Vygotsky, 1980). Providing encouraging feedback and creating an environment in which learners feel safe to 'fail' and allowing them opportunities to practice new skills is also posited to benefit engagement.

Similarly, if specific motivational changes are the primary goal, then this research points to particular experience categories to target. For example, encouraging affective engagement can influence changes in fascination and values. Selecting activities that are more student-centred (e.g. hands-on, authentic science practices, group/student-directed) and help learners find the relevance of an activity to their own lives can generate more positive experiences, which influence more stable motivations, particularly in those learners with lowered expectations of success (Chin & Osborne, 2008; Hulleman & Harackiewicz, 2009; Marks, 2000; Renninger & Bachrach, 2015). However, these approaches are not an uncomplicated solution, as different approaches are most effective at different levels of initial motivation. For example, learners with lower interest are influenced by different factors (e.g. novelty) and may need more support to engage and see opportunities for engagement than learners with higher interest (Durik & Harackiewicz, 2007; Renninger, 2010; Renninger & Bacrach, 2015). Educators desiring to best meet the needs of their learners should consider the initial motivations and expectations of their learners and how their activity content and procedures could best connect to learners' lives.

The high co-occurrence of cognitive and behavioural engagement is also important. Teachers are given relatively little access to student cognition in the class overall. Here we find that, at least in middle school science, when students are behaviourally on-task there is a good chance they are also cognitively on-task as well. However, we do note that future research may need to consider cognitive engagement in a more fine-grained and task-specific way that the current survey instrument was able to do in order to better understand the generality of that behavioural-cognition co-occurrence (Greene, 2015).

Future directions

Engagement structure

We showed a clear separation between affective engagement and behavioural–cognitive engagement in our data set, as well as a lack of separation of behavioural and cognitive engagement. However, there are limitations to generalising this finding. First, we expect situational context to affect the degree of this separation. Our scale was designed to measure behavioural and cognitive engagement distinctly, but the activities we studied required behaviours to go along with cognition. Behavioural engagement within a typical school classroom may be particularly prescribed and connected to cognitive engagement. As such, the two-factor structure found here may not be found across all experiences. Perhaps activities that allow learners to engage in more diverse behavioural and cognitive ways may demonstrate a stronger split between behavioural and cognitive engagement. For example, a child in a science museum may interact with a particular display in different behavioural ways (e.g. playing with the objects in a display, reading the associated explanations provided by museum signs, speaking with a museum educator). These activities could relate to cognitive engagement in less predictable ways (e.g. the child is playing with the objects and making no connection to the larger content to be learned or perhaps they are tying it all together). Similarly, students completing the highly routinised process (e.g. data collection or entry) could be behaviourally engaged but not cognitively engaged. However, routinisation in the current classroom context is unlikely, so that when one is behavioural off task, cognition is likely to follow (and vice versa). Engagement measured across a greater variety of contexts is needed to better understand how these dimensions function under different constraints and opportunities (Azevedo, 2015; Sinatra, Heddy, & Lombardi, 2015).

Second, we did not systematically examine different classroom practices across teachers so we cannot say whether some practices are more engaging for students than others. The type of activity, degree and quality of support from the teacher and peers, and the coherence of activities across classes likely influences engagement. Our use of hierarchical linear modelling controlled for differences across teachers, but the specific classroom practices that influence differences in engagement and perceived success can be systematically explored in future studies.

Third, our measure of behavioural and cognitive engagement did not include all ways in which students might cognitively or behaviourally engage in a science learning tasks. For example, other conceptualisations of cognitive engagement include particular cognitive strategies or ways of thinking, such as metacognitive monitoring. Perhaps using this approach would provide greater distinction among behavioural and cognitive aspects. Similarly, behaviours strongly related to a specific task (e.g. sharing ideas with a peer during group work) could provide further distinction, but are more context dependent (e.g. talking with a friend could be part of engagement with the activity or not, depending on the context). In order to be generalisable across classroom science tasks, we focused on a few indicators of behavioural and cognitive engagement that largely measure amount of each, rather than particular strategies. Prior work has explored a much larger set of indicators (Bathgate et al., 2013). But, they also produced a single behavioural–cognitive engagement factor, and further their psychometric properties were not as robust as the items that were included in our final measure. Here a more focused study of cognitive and behavioural engagement in fixed tasks may be useful so that very detailed task-specific measures can be deployed.

Malleability

We have emphasised experiential variables that are malleable to some extent. Next steps should be taken to understand the character of the context in which these experiences occur and their relationship with engagement (Azevedo, 2015; Sinatra et al., 2015). In other words, what features of science learning experiences lead to greater affective and behavioural–cognitive engagement, respectively? Understanding these relationships will require both qualitative and quantitative work and produce practical application to educational practices (Renninger & Bacrach, 2015; Ryu & Lombardi, 2015).

Demographic effects

Although not the focus of this article, there are relationships among minority, gender, home resources, and parental education with motivational and learning outcomes. These patterns are reflected in the previous literature examining these effects (e.g. Bembenutty 2007; Jones, Howe, & Rua, 2000; Reilly, Neumann, & Andrews, 2015; Wang & Eccles, 2012) and deserve further investigation to understanding the cultural effects that are giving rise to these differences. However, it is notable that home resources and parental education had a larger role in competency belief outcomes than minority or gender, both of which have previous relationships to such outcomes. Additionally, highest parental education had the largest relationship with learning content knowledge (controlling for time-1 test) – higher than minority status and any of the classroom experience variables. This finding warrants additional exploration to examine the possible mechanisms for these effects.

Note

1. ‘Change’ can be interpreted in multiple ways and is often interpreted as a delta score (i.e. time-1 subtracted from time-2 scores). However, since delta scores often have statistical artifacts, we use time-2 test motivational scores as our outcome, controlling for students’ initial (time-1) motivation scores, to represent change in the current study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by National Science Foundation [DRL-1348468].

References

Ainley, M., & Ainley, J. (2011). Student engagement with science in early adolescence: The contribution of enjoyment to students' continuing interest in learning about science. *Contemporary Educational Psychology*, *36*, 4–12. doi:10.1016/j.cedpsych.2010.08.001

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*(3), 261. doi:10.1037/0022-0663.84.3.261

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, *50*(1), 84–94. doi:10.1080/00461520.2015.1004069

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*(2), 117–148. doi:10.1207/s15326985ep2802_3

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Macmillan.

Baram-Tsabari, A., & Yarden, A. (2008). Girls' biology, boys' physics: Evidence from free-choice science learning settings. *Research in Science & Technological Education*, *26*(1), 75–92. doi:10.1080/02635140701847538

Bathgate, M. E., Schunn, C. D., & Correnti, R. (2013). Children's motivation towards science across contexts, manner-of-interaction, and topic. *Science Education*, *98*(2), 189–215.

Bathgate, M. E., & Schunn, C. D. (2016). Disentangling intensity from breadth of science interest. *Instructional Science*, *44*(5), 423–440. doi:10.1007/s11251-016-9382-0

Bathgate, M. E., & Schunn, C. D. (2017). Exploring motivational shifts in middle school: What deepens science utility value and what attenuates its decline? *Contemporary Educational Psychology*, *49*, 215–225.

Bathgate, M. E., & Schunn, C. D. (in press). The psychological characteristics of experiences that influence science motivation and content knowledge. *International Journal of Science Education*. doi:10.1080/09500693.2017.1386807

Beghetto, R. (2007). Factors associated with middle and secondary students' perceived science competence. *Journal of Research in Science Teaching*, *44*(6), 800–814. doi:10.1002/tea.20166

Belenky, D. M., & Nokes-Malach, T. J. (2012). Motivation and transfer: The role of mastery-approach goals in preparation for future learning. *Journal of the Learning Sciences*, *21*(3), 399–432.

Bembenutty, H. (2007). Self-regulation of learning and academic delay of gratification: Gender and ethnic differences among college students. *Journal of Advanced Academics*, *18*(4), 586–616. doi:10.4219/jaa-2007-553

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, *43*(5), 485–499. 10.1002/tea.20131

Bryan, R. R., Glynn, S. M., & Kittleson, J. M. (2011). Motivation, achievement, and advanced placement intent of high school students learning science. *Science Education*, *95*(6), 1049–1065. doi:10.1002/sce.20462

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105. doi:10.1111/j.1756-8765.2008.01005.x

Chin, C., & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, *44*(1), 1–39. doi:10.1080/03057260701828101

Cho, M. H., & Summers, J. (2012). Factor validity of the motivated strategies for learning questionnaire (MSLQ) in asynchronous online learning environments. *Journal of Interactive Learning Research*, *23*(1), 5–28. doi:10.1177/0013164493053003024

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self-processes and development: Minnesota symposium on child psychology* (Vol. 23, pp. 43–77). Chicago, IL: University of Chicago Press.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper-Perennial.

Deci, E. L., & Ryan, R. M. (2000). *Intrinsic motivation and self-determination in human behavior*. New York,NY: Plenum.
Desimone, L. M., & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, *26*(1), 1–22.
Dorph, R., Cannady, M., & Schunn, C. D. (2016). How science learning activation enables success for youth in science learning. *Electronic Journal of Science Education*, *20*(8), 49–85.
Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How personal interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology*, *99*, 597–610. doi:10.1037/0022-0663.99.3.597 597
Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132. doi:10.1146/annurev.psych.53.100901.135153
Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, *34*, 169–189. doi:10.1037/0022-3514.70.3.461
Elliot, A., & McGregor, H. (2001). A 2 X 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*, 501–509.
Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, *59*, 117–142. doi:10.3102/00346543059002117
Finn, J. D. (1993). *School engagement and students at risk* (NCES 93470). National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubs93/93470.pdf
Fredricks, J., Blumenfeld, P., & Paris, A. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. doi:10.3102/00346543074001059
Fredricks, J., McColskey, W., Meli, J., Mordica, J., Montrosse, B., & Mooney, K. (2011). *Measuring student engagement in upper elementary through high school: A description of 21 instruments* (Issues & Answers Report, REL 2011–No. 098). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs
Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, *48*(10), 1159–1176. doi:10.1002/tea.20442
Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology*, *93*, 3–13. doi:10.1037/0022-0663.93.1.3
Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist*, *50*(1), 14–30. doi:10.1080/00461520.2014.989230
Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127. doi:10.1159/000094368
Hilpert, J. C., Stempien, J., van der Hoeven Kraft, K. J., & Husman, J. (2013). Evidence for the latent factor structure of the MSLQ: A new conceptualization of an established questionnaire. *SAGE Open*, *3*(4), 1–10. doi:10.5408/1.3604828
Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, *326*, 1410–1412. doi:10.1126/science.1177067
Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure, but autonomy support and structure. *Journal of Educational Psychology*, *102*, 588–600. doi:10.1037/a0019682
Jenkins, E. W., & Pell, R. G. (2006). *The relevance of science education project (ROSE) in England: A summary of findings*. Leeds: Centre for Studies in Science and Mathematics Education, University of Leeds.
Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review*, *32*(2), 125–160. doi:10.1016/j.dr.2012.04.002
Jones, M. G., Howe, A., & Rua, M. J. (2000). Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Science Education*, *84*(2), 180–192. doi:10.1002/(SICI)1098-237X(200003)84:2 < 180::AID-SCE3 > 3.0.CO;2-X

Koballa, T. R., & Glynn, S. M. (2007). Attitudinal and motivational constructs in science education. In S. K. Abell & N. Lederman (Eds.), *Handbook for research in science education* (pp. 75–102). Mahwah, NJ: Erlbaum.

Linnenbrink-Garcia, L., & Pekrun, R. (2011). Students' emotions and academic engagement: Introduction to the special issue. *Contemporary Educational Psychology*, *36*, 1–3. doi:10.1016/j.cedpsych.2010.11.004

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98. doi:10.1037/0033-2909.116.1.75

Maltese, A. V., Melki, C. S., Wiebke, H. (2014). The nature of experiences responsible for the generation and maintenance of interest in STEM. *Science Education*, *98*(6), 937–962. doi:10.1002/sce.21132

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, *37*, 153–184.

National Research Council. (2008). *Research on future skill demands: A workshop summary*. Washington, DC: The National Academies Press.

National Research Council. (2009). Learning science in informal environments. In P. Bell, B. Lewenstein, A. W. Shouse, & M. A. Feder (Eds.), *Learning science in informal environments: People, places, and pursuits* (pp. 9–90). Washington, DC: The National Academies Press.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*(9), 1049–1079. doi:10.1080/0950069032000032199

Ouimet, J. A., Bunnage, J. C., Carini, R. M., Kuh, G. D., & Kennedy, J. (2004). Using focus groups, expert advice, and cognitive interviews to establish the validity of a college student survey. *Research in Higher Education*, *45*(3), 233–250.

Pajares, F. (1997). Current directions in self-efficacy research. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (pp. 1–49). Greenwith, CT: JAI Press.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 259–292). New York, NY: Springer.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation and learning and teaching contexts. *Journal of Educational Psychology*, *95*, 667–686.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*(4), 385–407.

President's Council of Advisors on Science and Technology. (2012). *Engage and excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics* (Report to the President). Washington, DC: Executive Office of the President.

Prokop, P., Prokop, M., & Tunnicliffe, S. D. (2007). Is biology boring? Student attitudes towards biology. *Educational Research*, *42*(1), 36–42. doi:10.1080/00219266.2007.9656105

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

Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of national assessment of educational progress assessments. *Journal of Educational Psychology*, *107*(3), 645. doi:10.1037/edu0000012

Renninger, A., & Bachrach, J. E. (2015). Studying triggers for interest and engagement using observational methods. *Educational Psychologist*, *50*(1), 58–69. doi:10.1080/00461520.2014.999920

Renninger, K. A. (2010). Working with and cultivating interest, self-efficacy, and self-regulation. In D. Preiss & R. Sternberg (Eds.), *Innovations in educational psychology: Perspectives on learning, teaching and human development* (pp. 158–195). New York, NY: Springer.

Richey, J. E., Nokes-Malach, T. J., & Wallace, A. (2014). Achievement goals, observed behaviors, and performance: Testing a mediation model in a college classroom. COGSCI 2014: The Annual Meeting of the Cognitive Science Society, Quebec, Canada.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. doi:10.1006/ceps

Ryu, S., & Lombardi, D. (2015). Coding classroom interactions for collective and individual engagement. *Educational Psychologist*, *50*(1), 70–83. doi:10.1080/00461520.2014.1001891
Schiefele, U. (2009). Situational and individual interest. In K. R. Wentzel, & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 197–222). New York, NY: Routledge.
Senko, C. (2016). Achievement goal theory. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (2nd ed., pp. 75–95). New York, NY: Routledge.
Sha, L., Schunn, C. D., & Bathgate, M. E. (2015). Measuring choice to participate in optional science learning experiences during early adolescence. *Journal of Research in Science Teaching*, *52*(5), 686–709. doi:10.1002/tea.21210
Shen, C., & Tam, H. P. (2008). The paradoxical relationship between student achievement and self-perception: A cross-national analysis based on three waves of TIMSS data. *Educational Research and Evaluation*, *14*, 87–100.
Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, *18*(2), 158–176. doi:10.1521/scpq.18.2.158.21860
Silvia, P. J. (2006). *Exploring the psychology of interest*. New York, NY: Oxford University Press.
Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology*, *42*(1), 70–83. doi:10.1037/0012-1649.42.1.70
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. doi:10.1080/00461520.2014.1002924
Tytler, R., & Osborne, J. (2012). Student attitudes and aspirations towards science. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 597–625). New York, NY: Springer International.
Vedder-Weiss, D., & Fortus, D. (2011). Adolescents' declining motivation to learn science: Inevitable or not? *Journal of Research in Science Teaching*, *48*, 199–216. doi:10.1002/tea.20398
Vedder-Weiss, D., & Fortus, D. (2012). Adolescents' declining motivation to learn science: A follow-up study. *Journal of Research in Science Teaching*, *49*(9), 1057–1095. doi:10.1002/tea.21049
Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Wang, M. T., & Degol, J. (2014). Staying engaged: Knowledge and research needs in student engagement. *Child Development Perspectives*, *8*(3), 137–143. doi:10.1111/cdep.12073
Wang, M. T., & Eccles, J. S. (2012). Adolescent behavioral, emotional, and cognitive engagement trajectories in school and their differential relations to educational success. *Journal of Research on Adolescence*, *22*(1), 31–39. doi:10.1111/j.1532-7795.2011.00753.x
Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, *47*(3), 633–662. doi:10.3102/0002831209361209
Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. doi:10.1006/ceps.1999.1015
Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.
Zimmerman, H. (2012). Participating in science at home: Recognition work and learning in biology. *Journal of Research in Science Teaching*, *49*(5), 597–630. doi:10.1002/tea.21014
Zusho, A., Pintrich, P. R., & Coppola, B. (2003). Skill and will: The role of motivation and cognition in the learning of college chemistry. *International Journal of Science Education*, *25*(9), 1081–1094. doi:10.1080/0950069032000052207