International Journal of Science Education

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

Science engagement and science achievement in the context of science instruction: a multilevel analysis of U.S. students and schools

Larry J. Grabau & Xin Ma

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2017
http://dx.doi.org/10.1080/09500693.2017.1313468

Routledge
Taylor & Francis Group

Check for updates

Science engagement and science achievement in the context of science instruction: a multilevel analysis of U.S. students and schools

Larry J. Grabau[a] and Xin Ma[b]

[a]College of Agriculture, Food and Environment, University of Kentucky, Lexington, KY, USA; [b]Department of Educational, School, and Counseling Psychology, University of Kentucky, Lexington, KY, USA

ABSTRACT
Using data from the 2006 Program for International Student Assessment (PISA), we explored nine aspects of science engagement (science self-efficacy, science self-concept, enjoyment of science, general interest in learning science, instrumental motivation for science, future-oriented science motivation, general value of science, personal value of science, and science-related activities) as outcomes and predictors of science achievement. Based on results from multilevel modelling with 4456 students nested within 132 schools, we found that all aspects of science engagement were statistically significantly and positively related to science achievement, and nearly all showed medium or large effect sizes. Each aspect was positively associated with one of the (four) practices (strategies) of science teaching. Focus on applications or models was positively related to the most aspects of science engagement (science self-concept, enjoyment of science, instrumental motivation for science, general value of science, and personal value of science). Hands-on activities were positively related to additional aspects of science engagement (science self-efficacy and general interest in learning science) and also showed a positive relationship with science achievement.

ARTICLE HISTORY
Received 21 September 2016
Accepted 27 March 2017

KEYWORDS
Science engagement; science achievement; science teaching practices; U.S. adolescents

Science achievement of U.S. students has lagged behind that of many other nations, and the gap between U.S. science achievement and that of the top-performing countries (regions) has particularly been wide, across several international student assessments such as the Programme for International Student Assessment (PISA) administered by the Organization for Economic Cooperation and Development (OECD). In the two recent assessment years with science as the major domain (2006 and 2015), U.S. science achievement was below the PISA international average (OECD, 2007b, 2016).[1] In an effort to address the science performance issue among U.S. students, some reforms have been recommended, such as improved school resources for science education (Wößmann, 2003), reduced class size and certification of a larger percentage of teachers (Fuchs & Wößmann, 2007), and adoption of enhanced science teaching methods (Perera, 2014). Improvement in affect toward science has also been given significant consideration (e.g. Acosta & Hsu, 2014a; Lau & Roeser, 2002; Perera, 2014; Tang & Neber,

CONTACT Xin Ma xin.ma@uky.edu

2008). In fact, improving affect is drawing more attention not only from science educators but also from math and literacy educators (e.g. Lee, 2014; Scherer & Siddiq, 2015). Central to this reform effort is the belief that improved affect is a very promising strategy to improve achievement. The inclusion of comprehensive affective measures in PISA was seen as a signal of OECD to educators worldwide that the development of positive affect was a favourable educational goal and could enhance student achievement (Fensham, 2009). We adopted the same vision in the current research.

The current research joined the nation-wide effort to identify better ways to improve science achievement in the U.S. and was particularly inspired by the importance of the affective domain in science education (Ma, Jong, & Yuan, 2013; Ma & Ma, 2014). Using data from PISA 2006, the PISA cycle with a focus on science education, the current research examined science engagement and science achievement. Specifically, the current research was carried out in two steps. The first focused on science engagement and the second on science achievement.

(1) What is the relationship between (aspects of) science engagement and school climate among U.S. 15-year-olds, with control over student background and school context? (Simply, can school climate predict science engagement?)
(2) Are there any effects (of aspects) of science engagement and school climate on science achievement among U.S. 15-year-olds, with control over student background and school context? (Simply, can science engagement and school climate predict science achievement?)

Therefore, the current research was unique in that it explored science engagement in the context of school climate and examined the effects of science engagement on science achievement in the context of a relationship of school climate to science achievement.

Conception of science engagement

What is currently coming to the forefront of education from the affective domain is the overarching concept of engagement. Fredricks, Blumenfeld, and Paris (2004) defined school engagement as a meta-construct consisting of cognitive, emotional, and behavioural factors. Cognitive engagement included students' investment in schooling and therefore a willingness to commit effort to master their work; emotional engagement related to (positive or negative) responses to the entire school environment (social and academic) and thus influenced both their connection with the school and their intention to participate in school activities; and finally, behavioural engagement connected with actual participation, both within and beyond the context of the school itself. Notably, the primary outcome of interest in Fredricks et al. (2004) was academic achievement. Therefore, from the perspective of school engagement enhancing academic achievement, the comprehensive conception of school engagement by Fredricks et al. (2004) was inclusive of cognitive, emotional, and behavioural strands. The three-strand approach of Fredricks et al. (2004) is a conceptual advancement over the conventional approaches that have included only behavioural and emotional factors (e.g. Skinner, Kindermann, & Furrer, 2009) or simply behavioural factors (e.g. Singh, Granville, & Dika, 2002), for the better capturing of factors involved in engagement which could potentially correlate with academic achievement.

Based on the conceptual framework of Fredricks et al. (2004), science engagement refers to 'a multidimensional concept that broadly encompasses three components, namely behavioural engagement, emotional engagement, and cognitive engagement' (Hampden-Thompson & Bennett, 2013, p. 1327). OECD (2009) has taken a similar stance to define science engagement as covering 'self-related cognitions, motivational preferences, emotional factors as well as behavioral-related variables (such as participation in science-related activities in and out of school)' (p. 55) (see also Chen, 2005; Demanet & Van Houtte, 2014; Lam et al., 2014). Nine aspects of science engagement were measured in PISA 2006 (OECD, 2007a) as indicators of the comprehensive concept of science engagement (Fredricks et al., 2004; Hampden-Thompson & Bennett, 2013).

Aspects of science engagement

Science self-efficacy

This concept refers to a student's 'confidence in performing science-related tasks' (OECD, 2009, p. 464). Uitto (2014) related this concept to a student's perception of capacity to apply acquired knowledge and skills when confronted with new science-based tasks. Four sources acted in an additive way to build science self-efficacy including mastery experience, vicarious experience, social persuasion, and physiological (psychological) state with mastery experience as the leading source (Chen & Usher, 2013). In PISA 2006, dimensions of science self-efficacy involved making use of scientific evidence, explaining observations in scientific ways, and identifying issues in scientific terms (OECD, 2009). Apart from applying laboratory skills, the work of Aydin and Uzuntiryaki (2009) affirmed the cognitive and application-oriented approach of OECD (see also Lin, Tan, & Tsai, 2013). Science self-efficacy was positively associated with science achievement in several countries such as Canada (Areepattamannil, Freeman, & Klinger, 2011), Finland (Lavonen & Laaksonen, 2009), Germany (Scherer, 2013), Hong Kong (Lam & Lau, 2014; Sun, Bradley, & Akers, 2012), Taiwan (Maynard Wang, Wu, & Iris Huang, 2007), and the U.S. (Lau & Roeser, 2002). Perera (2014) confirmed this relationship across 15 nations. Turkish high school students with high science self-efficacy tended to expend more effort when confronted with difficult science problems (Sungur, 2007).

Science self-concept

This concept commonly refers to a student's belief that he or she can easily learn and understand science. Parents, peers, teachers, and the resulting frames of reference influenced the formation of science self-concept (e.g. Jen, Lee, Chien, Hsu, & Chien, 2013). Steinmayr, Dinger, and Spinath (2012) considered dimensions of science self-concept as intrinsic task value, importance, and utility. Others identified science self-concept primarily as an expectation for successful performance in science (e.g. Bhanot & Jovanovic, 2009; Riegle-Crumb, Moore, & Ramos-Wada, 2011). PISA 2006 operationalised dimensions of science self-concept to indicate (by means of self-assessment) ease in learning in science, good understanding of newly presented science concepts, and good answers to questions about science (OECD, 2009). Science self-concept was positively related to science achievement in several regions of the world such as Canada (Areepattamannil, 2012b; Areepattamannil et al., 2011; Areepattamannil & Kaur, 2013), East Asian nations (regions) (Shen, 2005; Yu, 2012), Finland (Lavonen & Laaksonen, 2009), Saudi Arabia (Tighezza, 2014), Singapore (Mohammadpour, 2013), Taiwan (Jen et al., 2013), and the U.S. (Lau & Roeser, 2002; Shen, 2005; Yu, 2012). English students who continued to study physics after it was no longer compulsory showed high science self-concept (Mujtaba & Reiss, 2014). In contrast, Anagün (2011) found no relationship between science self-concept and science achievement for Turkish students, while Bouhlila (2011) observed a negative relationship for students from a group of MENA (Middle East and North Africa) nations. Finally, correlations between science self-efficacy and science self-concept were .60 for a 500 sample of U.S. students and .55 for the entire OECD sample in PISA 2006 (OECD, 2009).

Enjoyment of science

This concept is derived from a student's feelings of fun and happiness when engaging in science learning activities (see Shumow, Schmidt, & Zaleski, 2013). Enjoyment of science took its shape from positive relationships with peers and teachers (Jen et al., 2013) and science instructional strategies (e.g. hands-on activities) (Hampden-Thompson & Bennett, 2013; Shumow et al., 2013). A positive relationship between enjoyment of science and science achievement was observed in several countries such as Canada (Areepattamannil et al., 2011), Finland (Lavonen & Laaksonen, 2009), Hong Kong (Lam & Lau, 2014), Malaysia and Singapore (Ng, Lay, Areepattamannil, Treagust, & Chandrasegaran, 2012), Saudi Arabia (Tighezza, 2014), and Taiwan (Jen et al., 2013; Tsai & Yang, 2015). Enjoyment of science was associated with science achievement of second (but not first) generation immigrants to Canada (Areepattamannil, 2012b), and it stimulated continued study of physics by English students after this subject was no longer required (Mujtaba & Reiss, 2014). In contrast, enjoyment of science was negatively related to science achievement in a group of MENA nations (Bouhlila, 2011).

General interest in science

This concept refers to a student's expressed desire to learn about an array of science subjects and methodologies (see Areepattamannil, 2012a; Larson, Stephen, Bonitz, & Wu, 2014). Swarat, Ortony, and Revelle (2012) were more concerned about a psycho (physiological) state of interest likely more transient. Science learning environment (e.g. hands-on involvement, interaction through technology, and relevance to daily life) (Areepattamannil, 2012a; Jocz, Zhai, & Tan, 2014; Swarat et al., 2012) and friendship group support (Robnett & Leaper, 2013) stimulated general interest in science. While general interest in science was positively associated with science achievement in Qatar (Areepattamannil, 2012a), this relationship was negative for students in both Canada (Areepattamannil et al., 2011) and Finland (Lavonen & Laaksonen, 2009). Finally, correlations between enjoyment of science and general interest in science were .73 for the U.S. sample and .75 for the OECD sample in PISA 2006 (OECD, 2009). Although Anderson and Chen (2016) conflated these related constructs as a single concept of 'interest-enjoyment value' (p. 57), more researchers adopted the PISA differentiation between enjoyment of science and general interest in science with the former focusing more on a student's affect while involved in science activities and the latter focusing more on a student's preference for specific science disciplines (e.g. Jack, Lin, & Yore, 2014; Woods-McConney, Oliver, McConney, Maor, & Schibeci, 2013; Woods-McConney, Oliver, McConney, Schibeci, & Maor, 2014). Furthermore, Ainley and Ainley (2011) identified a transcendent (across four culturally diverse nations) of enjoyment of science on interest in science.

Instrumental motivation for science

This concept refers to a student's cognitive investment in science learning because the student perceives that science is of value for future study and work (see Acosta & Hsu, 2014a; Farmer, Wardrop, & Rotella, 1999; Hampden-Thompson & Bennett, 2013). Similarly, Spearman and Watt (2013) concerned the 'extrinsic utility value' of science (p. 223). Contexts of teaching and learning (e.g. classroom interaction) (Hampden-Thompson & Bennett, 2013) as well as science self-concept, peer context, and lack of teacher encouragement for science (George, 2006) determined instrumental motivation for science. The formation of instrumental motivation for science was also influenced by parents through the so-called parental 'utility-value intervention' (Rozek, Hyde, Svoboda, Hulleman, & Harackiewicz, 2015, p. 195). Instrumental motivation for science was positively related with science achievement of students in Hong Kong (Sun et al., 2012) and New Zealand (Acosta & Hsu, 2014a). Perera (2014) confirmed this relationship across 15 nations. Both East Asian and U.S. eighth grade students with higher instrumental motivation for college placement showed higher science achievement (Yu, 2012). English high school students with higher degree of instrumental motivation for science were more likely to continue to study physics after it was no longer compulsory (Mujtaba & Reiss, 2014).

Future-oriented science motivation

This concept focuses on a student's expressed desire to work, study, or spend life time in an area of science (see Lupart, Cannon, & Telfer, 2004; Riegle-Crumb et al., 2011). Hampden-Thompson and Bennett (2013) observed that future motivation for science could come from science instruction (teaching strategies) that emphasised interactions, hands-on activities, and applications. Both immigrant and non-immigrant students across Canada had positive associations between future motivation for science and science achievement (Areepattamannil & Kaur, 2013). Finally, instrumental motivation for science and future-oriented science motivation showed correlation in both U.S. (.67) and OECD (.72) samples (OECD, 2009). Shin et al. (2015) differentiated between instrumental motivation for science as a stepping stone for a future science-related career and future motivation for science as a desire to study or work in a science area in the future.

General value of science

This concept represents a student's belief that scientific advancement provides benefits to society as a whole (see Acosta & Hsu, 2014b; Khalijah et al., 1995). General value of science was built under the influence of parents who held high general value of science (Acosta & Hsu, 2014b). Students with higher general value of science had higher science achievement in Hong Kong (Acosta & Hsu, 2014b; Lam & Lau, 2014).

Personal value of science

This concept relates the relevance of science (science topics) to students as individuals (see Else-Quest, Mineo, & Higgins, 2013; Viljaranta, Nurmi, Aunola, & Salmela-Aro, 2009). Non-immigrant students in Canada showed no relationship between personal value of science and science achievement; in contrast, for immigrant students, this relationship was negative (Areepattamannil & Kaur, 2013). Personal value of science was related to science achievement for East Asian but not U.S. students (Yu, 2012). Finally, general value of science and personal value of science were related in the U.S. sample (.77) and the OECD sample (.78) (OECD, 2009). Woods-McConney et al. (2013) observed that students tended to have stronger general value of science (value relating to society at large) than personal value of science (value relating to individuals) (see also Vázquez Alonso & Manassero Mas, 2009).

Science-related activities

This concept refers to a student's self-chosen science activities outside of the context of school (see Ho, 2010). Gerber, Cavallo, and Marek (2001) described such activities as creating 'enriched informal learning environments' (p. 539). While VanMeter-Adams, Frankenfeld, Bases, Espina, and Liotta (2014) emphasised (formal) extracurricular activities as the backbone of science-related activities, Ho (2010) considered any (outside) experiences of children with science as science-related activities (e.g. watching TV programmes about science, reading science stories on the internet). Similarly, Lin, Lawrenz, Lin, and Hong (2013) referred to science-related activities as 'engagement with leisure science learning' (p. 945). Engagement in science-related activities has been shown to enhance science achievement in Hong Kong (Ho, 2010), Turkey (Kalendar & Berberoglu, 2009), and the U.S. (Sha, Schunn, & Bathgate, 2015; Tran, 2011). Gerber et al. (2001) observed enhanced scientific reasoning ability among 7th through 10th grade U.S. students who had opportunities for out-of-school science activities.

Theoretical basis for analytical framework

We adopted the input-process-output (IPO) model (see Ilgen, Hollenbeck, Johnson, & Jundt, 2005) that has been widely applied in school effects research to guide the selection of variables and specification of statistical models (see Ma, Ma, & Bradley, 2008). In the IPO model, students bring different individual and family characteristics and different cognitive and affective conditions into their schools. Schools then process, by means of context and climate, students with different backgrounds into different categories of outcome measures (e.g. attitude, achievement). As illustrated in Ma et al. (2008), researchers who use the IPO model carefully control student background and school context in order to examine the relationship between outcome measures and school climate. This analytical strategy is straightforward in that educators have a greater deal of 'power' to change school climate (e.g. school policy, classroom practice) than individual and family characteristics of students and school context (e.g. school size, student composition). We adopted this theoretical basis also because it matches well with our primary statistical technique of multilevel modelling that deals with data with hierarchical structure such as students nested within schools. Essentially, we controlled individual and family characteristics of students at the student level and contextual characteristics of schools at the school level so that we were able to focus on the effects of climate characteristics of schools on outcome measures.

Method

Data

Data on 15-year-old U.S. students were acquired from the 2006 PISA dataset (OECD, 2007a). Multi-stage stratified random sampling was used to sample the U.S. 15-year-old student population (OECD, 2009). Schools were sampled first from a comprehensive national list of all eligible schools with sampling probability proportional to the size of a school. Once schools were selected, approximately 35 students were randomly selected from the eligible list based on age. If a selected school had less than 35 students of the targeted age, all students were selected. The U.S. sample included 4456 students in 132 schools. The PISA dataset included both student and school sampling weights (OECD, 2007a), and they were included in our analysis.

Outcome measures

Nine aspects of science engagement were used as outcome (dependent) variables in the first stage of the current research (science self-efficacy, science self-concept, enjoyment of science, general interest in science, instrumental motivation for science, future-oriented science motivation, general value of science, personal value of science, and science-related activities). Each aspect was measured with a scale consisting of several items on the student questionnaire (OECD, 2007d). Appendix A presents items that describe those aspects of science engagement. For all aspects, the PISA team implemented item response theory (IRT) to derive final student measures (composite variables) with higher values consistently indicating more positive aspects of science engagement.[2]

In the second stage of the current research, science achievement was the outcome variable (aspects of science engagement functioned as key independent variables). The definition for scientific literacy by PISA refers to an individual's

> scientific knowledge and use of that knowledge to identify questions, to acquire new knowledge, to explain scientific phenomena, and to draw evidence-based conclusions about science-related issues, understanding of the characteristic features of science as a form of human knowledge and inquiry, and awareness of how science and technology shape our material, intellectual, and cultural environments. (OECD, 2007c, p. 35)

This definition was operationalised as science achievement based on a combined literacy scale of using scientific evidence, identifying scientific issues, and explaining phenomena scientifically (see OECD, 2007c). Five plausible values of science achievement for each student were estimated by PISA; these plausible values were properly combined to use as science achievement in the current research.

Independent variables

Based on the review of research literature (e.g. Ma et al., 2008), independent variables were selected at both student and school levels. Six student-level variables, all considered to be exogenous to the outcome variables, were selected from the PISA dataset. Those variables included gender (dichotomous with 1 as male), age, father's socioeconomic status (SES) and mother's SES (PISA standardised indices), immigration background (dichotomous with 1 as at least one parent born in the U.S.), and language spoken at home (dichotomous with 1 as English). According to Ma et al. (2008), these variables provided a good control of student and family characteristics.

At the school level, school context variables included school size (total enrolment), school type (dichotomous with 1 as public), proportion of girls, school mean father's SES and school mean mother's SES (aggregated from SES of students within a school), proportion of teachers certified, student-teacher ratio, teacher shortage (PISA index), and quality of educational resources (PISA index). According to Ma et al. (2008), these variables provided a good control of school contextual characteristics.

Meanwhile, school climate variables included school responsibility (autonomy) for resource allocation (PISA index), school responsibility (autonomy) for curriculum and assessment (PISA index), ability grouping (dichotomous with 1 as grouping either between or within classes), parent influence (count of 'yes' to four items about parent groups exerting direct influences on educational decisions), teacher influence (count of 'yes' to four items about teacher groups exerting direct influences on educational decisions), school activities to promote the learning of science (count of 'yes' to five school activities), school mean science teaching – focus on models or applications, school mean science teaching – hands-on activities, school mean science teaching – interaction, and school mean science teaching – student investigations (see Appendix B for a description of these science teaching practices).[3]

Statistical procedure

To address the data hierarchy of students nested within schools in the PISA dataset, hierarchical linear modelling (HLM) or multilevel modelling was employed (see Ma et al., 2008; Raudenbush & Bryk, 2002), using HLM7 software (Raudenbush, Bryk, & Congdon, 2011). To address our first research question, aspects of science engagement were modelled separately with student and family characteristics at the student level and school context and school climate at the school level. With control over individual and family differences at the student level and school contextual differences at the school level, the relationship between science engagement and school climate was examined.[4] To address our second research question, science achievement was modelled with aspects of science engagement (separately) as well as student and family characteristics at the student level and school context and school climate at the school level. With control over individual and family differences at the student level and school contextual differences at the school level, the effects of science engagement and school climate on science achievement were examined. In both stages of analysis, in pursuit of parsimonious models, we deleted variables not statistically significant in a model from both student and school levels one at a time (the one with the largest *p* value) until all variables in the model were statistically significant.[5] Finally, to evaluate the performance of our HLM models, we calculated proportion of variance explained at the student and school levels for each (final) parsimonious model (see Raudenbush & Bryk, 2002).[6]

Technically, we employed the random intercept model in HLM in which all student-level parameters (variables) were fixed at the school level except the intercept. We had two reasons to do so. First, Thum and Bryk (1997) recommended that if the variation (variance) in student-level effects among schools is not any part of the research question then student-level variables need to be fixed. Second, only when student-level variables are fixed, proportion of variance explained by the HLM model can be calculated as a measure of the overall model performance.

Our large number of students and schools and our purpose of comparing results across aspects of science engagement created the need to work with effect size. We made use of Cohen's *d* to estimate effect size.[7] Cohen's *d* was calculated for each statistically significant independent variable in each parsimonious model by dividing the coefficient by its corresponding standard deviation (SD) of the outcome. In general, effect sizes of .20–.49 are considered small, .50–.79 are considered medium, and those of .80 or greater are considered large (Cohen, 1988).

Results

We used student and family characteristics at the student level and school contextual characteristics at the school level as control to produce 'purer' effects of science engagement and school climate. Therefore, we omitted any interpretations of the control variables to save space. Also for the sake of space and for an emphasis on comparison across results, we highlighted effect size for any statistically significant variable at either student or school level.

Aspects of science engagement as outcome variables

Science self-efficacy. According to Table 1, after control for student and family variables at the student level and school contextual variables at the school level, school climate that correlated statistically significantly with science self-efficacy was school mean science teaching – hands-on activities; effect was positive and effect size was small (.26). In terms of variance in science self-efficacy, the multilevel model explained 86% at the school level and 7% at the student level.

Science self-concept. School climate that correlated statistically significantly with science self-concept was school mean science teaching – focus on applications or models; effect was positive and effect size was small (.33). The multilevel model explained 67% of the school-level variance and 7% of the student-level variance in science self-concept.

Enjoyment of science. School climate that correlated statistically significantly with enjoyment of science was school mean science teaching – focus on applications or models, with positive effect and small effect size (.32). Proportion of variance in enjoyment

Table 1. Statistical results for multilevel models of student-level and school-level effects on aspects of science engagement.

	Science self-efficacy		Science self-concept		Enjoyment of science		General interest in science	
	Effect	SE	Effect	SE	Effect	SE	Effect	SE
Student-level variables								
Gender (D)	−.29	(.05)	−.40	(.05)	−.12	(.06)		
Age	−.15	(.07)					−.20	(.08)
Father SES	.01	(.00)	.01	(.00)	.01	(.00)		
Immigration status (D)							−.28	(.06)
Language at home (D)	.38	(.07)						
School context variables								
Proportion of Girls			.73	(.20)				
Teacher shortage							−.06	(.03)
School climate variables								
Promotion of science							−.06	(.02)
Focus on applications			.58	(.16)	.43	(.12)		
Hands-on activities	.39	(.13)					.39	(.09)

	Instrumental motivation in science		Future motivation in science		General value of science		Personal value of science		Science-related activities	
	Effect	SE	Effect	SE	Effect	SE	Effect	SE	Effect	SE
Student-level variables										
Gender (D)			−.18	(.06)	−.20	(.06)	−.15	(.05)	−.27	(.05)
Age							−.16	(.08)		
Father SES	.01	(.00)			.01	(.00)			.01	(.00)
Immigration status (D)	−.15	(.07)	−.19	(.04)						
Language at home (D)									−.35	(.08)
School context variables										
School size			−.01	(.00)					.01	(.00)
Proportion of girls			.37	(.16)						
School mean father SES			.01	(.00)						
Student–teacher ratio							−.01	(.00)	−.02	(.01)
School climate variables										
Ability grouping (D)			.19	(.05)						
Promotion of science			.04	(.01)						
Parent influence			.03	(.01)						
Focus on applications	.33	(.12)			.91	(.17)	.64	(.15)		
Interaction					−.67	(.15)			.39	(.16)
Student investigations			.19	(.06)						

Notes: SE, standard error; SES, socioeconomic status; D, dichotomous variable. All estimates are statistically significant at the alpha level of .05.

of science explained by the multilevel model was 61% at the school level and 3% at the student level.

General interest in learning science. School climate that correlated statistically significantly and positively with general interest in learning science was school mean science teaching – hands-on activities, with small effect size (.38). School climate that correlated statistically significantly and negatively with general interest in learning science was school activities to promote the learning of science, with small effect size (.29). The negative effect made sense in that schools where students lacked general interest in learning science might attempt to use more activities to promote the learning of science. Proportion of variance in general interest in learning science explained by the multilevel model was 68% at the school level and 3% at the student level.

Instrumental motivation for science. Stronger school mean science teaching – focus on applications or models was statistically significantly associated with higher instrumental motivation for science; effect size was small (.24). While proportion of variance explained at the school level was 60%, the corresponding proportion at the student level was 1%.

Future-oriented science motivation. Stronger school mean science teaching – student investigations was statistically significantly associated with higher future-oriented science motivation; effect size was small (.26). School practice of ability grouping, more school activities to promote the learning of science, and stronger parent influence were also statistically significantly associated with higher future-oriented science motivation; effect sizes were all small (.36, .24, and .18, respectively). While proportion of variance explained at the school level was 92%, the corresponding proportion at the student level was 2%.

General value of science. While school mean science teaching – focus on applications or models was statistically significantly and positively associated with general value of science, school mean science teaching – interaction was statistically significantly and negatively associated with this aspect of science engagement. The negative effect made sense in that schools where students lacked general value of science might attempt to involve students more in science classrooms. While effect size for the former was (nearly) medium (.48), effect size for the latter was small (.39). Model performance indicated that 73% of the variance in general value of science was explained at the school level and 9% at the student level.

Personal value of science. School mean science teaching – focus on applications or models was statistically significantly and positively associated with personal value of science. Effect size was small (.39). Model performance indicated that 72% of the variance in personal value of science was explained at the school level and 1% at the student level.

Science-related activities. School mean science teaching – interaction was statistically significantly and positively associated with science-related activities (outside of school); effect size was small (.22). In terms of variance in science-related activities, the multilevel model accounted for 36% at the school level and 8% at the student level.

Science achievement as outcome variable

Before we interpret Table 2 that presents the effects of science engagement (and school climate) on science achievement, we provide some essential estimates on science achievement. The grand average of U.S. 15-year-olds was 505.55 in science achievement; intraclass correlation (proportion of variance attributable to schools) was .18. We estimated a multilevel model without any aspects of science engagement (as independent variables at the student level) (see Appendix C). Three school climate variables emerged with statistically significant effects on science achievement: school responsibility for curriculum and assessment (positive), school mean science teaching – hands-on activities (positive), and school mean science teaching – student investigations (negative). The negative effect appeared to indicate the importance of time and activities of teachers rather than students in science classrooms. Effect sizes were all small (.21, .24, and .37, respectively). Proportion of variance in science achievement explained by the model was 78% at the school level and 8% at the student level.[8] We used this model as our baseline model for comparison with each of the nine successive multilevel models, each of which included a single aspect of science engagement.

Table 2. Statistical results for multilevel models of aspects of science engagement, student-level, and school-level effects on science achievement.

	Science self-efficacy on science achievement		Science self-concept on science achievement		Enjoyment of science on science achievement		General interest in science on science achievement	
	Effect	SE	Effect	SE	Effect	SE	Effect	SE
Student-level variables								
Aspect	33.43	(2.33)	29.59	(2.60)	29.07	(2.67)	18.56	(2.45)
Father SES	.39	(.19)	.56	(.18)	.52	(.19)	.60	(.20)
Mother SES	.50	(.23)	.56	(.21)	.50	(.22)	.57	(.24)
Language at Home (D)	18.65	(7.61)	27.11	(7.85)	31.12	(8.83)	31.81	(7.79)
School context variables								
School size	−.01	(.00)	−.01	(.00)	−.01	(.00)	−.01	(.00)
School mean father SES	2.01	(.55)	2.05	(.55)	1.63	(.57)	1.88	(.64)
School climate variables								
Curriculum & assessment	8.84	(4.41)						
Hands-on activities					34.32	(13.89)	36.86	(15.85)
Student investigations	−42.26	(12.95)	−53.93	(12.38)	−56.54	(11.28)	−53.65	(12.50)

	Instrumental motivation in science on science achievement		Future motivation in science on science achievement		General value of science on science achievement		Personal value of science on science achievement		Science-related activities on science achievement	
	Effect	SE	Effect	SE	Effect	SE	Effect	SE	Effect	SE
Student-level variables										
Aspect	19.35	(2.56)	22.44	(3.33)	22.20	(2.11)	24.34	(2.31)	21.26	(2.41)
Father SES	.63	(.20)	.64	(.17)	.65	(.18)	.63	(.18)	.55	(.20)
Mother SES	.59	(.22)	.58	(.22)			.52	(.25)	.52	(.25)
Language at Home (D)	31.13	(8.28)	32.11	(8.93)	38.43	(7.56)	30.35	(8.76)	36.09	(8.71)
School context variables										
School size	−.01	(.00)	−.01	(.00)	−.01	(.00)	−.01	(.00)	−.01	(.00)
School mean father SES	2.35	(.61)	2.20	(.58)	1.94	(.60)	1.86	(.62)	1.87	(.63)
School climate variables										
Hands-on activities					31.87	(14.75)	28.88	(13.71)	34.62	(15.37)
Student investigations	−43.12	(16.55)	−44.30	(16.24)	−52.41	(12.23)	−57.53	(12.51)	−55.40	(12.69)

Notes: SE, standard error; SES, socioeconomic status; D, dichotomous variable. All estimates are statistically significant at the alpha level of .05.

Science self-efficacy on science achievement. With the baseline model as the background, Table 2 indicates that science self-efficacy was statistically significantly and positively related to science achievement even after control over statistically significant student and school characteristics, with large effect size of .98. Over and above this large effect of science self-efficacy on (individual) science achievement, stronger school responsibility for curriculum and assessment was statistically significantly associated with higher (school average) science achievement, and stronger school mean science teaching – student investigations was statistically significantly associated with lower (school average) science achievement. Effect sizes were small (.24 and .29, respectively). Proportion of variance in science achievement accounted for by the multilevel model was 80% at the school level and 23% at the student level. Since the multilevel model with no aspects of science engagement explained only 8% of the student-level variance, this impressive improvement in model performance (8–23%) was brought in entirely by the addition of science self-efficacy, indicating the highly important contribution of science self-efficacy to the explanation of variance in science achievement.

Science self-concept on science achievement. Science self-concept was statistically significantly and positively related to science achievement even after control over statistically significant student and school characteristics, with large effect size of 1.07. Over and above this large effect of science self-concept on (individual) science achievement, weaker school mean science teaching – student investigations was statistically significantly associated with higher (school average) science achievement, with small effect size of .38. Proportion of variance in science achievement explained was 75% at the school level and 20% at the student level. Again, since the multilevel model with no aspects of science engagement explained only 8% of the student-level variance, this result indicated the highly important contribution of science self-concept to the explanation of variance in science achievement. For the sake of space, we simplified the interpretation on other aspects of science engagement.

Enjoyment of science on science achievement. This aspect of science engagement was statistically significantly and positively related to science achievement, with (nearly) medium effect size of .49. Both stronger school mean science teaching – hands-on activities and weaker school mean science teaching – student investigations were statistically significantly associated with higher science achievement. Effect sizes were small (.22 and .44 respectively). Proportion of variance in science achievement accounted for by the multilevel model was 81% at the school level and 18% at the student level.

General interest in learning science on science achievement. This aspect of science engagement was statistically significantly and positively related to science achievement, with medium effect size of .56. Both stronger school mean science teaching – hands-on activities and weaker school mean science teaching – student investigations were statistically significantly associated with science achievement, with small effect sizes (.21 and .38, respectively). The multilevel model explained 78% of the school-level variance and 12% of the student-level variance in science achievement.

Motivations on science achievement. Both instrumental motivation for science and future-oriented science motivation were statistically significantly and positively related to science achievement; effect sizes were medium (.58) for the former and small (.16) for the latter. Weaker school mean science teaching – student investigations was statistically significantly tied to higher science achievement in both cases, with small effect sizes of .23 and .24, respectively. Also respectively, proportion of variance in science achievement explained was 69% and 67% at the school level and 13% and 15% at the student level.

Values on science achievement. Both general value of science and personal value of science indicated statistically significant effects on science achievement, with large effect sizes of 1.38 and 1.51, respectively. In both cases, stronger school mean science teaching – hands-on activities and weaker school mean science teaching – student investigations were statistically significantly associated with higher science achievement, with small effect sizes ranging from .19 to .41. Respectively, proportion of variance in science achievement explained was 78% (in both cases) at the school level and 10% and 15% at the student level.

Science-related activities on science achievement. This aspect of science achievement was statistically significantly and positively related to science achievement, with medium effect size of .61. Both stronger school mean science teaching – hands-on activities and weaker school mean science teaching – student investigations were statistically significantly tied to higher science achievement, with small effect sizes (.20 and .39, respectively). Proportion of variance in science achievement explained by the multilevel model was 77% at the school level and 12% at the student level.

Discussion

Summary of principal findings

All nine aspects of science engagement were statistically significantly and positively related to science achievement, and more importantly, nearly all showed medium or large effect sizes. Paired but different aspects of science engagement were noticeable in the pattern of the effects. Large effect sizes were noted for general value of science and personal value of science as well as for science self-efficacy and science self-concept. Medium effect sizes were noted for enjoyment of science and general interest in learning science. This pattern was less evident for instrumental motivation for science (medium) and future-oriented science motivation (small). Finally, science-related activities showed medium effect size.

Among school climate variables, school mean science teaching measures (practices) connected most frequently with aspects of science engagement; in fact, each of the nine aspects was positively associated with one of the (four) practices (strategies) of science teaching. Focus on applications or models was positively related to the most aspects of science engagement (science self-concept, enjoyment of science, instrumental motivation for science, general value of science, and personal value of science). Hands-on activities were positively related to additional aspects of science engagement (science self-efficacy and general interest in learning science). Student investigations and interaction were each positively related to one aspect of science engagement (future-oriented science motivation and science-related activities, respectively). Finally, hands-on activities showed a (direct) positive relationship not only with science engagement but also with science achievement.

Science engagement matters to science achievement

What is unique about the current research is our focus on aspects of science engagement above and beyond a suite of 'powerful' predictors of science achievement (student and family background characteristics). In fact, all nine aspects of science engagement showed effects on science achievement, and eight had either medium or large effect sizes (including .49 for enjoyment of science). Therefore, we conclude that science engagement matters greatly to science achievement. Enhancement of various aspects of science engagement would improve science achievement of U.S. students. Our policy recommendation is for U.S. education policymakers, school leaders, and classroom teachers to enhance aspects of science engagement of students.

Based on our estimates, improvement in science achievement of U.S. students, associated with improvement in their science engagement, should be quite substantial. For example, our largest effect size of 1.51 for personal value of science indicates that improvement in personal value of science (1 point increase on a scale of 1–4) of an average-achieving U.S. student would move this student from 50th percentile to 95th percentile in science achievement (see McGough & Faraone, 2009). Even our smallest effect size of .16 for future-oriented science motivation indicates that (the same) improvement in future-oriented science motivation of an average-achieving U.S. student would move this student from 50th percentile to 58th percentile in science achievement. Priorities can be particularly given to the improvement of general value of science and personal value of science as well as science self-efficacy and science self-concept because they demonstrated large effect sizes.

Science teaching matters to science engagement

Given that science engagement matters so much to science achievement, how can we enhance science engagement? We are able to provide insights on issues relevant to appropriate school climate strategies that improve aspects of science engagement. Perhaps expected, among all school climate variables, those concerning science teaching practices (strategies) stood out as consistent predictors of aspects of science engagement. We recommend two science teaching practices as the key mechanisms for enhancing many aspects of science engagement (science teaching with a focus on applications or models and science teaching with a focus on hands-on activities). In fact, these two science teaching practices were positively related to seven of the nine aspects of science engagement. Focus on applications or models was positively related to five aspects of science engagement (science self-concept, enjoyment of science, instrumental motivation for science, general value of science, and personal value of science). Thus, this science teaching practice appears to hold broad merit in fostering science engagement. Hands-on activities were related to two (additional) aspects of science engagement (science self-efficacy and general interest in learning science). Overall, the powerful relationships of these aspects of science engagement to science achievement (as we discussed earlier) call for the use of every available strategy to enhance these aspects of science engagement.

According to our estimates, among effect sizes for focus on applications or models and hands-on activities, the largest effect size was .48 for focus on applications or models, indicating that improvement in this science teaching practice (1 point increase on a scale of 1–4) would move an average U.S. student from 50th percentile to 69th percentile in general value of science. The smallest effect size was .24 for focus on applications or models, indicating that (the same) improvement in this science teaching practice would move an average U.S. student from 50th percentile to 58th percentile in instrumental motivation in science. These improvements in aspects of science engagement are expected to correlate with improved science achievement as we discussed earlier.

Science teaching matters to science achievement

We especially emphasise one science teaching practice (strategy) that indicated a direct relationship not only with science engagement but also with science achievement, science teaching with a focus on hands-on activities. In five out of the nine multilevel models containing individual aspects of science engagement, this science teaching practice (hands-on activities) demonstrated a positive relationship with science achievement. More importantly, we emphasise that this science teaching practice showed effects first over and above the strong effects associated with various aspects of science engagement and the effects of powerful predictors of science achievement (student and family background characteristics) and second above and beyond other important (statistically significant) characteristics associated with school context and school climate. Therefore, the unique contribution or importance of hands-on activities to science achievement simply cannot be emphasised enough. Our policy recommendation is for the U.S. science education community to pay close attention to this science teaching practice. Indeed, this science teaching practice, together with the science teaching practice that focuses on applications or models, should become the centre piece of professional development in science education, with a great anticipation of benefits for both cognitive and affective domains in science education.

To quantify our estimates on this issue for illustration, we use the same approach. In this case, effect sizes associated with hands-on activities as science teaching practice (strategy) were very similar (ranging from .19 to .22), indicating that improvement in this science teaching practice (1 point increase on a scale of 1–4) would move an average-achieving U.S. student from 50th percentile to 58th percentile in science achievement. This unique contribution of hands-on activities makes this science teaching practice an extremely valuable 'player' in any 'team' effort aimed at improving science achievement of U.S. students.

Limitations and further research

There are several ways to improve and extend the current research. Generally, our results on science engagement and science achievement were established on the PISA 2006 database (with science as the major domain). With the release of the PISA 2015 database (with science as the major domain), our analytical framework can be re-run to either replicate our results or reveal departures caused by nearly a decade of improvement in science education (i.e. the stability issue). Specifically, five aspects of the current research are in need of improvement and advancement. First, instead of treating the nine aspects of science engagement as separate outcomes, a multiple outcome model can be developed to allow the sharing of information among these aspects of science engagement, using MLwiN2 software (Rasbash, Steele, Browne, & Goldstein, 2012). Second, instead of specifying a random intercept model, a random slope model (e.g. allowing effects of science engagement on science achievement to vary across schools) can be tested to allow a (selective) comparison of model performance between the two models. Third, aspects of science engagement with strong effects on science achievement can be investigated together for potential interaction effects, which was beyond the scope of the current research. Fourth, instead of assuming science engagement as a 'cause' of science achievement (i.e. science engagement as predictors), a more sophisticated model can be developed to treat this relationship as reciprocal (i.e. engagement improves achievement and achievement enhances engagement at the same time). Finally, efforts are needed to improve the low proportion of variance accounted for at the student level by our multilevel models with science engagement as outcomes. Although we noticed the lack of a good identifier for race/ethnicity typically important within the U.S. context, traditional exogenous predictors of science achievement may simply not apply to the prediction of science engagement. Researchers may need to explore student characteristics far beyond traditional student and family background. These improvements and advancements will produce more insightful empirical evidence on the complex relationship between science engagement and science achievement.

Notes

1. Overall, students' performance in science remained essentially unchanged from 2006 to 2015 in the majority of countries (including the U.S.) with comparable data (OECD, 2016). Also similar between the two years is that only a minority of students took part in science activities and expressed an interest in working in science although a majority of them reported an interest in science (OECD, 2007b, 2016).
2. Kolmogorov–Smirnov test of normality did not indicate any serious concern about any aspect of science engagement. This has to do with the IRT procedure that the PISA staff carried out in creating each composite variable. Although the items were measured on a Likert scale, the composite variables were actually indices.
3. The nature of the PISA sample has some implications for school climate effects. PISA samples schools first and then students. Because teachers (classrooms) are not a unit of sampling (and thus not a level of analysis), measures of school climate derived from students, teachers, and principals are all general measures. For example, classroom practice concerns general pedagogical approaches (within a school similar to a form of aggregation) rather than specific practices in a science classroom.
4. We treated the nine aspects of science engagement separately in all our analyses. We made this decision based on the range of correlations among these aspects from .30 to .63 with the vast majority of correlations below .50 (the only exceptional correlation was .70). We also considered these aspects as theoretically well distinct constructs. Finally, empirical studies just began to explore the largely uncharted theoretical framework of Fredricks et al. (2004), and we thought that detailed examination of each aspect would be beneficial as the foundation for future research to build from. The caveat of this approach is that multiple comparisons were performed which might potentially compromise the significance (alpha) level of .05.
5. There are different ways to arrive at a final statistical model. Some researchers keep all variables that can be theoretically justified in the model even though some of them may not be statistically significant. We adopted a different approach that seeks a parsimonious final model (see Vandekerckhove, Matzke, & Wagenmakers, 2015). This approach often results in stronger effects once (competing) variables not significant are removed from the model. A parsimonious model may also reduce the complexity of (hidden or not modelled) interactions among variables. The more variables, the more complex of how the effects of a key variable are influenced by the hidden interactions among a large number of variables.
6. Statistically, proportion of variance explained by a final HLM model is calculated as the difference in variance between the final model and a null model (that contains no predictors at any level) divided by the variance from the null model (see Raudenbush & Bryk, 2002).
7. How to calculate effect size in HLM remains unclear (very complex to say the least) (Peugh, 2010; Roberts & Monaco, 2006). The conventional way is to divide a coefficient by the (pooled) within-school standard deviation (SD) (e.g. Von Secker & Lissitz, 1999). This is equivalent to Cohen's *d*. Common effect size measures such as η^2 (and thus Cohen's *d*) may underestimate effect size in HLM (Louwerse, Hutchinson, Tillman, & Recchia, 2015). The other caveat of using Cohen's *d* is that Cohen's classification of effect size as small, medium, and large may not be entirely applicable to large-scale analyses of survey data.
8. HLM partitions the total variance in science achievement into variance among students and variance among schools, allowing the calculation of proportion of variance attributable to

students and schools (e.g. ICC). Proportion of variance explained by a final model is a different issue. Typically, although variance attributable to schools can be very small (e.g. only 18% of the total variance in science achievement in our case), proportion of variance explained by a final model at the school level can be quite large if school-level variables are important predictors of the outcome variable (e.g. 78% of the variance among schools or at the school level in science achievement was explained in our case). On the other hand, although variance attributable to students can be very large (e.g. 82% of the total variance in science achievement in our case), proportion of variance explained by a final model at the student level can be quite small if student-level variables are not important predictors of the outcome variable (e.g. only 8% of the variance among students or at the student level in science achievement was explained in our case).

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Acosta, S. T., & Hsu, H. Y. (2014a). Negotiating diversity: An empirical investigation into Family, school and student factors influencing New Zealand adolescents' science literacy. *Educational Studies*, *40*(1), 98–115.

Acosta, S. T., & Hsu, H. Y. (2014b). Shared academic values: Testing a model of the association between Hong Kong parents' and adolescents' perception and the general value of science and scientific literacy. *Educational Studies*, *40*(2), 174–195.

Ainley, M., & Ainley, J. (2011). A cultural perspective on the structure of student interest in science. *International Journal of Science Education*, *33*(1), 51–71.

Anagün, S. S. (2011). The impact of teaching-learning process variables to students' scientific literacy levels based on PISA 2006 results. *Education and Science*, *36*(162), 84–102.

Anderson, L., & Chen, J. A. (2016). Do high-ability students disidentify with science? A descriptive study of U.S. ninth graders in 2009. *Science Education*, *100*(1), 57–77.

Areepattamannil, S. (2012a). Effects of inquiry-based science instruction on science achievement and interest in science: Evidence from Qatar. *The Journal of Educational Research*, *105*(2), 134–146.

Areepattamannil, S. (2012b). First- and second-generation immigrant adolescents' multidimensional mathematics and science self-concepts and their achievement in mathematics and science. *International Journal of Science and Mathematics Education*, *10*(3), 695–716.

Areepattamannil, S., Freeman, J. G., & Klinger, D. A. (2011). Influences of motivation, self-beliefs, and instructional practices on science achievement of adolescents in Canada. *Social Psychology of Education*, *14*(2), 233–259.

Areepattamannil, S., & Kaur, B. (2013). Factors predicting science achievement of immigrant and non-immigrant students: A multilevel analysis. *International Journal of Science and Mathematics Education*, *11*(5), 1183–1207.

Aydin, Y. Ç., & Uzuntiryaki, E. (2009). Development and psychometric evaluation of the high school chemistry self-efficacy scale. *Educational and Psychological Measurement*, *69*(5), 868–880.

Bhanot, R. T., & Jovanovic, J. (2009). The links between parent behaviors and boys' and girls' science achievement beliefs. *Applied Developmental Science*, *13*(1), 42–59.

Bouhlila, D. S. (2011). The quality of secondary school education in the Middle East and North Africa: What can we learn from TIMSS' results? *Compare*, *41*(3), 327–352.

Chen, J. A., & Usher, E. L. (2013). Profiles of the sources of science self-efficacy. *Learning and Individual Differences*, *24*, 11–21.

Chen, J. J. L. (2005). Relation of academic support from parents, teachers, and peers to Hong Kong adolescents' academic achievement: The mediating role of academic engagement. *Genetic, Social, and General Psychology Monographs*, *131*(2), 77–127.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Demanet, J., & Van Houtte, M. (2014). Social-ethnic school composition and disengagement: An inquiry into the perceived control explanation. *The Social Science Journal*, *51*(4), 659–675.

Else-Quest, N. M., Mineo, C. C., & Higgins, A. (2013). Math and science attitudes and achievement at the intersection of gender and ethnicity. *Psychology of Women Quarterly*, *37*(3), 293–309.

Farmer, H. S., Wardrop, J. L., & Rotella, S. C. (1999). Antecedent factors differentiating women and men in science/nonscience careers. *Psychology of Women Quarterly*, *23*(4), 763–780.

Fensham, P. J. (2009). Real world contexts in PISA science: Implications for context-based science education. *Journal of Research in Science Teaching*, *46*(8), 884–896.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109.

Fuchs, T., & Wößmann, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data. *Empirical Economics*, *32*(2–3), 433–464.

George, R. (2006). A cross-domain analysis of change in students' attitudes toward science and attitudes about the utility of science. *International Journal of Science Education*, *28*(6), 571–589.

Gerber, B. L., Cavallo, A. M. L., & Marek, E. A. (2001). Relationships among informal learning environments, teaching procedures and scientific reasoning ability. *International Journal of Science Education*, *23*(5), 535–549.

Hampden-Thompson, G., & Bennett, J. (2013). Science teaching and learning activities and students' engagement in science. *International Journal of Science Education*, *35*(8), 1325–1343.

Ho, E. S. C. (2010). Family influences on science learning among Hong Kong adolescents: What we learned from PISA. *International Journal of Science and Mathematics Education*, *8*(3), 409–428.

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, *56*, 517–543.

Jack, B. M., Lin, H. S., & Yore, L. D. (2014). The synergistic effect of affective factors on student learning outcomes. *Journal of Research in Science Teaching*, *51*(8), 1084–1101.

Jen, T. H., Lee, C. D., Chien, C. L., Hsu, Y. S., & Chien, K. M. (2013). Perceived social relationships and science learning outcomes for Taiwanese eighth graders: Structural equation modeling with a complex sampling consideration. *International Journal of Science and Mathematics Education*, *11*(3), 575–600.

Jocz, J. A., Zhai, J., & Tan, A. L. (2014). Inquiry learning in the Singaporean context: Factors affecting student interest in school science. *International Journal of Science Education*, *36*(15), 2596–2618.

Kalendar, I., & Berberoglu, G. (2009). An assessment of factors related to science achievement of Turkish students. *International Journal of Science Education*, *31*(10), 1379–1394.

Khalijah, M. S., Deraman, M., Omar, R., Othman, M. Y. H., Samat, S., Jumali, H., & Yatim, B. (1995). Perceptions of Malaysian secondary school science students towards physics. *Journal of Instructional Psychology*, *22*(3), 238–241.

Lam, S. F., Jimerson, S., Wong, B. P. H., Kikas, E., Shin, H., Veiga, F. H., … Zollneritsch, J. (2014). Understanding and measuring student engagement in school: The results of an international study from 12 countries. *School Psychology Quarterly*, *29*(2), 213–232.

Lam, T. Y. P., & Lau, K. C. (2014). Examining factors affecting science achievement of Hong Kong in PISA 2006 using hierarchical linear modeling. *International Journal of Science Education*, *36*(15), 2463–2480.

Larson, L. M., Stephen, A., Bonitz, V. S., & Wu, T. F. (2014). Predicting science achievement in India: Role of gender, self-efficacy, interests, and effort. *Journal of Career Assessment*, *22*(1), 89–101.

Lau, S., & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment*, *8*(2), 139–162.

Lavonen, J., & Laaksonen, S. (2009). Context of teaching and learning school science in Finland: Reflections on PISA 2006 results. *Journal of Research in Science Teaching*, *46*(8), 922–944.

Lee, J. S. (2014). The relationship between student engagement and academic performance: Is it a myth or reality? *The Journal of Educational Research*, *107*(3), 177–185.

Lin, H. S., Lawrenz, F., Lin, S. F., & Hong, Z. R. (2013). Relationships among affective factors and preferred engagement in science-related activities. *Public Understanding of Science*, *22*(8), 941–954.

Lin, T. J., Tan, A. L., & Tsai, C. C. (2013). A cross-cultural comparison of Singaporean and Taiwanese eighth graders' science learning self-efficacy from a multi-dimensional perspective. *International Journal of Science Education*, *35*(7), 1083–1109.

Louwerse, M. M., Hutchinson, S., Tillman, R., & Recchia, G. (2015). Effect size matters: The role of language statistics and perceptual simulation in conceptual processing. *Language, Cognition and Neuroscience*, *30*, 430–447.

Lupart, J. L., Cannon, E., & Telfer, J. A. (2004). Gender differences in adolescent academic achievement, interests, values and life-role expectations. *High Ability Studies*, *15*(1), 25–42.

Ma, V. J., & Ma, X. (2014). A comparative analysis of the relationship between learning styles and mathematics performance. *International Journal of STEM Education*, *1*(3), 1–13.

Ma, X., Jong, C., & Yuan, J. (2013). Exploring reasons for the East Asian success in PISA. In H. D. Meyer & A. Benavot (Eds.), *PISA, power, and policy: The emergence of global educational governance* (pp. 225–245). Oxford: Symposium Books.

Ma, X., Ma, L., & Bradley, K. (2008). Using multilevel modeling to investigate school effects. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 59–110). Charlotte, NC: Information Age.

Maynard Wang, M. N., Wu, K. C., & Iris Huang, T. C. (2007). A study on the factors affecting biological concept learning of junior high school students. *International Journal of Science Education*, *29*(4), 453–464.

McGough, J. J., & Faraone, S. V. (2009). Estimating the size of treatment effects: Moving beyond *p* values. *Psychiatry*, *6*(10), 21–29.

Mohammadpour, E. (2013). A three-level multilevel analysis of Singaporean eighth-graders science achievement. *Learning and Individual Differences*, *26*, 212–220.

Mujtaba, T., & Reiss, M. J. (2014). A survey of psychological, motivational, family and perceptions of physics education factors that explain 15-year-old students' aspirations to study physics in post-compulsory English schools. *International Journal of Science and Mathematics Education*, *12*(2), 371–393.

Ng, K. T., Lay, Y. F., Areepattamannil, S., Treagust, D. F., & Chandrasegaran, A. L. (2012). Relationship between affect and achievement in science and mathematics in Malaysia and Singapore. *Research in Science and Technological Education*, *30*(3), 225–237.

Organization for Economic Cooperation and Development. (2007a). *Database: PISA 2006*. Retrieved from http://pisa2006.acer.edu.au/

Organization for Economic Cooperation and Development. (2007b). *PISA 2006: Science competencies for tomorrow's world*. Retrieved from http://www.oecd.org/unitedstates/39722597.pdf

Organization for Economic Cooperation and Development. (2007c). *PISA 2006: Science competencies for tomorrow's world, executive summary*. Retrieved from http://www.oecd.org/pisa/pisaproducts/39725224.pdf

Organization for Economic Cooperation and Development. (2007d). *Student questionnaire: PISA 2006*. Retrieved from http://nces.ed.gov/surveys/pisa/pdf/quest_pisa_2006_student.pdf

Organization for Economic Cooperation and Development. (2007e). *School questionnaire: PISA 2006*. Retrieved from: http://nces.ed.gov/surveys/pisa/pdf/quest_pisa_2006_school.pdf.

Organization for Economic Cooperation and Development. (2009). *PISA 2006 technical report*. Retrieved from http://www.oecd.org/pisa/pisaproducts/42025182.pdf

Organization for Economic Cooperation and Development. (2016). *PISA 2015 results in focus*. Retrieved from https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf

Perera, L. D. H. (2014). Parents' attitudes towards science and their children's science achievement. *International Journal of Science Education*, *36*(18), 3021–3041.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, *48*, 85–112.

Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2012). *A user's guide to MLwiN, v2.26*. Bristol: Centre for Multilevel Modelling, University of Bristol.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2011). *HLM7 for windows*. Skokie, IL: Scientific Software International.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: SAGE.

Riegle-Crumb, C., Moore, C., & Ramos-Wada, A. (2011). Who wants to have a career in science or math? Exploring adolescents' future aspirations by gender and race/ethnicity. *Science Education*, *95*(3), 458–476.

Roberts, J. K., & Monaco, J. P. (April, 2006). *Effect size measures for the two- level linear multilevel model*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Robnett, R. R., & Leaper, C. (2013). Friendship groups, personal motivation, and gender in relation to high school students' STEM career interest. *Journal of Research on Adolescence*, *23*(4), 652–664.

Rozek, C. S., Hyde, J. S., Svoboda, R. C., Hulleman, C. S., & Harackiewicz, J. M. (2015). Gender differences in the effects of a utility-value intervention to help parents motivate adolescents in mathematics and science. *Journal of Educational Psychology*, *107*(1), 195–206.

Scherer, R. (2013). Further evidence on the structural relationship between academic self-concept and self-efficacy: On the effects of domain specificity. *Learning and Individual Differences*, *28*, 9–19.

Scherer, R., & Siddiq, F. (2015). The Big-Fish-Little-Pond-Effect revisited: Do different types of assessments matter? *Computers & Education*, *80*, 198–210.

Sha, L., Schunn, C., & Bathgate, M. (2015). Measuring choice to participate in optional science learning experiences during early adolescence. *Journal of Research in Science Teaching*, *52*(5), 686–709.

Shen, C. (2005). How American middle schools differ from schools of five Asian countries: Based on cross-national data from TIMSS 1999. *Educational Research and Evaluation*, *11*(2), 179–199.

Shin, J., Lee, H., McCarthy-Donovan, A., Hwang, H., Yim, S., & Seo, E. (2015). Home and motivational factors related to science-career pursuit: Gender differences and gender similarities. *International Journal of Science Education*, *37*(9), 1478–1503.

Shumow, L., Schmidt, J. A., & Zaleski, D. J. (2013). Multiple perspectives on student learning, engagement, and motivation in high school biology labs. *The High School Journal*, *96*(3), 232–252.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research*, *95*(6), 323–332.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, *69*(3), 493–525.

Spearman, J., & Watt, H. M. G. (2013). Perception shapes experience: The influence of actual and perceived classroom environment dimensions on girls' motivation for science. *Learning Environments Research*, *16*(2), 217–238.

Steinmayr, R., Dinger, F. C., & Spinath, B. (2012). Motivation as a mediator of social disparities in academic achievement. *European Journal of Personality*, *26*(3), 335–349.

Sun, L., Bradley, K. D., & Akers, K. (2012). A multilevel modeling approach to investigating factors impacting science achievement for secondary school students: PISA Hong Kong sample. *International Journal of Science Education*, *34*(14), 2107–2125.

Sungur, S. (2007). Modeling the relationships among students' motivational beliefs, metacognitive strategy use, and effort regulation. *Scandinavian Journal of Educational Research*, *51*(3), 315–326.

Swarat, S., Ortony, A., & Revelle, W. (2012). Activity matters: Understanding student interest in school science. *Journal of Research in Science Teaching*, *49*(4), 515–537.

Tang, M., & Neber, H. (2008). Motivation and self-regulated science learning in high-achieving students: Differences related to nation, gender, and grade-level. *High Ability Studies*, *19*(2), 103–116.

Thum, Y. M., & Bryk, A. S. (1997). Value-added productivity indicators: The Dallas system. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure* (pp. 100–109). Thousand Oaks, CA: Corwin.
Tighezza, M. (2014). Modeling relationships among learning, attitude, self-perception, and science achievement for grade 8 Saudi students. *International Journal of Science and Mathematics Education*, *12*(4), 721–740.
Tran, N. A. (2011). The relationship between students' connections to out-of-school experiences and factors associated with science learning. *International Journal of Science Education*, *33*(12), 1625–1651.
Tsai, L. T., & Yang, C. C. (2015). Hierarchical effects of school-, classroom-, and student-level factors on the science performance of eighth-grade Taiwanese students. *International Journal of Science Education*, *37*(8), 1166–1181.
Uitto, A. (2014). Interest, attitudes and self-efficacy beliefs explaining upper-secondary school students' orientations towards biology-related careers. *International Journal of Science and Mathematics Education*, *12*(6), 1425–1444.
Vandekerckhove, J., Matzke, D., & Wagenmakers, E. J. (2015). Model comparison and the principle of parsimony. In J. R. Busemeyer, Z. Townsend, & J. Wang (Eds.), *Oxford handbook of computational and mathematical psychology* (pp. 300–319). Oxford: Oxford University Press.
VanMeter-Adams, A., Frankenfeld, C. L., Bases, J., Espina, V., & Liotta, L. A. (2014). Students who demonstrate strong talent and interest in STEM are initially attracted to STEM through extracurricular experiences. *Life Sciences Education*, *13*(4), 687–697.
Vázquez Alonso, A., & Manassero Mas, M. A. (2009). The relevance of science education: Students' values and attitudes related to science and technology. *Enseñaza de las Ciencias*, *27*(1), 33–48.
Viljaranta, J., Nurmi, J. E., Aunola, K., & Salmela-Aro, K. (2009). The role of task values in adolescents' educational tracks: A person-oriented approach. *Journal of Research on Adolescence*, *19*(4), 786–798.
Von Secker, C. E., & Lissitz, R. W. (1999). Estimating the impact of instructional practices on student achievement in science. *Journal of Research in Science Teaching*, *36*(10), 1110–1126.
Wößmann, L. (2003). Schooling resources, educational institutions and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics*, *65*(2), 117–170.
Woods-McConney, A., Oliver, M. C., McConney, A., Maor, D., & Schibeci, R. (2013). Science engagement and literacy: A retrospective analysis for Indigenous and non-Indigenous students in Aotearoa New Zealand and Australia. *Research in Science Education*, *43*(1), 233–252.
Woods-McConney, A., Oliver, M. C., McConney, A., Schibeci, R., & Maor, D. (2014). Science engagement and literacy: A retrospective analysis for students in Canada and Australia. *International Journal of Science Education*, *36*(10), 1588–1608.
Yu, C. H. (2012). Examining the relationships among academic self-concept, instrumental motivation, and TIMSS 2007 science scores: A cross-cultural comparison of five East Asian countries/regions and the United States. *Educational Research and Evaluation*, *18*(8), 713–731.

Appendix A

PISA items descriptive of various aspects of science engagement.

Science self-efficacy
How easy do you think it would be for you to perform the following tasks on your own?
(a) Recognize the science question that underlies a newspaper report on a health issue. (b) Explain why earthquakes occur more frequently in some areas than in others. (c) Describe the role of antibiotics in the treatment of disease. (d) Identify the science question associated with the disposal of garbage. (e) Predict how changes to an environment will affect the survival of certain species. (f) Interpret the scientific information provided on the labelling of food items. (g) Discuss how new evidence can lead you to change your understanding about the possibility of life on Mars. (h) Identify the better of two possible explanations about the formation of acid rain. (1 = I could do this easily, 2 = I could do this with a bit of effort, 3 = I would struggle to do this on my own, 4 = I couldn't do this).
Science self-concept
How much do you agree with the statements below?
(a) Learning advanced science topics would be easy for me. (b) I can usually give good answers to test questions in science. (c) I learn science topics quickly. (d) Science is easy for me. (e) When I am being taught science, I can understand the concepts very well. (f) I can easily understand new ideas in science. (1 = strongly agree, 2 = agree, 3 = disagree, 4 = strongly disagree).
Enjoyment of science
How much do you agree with the statements below?
(a) I generally have fun when I am learning science topics. (b) I like reading about science. (c) I am happy doing science problems. (d) I enjoy acquiring new knowledge in science. (e) I am interested in learning about science. (1 = strongly agree, 2 = agree, 3 = disagree, 4 = strongly disagree).
General interest in learning science
How much interest do you have in learning about the following science topics?
(a) Topics in physics. (b) Topics in chemistry. (c) The biology of plants. (d) Human biology. (e) Topics in astronomy. (f) Topics in geology. (g) Ways scientists design experiments. (h) What is required for scientific explanations? (1 = high interest, 2 = medium interest, 3 = low interest, 4 = no interest.
Instrumental motivation for science
How much do you agree with the statements below?
(a) Making an effort in my science class is worth it because this will help me in the work I want to do later on. (b) What I learn in my science class is important for me because I need this for what I want to study later on. (c) I study science because I know it is useful for me. (d) Studying science is worthwhile for me because what I learn will improve my career prospects. (e) I will learn many things in my science class that will help me get a job. (1 = strongly agree, 2 = agree, 3 = disagree, 4 = strongly disagree).
Future-oriented science motivation
How much do you agree with the statements below?
(a) I would like to work in a career involving science. (b) I would like to study science after high school. (c) I would like to spend my life doing advanced science. (d) I would like to work on science projects as an adult. (1 = strongly agree, 2 = agree, 3 = disagree, 4 = strongly disagree).
General value of science
How much do you agree with the statements below?
(a) Advances in science and technology usually improve people's living conditions. (b) Science is important for helping us to understand the natural world. (c) Science and technology usually help improve the economy. (d) Science is valuable to society. (e) Advances in science and technology usually bring social benefits. (1 = strongly agree, 2 = agree, 3 = disagree, 4 = strongly disagree).
Personal value of science
How much do you agree with the statements below?
(a) Some concepts in science help me see how I relate to other people. (b) I will use science in many ways when I am an adult. (c) Science is very relevant to me. (d) I find that science helps me to understand the things around me. (e) When I leave school there will be many opportunities for me in science. (1 = strongly agree, 2 = agree, 3 = disagree, 4 = strongly disagree).
Science-related activities
How often do you do these things?
(a) Watch TV programmes about science. (b) Borrow or buy books on science topics. (c) Visit web sites on science topics. (d) Listen to radio programmes about advances in science. (e) Read science magazines or science articles in newspapers. (f) Attend a science club. (1 = very often, 2 = regularly, 3 = sometimes, 4 = never or hardly ever).

Sources: OECD (2007e, 2009)

Appendix B

PISA items descriptive of various dimensions of science teaching.

Focus on models or applications
When learning science topics at school, how often do the following activities occur?
(a) The teacher explains how a science idea can be applied to a number of different phenomena (e.g. the movement of objects, substances with similar properties. (b) The teacher uses science to help students understand the world outside school. (c) The teacher clearly explains the relevance of science concepts to our lives. (d) The teacher uses examples of technological application to show how science is relevant to society. (1 = never or hardly at all, 2 = in some lessons, 3 = in most lessons, 4 = in all lessons)
Hands-on activities
When learning science topics at school, how often do the following activities occur?
(a) Students spend time in the laboratory doing practical experiments. (b) Students are required to design how a science question could be investigated in a laboratory. (c) Students are asked to draw conclusions from an experiment they have conducted. (d) Students do experiments by following the instructions of the teacher. (1 = never or hardly at all, 2 = in some lessons, 3 = in most lessons, 4 = in all lessons).
Interaction
When learning science topics at school, how often do the following activities occur?
(a) Students are given opportunities to explain their own ideas. (b) The lessons involve students' opinions about the topics. (c) There is a class debate or discussion. (d) Students have discussions about the topics. (1 = never or hardly at all, 2 = in some lessons, 3 = in most lessons, 4 = in all lessons).
Student investigations
When learning science topics at school, how often do the following activities occur?
(a) Students are allowed to design their own experiments. (b) Students are given the chance to choose their own investigations. (c) Students are asked to do an investigation to test out their own ideas. (1 = never or hardly at all, 2 = in some lessons, 3 = in most lessons, 4 = in all lessons).

Sources: OECD (2007e, 2009)

Appendix C

Statistical results for baseline multilevel model of science achievement without aspects of science engagement.

	Effect	SE
Student-level variables		
Father socioeconomic status (SES)	.66	(.18)
Mother SES	.54	(.25)
Language at home (D)	30.49	(9.31)
School context variables		
School size	−.01	(.00)
School mean father SES	1.57	(.66)
School climate variables		
Curriculum & assessment	9.21	(4.54)
Hands-on activities	39.74	(15.03)
Student investigation	−53.46	(12.98)

Note: SE = standard error. D = dichotomous variable. All estimates are statistically significant at the alpha level of .05.