The mediation effect of student self-efficacy between teaching approaches and science achievement: findings from 2011 TIMSS US data

Su Gao[1] · Haiying Long[2] · Dan Li[3] · Luxi Yang[4]

Abstract

Over the past two decades, researchers consistently demonstrated the importance of science teaching approaches and student self-efficacy in influencing their science achievement. These findings have become the foundation of science education reform. However, empirical supports of these relationships are limited to direct relationships and small-scale studies. Therefore, little is known about the mechanism of how teaching approaches and student self-efficacy affect student achievement. In order to fill these gaps, this study used a multilevel structural equation modeling approach to analyze the direct and indirect relationships between teaching approaches, student self-efficacy, and science achievement by using the data of US eighth grade students in the 2011 TIMSS assessment. The results indicated that none of the teaching approaches identified in this study were directly associated with student science achievement, but significant mediation effect was found between generic teaching and student science achievement through student self-efficacy. Implications of these results for US educational system and reform were discussed.

Keywords Science teaching · Student self-efficacy · Science achievement

1 Introduction

The quality of students' science learning at middle-grade levels plays an important role in shaping their later science outcomes (Kwon and Lawson 2000; Jackson and Davis 2000), such as their decisions to enroll in science classes in high school, to choose a STEM-related college major, and finally to pursue a STEM career (Britner and Pajares 2006; Chen and Pajares 2010). These educational outcomes further affect economic prosperity, democratic process, and individual pursuit of equality and happiness in a nation (Quinn and Cooc 2015). However, U.S. middle-grade

✉ Su Gao
  Su.Gao@ucf.edu

Extended author information available on the last page of the article

students showed poor science performance in international assessments (Kastberg et al. 2016; Provasnik et al. 2016), which has become a national concern during the past several decades (National Science Board 2018). Researchers have paid special attention to this issue and suggested a multitude of factors that are related to it. How teachers teach science and how confident students are in learning science are two key factors identified by many researchers (Bandura et al. 2001; Fogleman et al. 2011; Høigaard et al. 2015; Sweller 2009).

Science teaching approaches are at the center of students' science learning quality (Fogleman et al. 2011). Although many teaching approaches are general to almost all subjects (Dole et al. 1991; Hattie et al. 1996), a set of approaches specific to science teaching are implemented in science classes. These approaches have often been categorized into traditional didactic teaching and practice-based science teaching (National Research Council 2012). Substantial research studies have been carried out to examine the relationship between these two main types of teaching and student achievement. However, the studies have corroborated the significant effects of both type of teaching on student achievement, thus, making inconclusive the question what kind of teaching approach is more effective for students' science learning (Cairns and Areepattamannil 2019; Furtak et al. 2012; Jerrim et al. 2019; Lazonder and Harmsen 2016; Minner et al. 2010; Rönnebeck et al. 2016). In addition, the mechanism of how teaching approaches affect students' science achievement has seldom been addressed in the literature. Researchers have suggested that students' academic self-efficacy may play a role in the relationship (Chen and Pajares 2010). However, this hypothesis has not yet been empirically examined.

As another key factor affecting students' science achievement, students' self-efficacy, has also been examined in plenty of studies (Aurah 2017; Bandura et al. 2001; Boz et al. 2016; Britner 2008; Høigaard et al. 2015; Schraw et al. 2006). However, many studies have focused on students in elementary schools (Griggs et al. 2013), high school (Britner 2008; Kupermintz 2002; Lau and Roeser 2002; Tsai et al. 2011), and college (Alt 2015; Bartimote-Aufflick et al. 2016; Bilgin et al. 2015; Gormally et al. 2009), and there have been comparatively fewer studies focusing on middle school students (Britner and Pajares 2001, 2006; Chen and Pajares 2010). Additionally, although extensive research has supported the significant relationships between self-efficacy and science achievement as well as between teaching approaches and science achievement, little empirical evidence has been obtained from real school environments and incorporated the hierarchical relationship between teachers and students (Jiang and McComas 2015). Very few studies collected large-scale and nationally representative data. What's more, previous studies have demonstrated that self-efficacy had a mediating effect on the relationship between previous and current academic achievement as well as learning environment and achievement (Diseth 2011; Moriarty et al. 1995). Nevertheless, the mediating effect has not been investigated on the relationship of between teaching approaches and academic achievement. Further, limited literature examining the relationship between science teaching approaches and student self-efficacy has also yielded mixed conclusions (Britner and Pajares 2001; Sungur and Tekkaya 2006).

Given these gaps in the literature, the current study aims to examine the direct and indirect relationships between teaching approaches, student self-efficacy in

science, and student science achievement by using US eighth-grade dataset from the 2011 TIMSS large-scale assessment. It addresses the following research questions:

1. How do teaching approaches directly associate with student science achievement?
2. How do teaching approaches directly associate with student self-efficacy?
3. How does student self-efficacy in science directly associate with student science achievement?
4. Does student self-efficacy in science mediate the relationship between teaching approaches and student science achievement?

2 Theoretical framework and literature review

2.1 Teaching approaches and student science achievement

Two theoretical lenses frame this study. The first one is the theories about how different teaching approaches influence student achievement. In this study, teaching approaches are categorized into generic teaching and science-domain specific teaching approaches (Russ et al. 2016). Generic teaching approaches refer to strategies designed to initiate, sustain, and foster students' cognitive engagement across different subject areas and grade levels. These approaches can lead to positive science learning outcomes (Azevedo 2015) by influencing students' attention allocation, metacognitive monitoring, positive emotions (Broughton et al. 2011), setting learning goals (Zimmerman 1990), and involvement in academic tasks (Kyriakides et al. 2013). For example, empirical studies showed that questioning strategy, providing or reinforcing objectives, directly influence achievement (Schroeder et al. 2007).

Science-domain specific teaching approaches reflect specific scientific processes (Bransford et al. 1999) and help students develop an understanding of scientific content, processes, and values at the same time (Southerland et al. 2007). Two different type of teaching approaches are currently found in most science classrooms (Windschitl and Calabrese Barton 2016). The first one draws on a practice-based view of teaching and of science itself. This practice-based science teaching includes activities such as designing, planning, conducting investigations, conducting experiments, and developing explanations based on observations. These science practices provide students with opportunities to apply reason to key concepts in science, participate in the discussion of science, and solve authentic problems (Sinatra et al. 2015; Windschitl and Calabrese Barton 2016). This notion has been directing science education reform in the US and has helped shift the focus of science teaching from didactic teaching to the approaches that engages students in science practices (NGSS Lead States 2013; NRC 1996). At the same time, traditional didactic science teaching is often seen as a popular, yet ineffective, science teaching method as it transmits science facts to students mainly through teachers' lecture and students' drill and practices in order to memorize factual knowledge (Crawford et al. 2005; Smerdon et al. 1999).

Although there have been many discussions about these two types of science teaching in the literature, studies that examined the effects of different science teaching approaches on student achievement led to mixed results. In a study by Wilson, Taylor, Kowalski, and Carlson (2010), researchers randomly assigned 58 14–16-year-old students to either a science practices group or a control group. In the science practices group, students were exposed to learning activities, such as designing and implementing investigations, analyzing and interpreting data, constructing explanations and designing solutions, and evaluating scientific arguments. In the control group, students received traditional didactic instruction. Students in the science practices group performed significantly better than those in the control group on science achievements. This result is consistent with another study (Taraban et al. 2007) that assessed the achievement of 408 high-school students from six classrooms in Texas. Among the six classrooms, some were characterized with science practices focusing on lab activities, while others were traditional teaching classrooms with the characteristics of teachers' direct transmission of information, whole-class activities, and "cookbook" experiments. However, these results are challenged by a study conducted by Wolf and Fraser (2007), which involved two science teachers and 165 7th grade students in eight classes located in Long Island, New York. The researchers did not find a significant difference between the effects of these two teaching strategies on student science achievement. Four classes of students who were engaged in science practices, including developing and planning their own investigations, working collectively toward unique goals, and sharing individual findings through conducting several laboratory experiments based on the physical science curriculum, scored slightly higher but not statistically significant than another four classes of students that followed a set of procedures provided by teachers and completed a findings' worksheet. Similarly, Glasson (1989) found that ninth-graders from a classroom focusing on science practices did not outperform their peers from a classroom using direct demonstration for instruction. A limitation of these above-cited studies is that student samples were mostly selected from limited geographic areas, which cannot represent students from the whole nation. The results from some large datasets, such as Programme for International Student Assessment (PISA), indicated that science practice-based teaching was negatively related to science achievements of adolescents in Canada (Areepattamannil et al. 2011), Finland (Lavonen and Laaksonen 2009), England (Jerrim et al. 2019), and even across 54 countries (Cairns and Areepattamannil 2019). The current study intends to join this ongoing conversation by examining the relationship between student science achievement and science teaching approaches based on a nationally representative US sample collected in the 2011 TIMSS assessment.

2.2 Self-efficacy in science

The second theory guiding this study is that student self-efficacy has a positive impact on academic achievement (Bandura 1997). According to Bandura (1997), self-efficacy in human functioning, like people's other motivational and affective states and actions, is "based more on what they believe than on what is objectively

true'' (p. 2). A wealth of research has supported the positive role of student self-efficacy in their achievement, performance, and other outcomes in different subjects (Artino 2012; Greene 2015; Greene and Miller 1996; Miller et al. 1996; Linnenbrink and Pintrich 2003), including in science (Andrew 1998; Britner and Pajares 2006; Chen and Usher 2013; Jansen et al. 2015; Luzzo et al. 1999; Pajares et al. 2000; Yazici et al. 2011).

Previous studies have demonstrated that students' self-efficacy in science is affected by student demographic backgrounds, such as gender, race, grade, student beliefs, such as implicit beliefs of ability, epistemic beliefs, beliefs of learning science (Britner and Pajares 2001, 2006; Chen and Usher 2013; Griggs et al. 2013; Tsai et al. 2011), and mastery experiences, one of the primary sources of self-efficacy (Bandura 1997; Britner and Pajares 2006; Luzzo et al. 1999). Past research has also shown that student self-efficacy in learning is influenced by how academic tasks are provided, how students are encouraged to do their work, and other teaching approaches, such as recognizing students' effort, providing opportunities for improvement, focusing on individual learning progress, help students improve their self-efficacy in learning (Ames 1992; Ryan et al. 1998).

In addition, a few studies also revealed that student self-efficacy improved through innovative teaching or curricula, such as engaging students in authentic inquiry-oriented science investigations (Britner and Pajares 2001), game-based learning (Meluso et al. 2012), the embedment of technology (Liu et al. 2011), and the use of animation (van der Meij et al. 2015). One study implementing the *Responsive Classroom* approach focusing on social-emotional learning reported significant increase in student self-efficacy and decrease in anxiety in math and science (Griggs et al. 2013).

However, the literature is limited in the number of studies that examine the effects of teaching practices on student self-efficacy in science. These studies are also limited methodologically because they have employed mostly experiments or quasi-experiments and reported the effects of the interventions on the outcomes through statistical analyses of mean differences. What is lacking in the literature is probably a study that takes into account hierarchies existing between teachers and students and examines the relationships through a modelling approach. Furthermore, the limited number of empirical studies also yielded mixed results about the relationship between teaching approaches and self-efficacy. For instance, a study found that more inquiry-oriented science investigations provide middle school students with mastery experiences that are necessary to the development of strong science self-efficacy (Britner and Pajares 2001). However, another study did not find the same results in a 6-week problem-based teaching session with high school students in a biology class (Sungur and Tekkaya 2006).

Despite the connections between student self-efficacy in science and science achievement as well as between teaching approaches and student self-efficacy, there were surprisingly few empirical studies that have put these three aspects together in one study and examined their direct and indirect relationships. However, self-efficacy was corroborated in a few studies as a mediator (Diseth 2011; Moriarty et al. 1995). For instance, in one experimental study that examined the effects of learning environments on achievement and behavior, researchers found that competitive learning environments

enhanced students' performance but this effect was difficult to maintain if students had a low self-efficacy (Moriarty et al. 1995). A more recent study showed that the relationship between students' past and current academic performance was mediated by self-efficacy (Diseth 2011). Further, past studies have found a mediating effect of factors similar to self-efficacy, such as positive affect, on the relationship of teaching practices and students' science achievement (Long 2016).

2.3 Hypotheses

Based on previous literature review, we hypothesized:

1. Student self-efficacy in science has a significant, direct, and positive effect on student science achievement.
2. Three types of teaching approaches (one generic teaching approach and two science-domain specific teaching approaches) have significant direct effects on student science achievement. Generic teaching approaches and science practice-based teaching have positive effects while traditional didactic science teaching has negative effect.
3. Three types of teaching approaches have significant direct effects on student self-efficacy in science. Generic teaching approach and science practice-based teaching approach have positive effects and traditional didactic science teaching has negative effects.
4. Three types of teaching approaches have significant indirect effects on student science achievement through student self-efficacy in science.

3 Methods

3.1 Data source and sample

In this study, we used the dataset from the 2011 Trends in International Mathematics and Science Study (TIMSS), which is the fifth cycle of the assessment. The data were collected from 10,382 8th—grade students and 865 teachers in the U.S. They were selected through the approach of two-stage stratified cluster sampling with probabilities proportional to size (PPS), which is the sampling design consistently used in international large-scale assessments, such as TIMSS and PISA. Specifically, schools were first randomly selected with PPS at the first stage and teachers and students in one or more classes were randomly selected at the second stage. The use of this sampling approach ensures the randomness and representativeness of the sample, thus increases the generalizability of the results (Foy et al. 2013).

3.2 Measures

3.2.1 Student science achievement

Student science achievement was an endogenous/outcome variable in this study. It was measured by test items about science content and cognitive domains. Four content domains were included in the assessment and each domain contained a few topics. The domain and its proportion accounting for in the assessment were as follows: biology (35%), chemistry (20%), physics (25%), and earth science (20%). Three cognitive domains were also included in the assessment. They were: knowing facts, procedures, and concepts; applying knowledge and conceptual understanding in real-life problems; and reasoning with unfamiliar situations, complex context, and multi-step problems. A multi-matrix design was employed to collect students' science achievement, in which each student received only a subset of test items out of a test pool (Martin and Mullis 2012a, b). After the assessment data were collected, an item response theory approach was used to scaling student science assessment for each student, which was indicated by a set of five plausible values (Martin and Mullis 2012a, b). These values are randomly selected from a distribution of achievement scores that approximated the student true ability (Martin and Mullis 2012a, b) (see Table 1). Student final science achievement score used in the study was the average of the five plausible values.

3.2.2 Student self-efficacy in science

We treated student self-efficacy in science both as an endogenous and exogenous variable. Three items in student questionnaire measured this construct, including "I usually do well in science", "I learn things quickly in science", and "I am good at working out difficult science problems". A 1–4 Likert scale was used, where 1 referred to "Agree a lot" and 4"Disagree a lot". These three items were selected because they were strong expressions of student self-efficacy in learning science. All the items were reverse coded so that a higher value means a higher degree of agreement. The reliability among the items estimated by Cronbach's α was.82, which was high (see Table 1).

3.2.3 Teachers' generic teaching

We treated teachers' generic teaching as another exogenous variable in the study. Four items in teacher questionnaire measured this construct. They were "Summarize what students should have learned from the lesson", "Relate the lesson to students' daily lives", "Use questioning to elicit reasons and explanations" and "Bring interesting materials to class" based on a 1–4 Likert scale (1-"Every or almost every lesson", 4-"Never"). All the items were reverse coded so that a

Table 1 Means and standard deviations of items

Variables	Items	Mean (SD)
Science performance		523.88 (77.11)
Student self-efficacy ($\alpha=.82$)	I usually do well in science (Y_1)	3.28 (.83)
	I learn things quickly in science (Y_2)	2.95 (.93)
	I am good at working out difficult science problems (Y_3)	2.68 (.98)
Science practice-based teaching ($\alpha=.72$)	Observe natural phenomena and describe what they see (Y_4)	2.82 (.77)
	Design or plan experiments or investigations (Y_5)	2.47 (.73)
	Conduct experiments or investigations (Y_6)	2.73 (.69)
	Relate what they are learning in science to their daily lives (Y_7)	3.37 (.74)
Traditional didactic science teaching ($\alpha=.69$)	Read their textbooks or other resource materials (Y_8)	2.70 (.85)
	Have students memorize facts and principles (Y_9)	2.40 (.82)
	Take a written test or quiz (Y_{10})	2.60 (.76)
Generic teaching ($\alpha=.53$)	Summarize what students should have learned from the lesson (Y_{11})	3.69 (.60)
	Relate the lesson to students' daily lives (Y_{12})	3.57 (.61)
	Use questioning to elicit reasons and explanations (Y_{13})	3.75 (.53)
	Bring interesting materials to class (Y_{14})	3.23 (.71)
SES		2.14 (.53)
Homework frequency		3.04 (1.12)

Valid N=5757

higher value means a higher frequency. The reliability among the items estimated by Cronbach's α was .53, which was acceptable (see Table 1).

3.2.4　Teachers' science practice-based teaching

We treated science practice-based teaching as an exogenous variable. Four items in teacher questionnaire measured this construct. They were "Observe natural phenomena and describe what they see", "Design or plan experiments or investigations", "Conduct experiments or investigations", and "Relate what they are learning in science to their daily lives". These items were representations of the most popular inquiry-based teaching practices in science. They were measured on a 1–4 Likert scale, where 1 referred to "Every or almost every lesson" and 4 "Never". All the items were reverse coded so that a higher value means a higher frequency. The reliability among the items estimated by Cronbach's α was .72, which was adequate (see Table 1).

3.2.5　Teachers' traditional didactic science teaching

We treated teachers' traditional didactic science teaching as an exogenous variable. Three items in the teacher questionnaire measured this construct. They were "Read their textbooks or other resource materials", "Have students memorize facts and principles", and "Take a written test or quiz". A 1–4 Likert scale was used, where 1 referred to "Every or almost every lesson" and 4 "Never". All the items were reverse coded so that a higher value means a higher frequency. The reliability among the items estimated by Cronbach's alpha was .69, which was very close to adequate (see Table 1).

3.2.6　Control variables

We selected student socioeconomic status (SES) and the frequency of science homework as control variables. SES has long been considered one of the most important variables that affect student achievement and should be treated as a control variable (Bornstein and Bradley 2003; Bradley and Corwyn 2002; Coley 2002; Long and Pang 2017; Milne and Plourde 2006). Although whether homework positively or negatively affects student achievement is still controversial, it was suggested to be one representation of Opportunity to Learn (OTL) (Schmidt et al. 2015) and has been found to be a significant predictor of student achievement (Cooper and Robinson 2006; Leone and Richards 1989; Maltese et al. 2012; Trautwein 2007). In addition, Trautwein (2007) differentiated the effects of homework time, homework frequency, and homework effort on student achievement and found that homework frequency, rather than the other two, was a significant predictor of achievement.

In this study, student SES was measured by the home educational resources index provided by TIMSS in a student questionnaire. It was based on a 1–3 Likert scale (1-"Many resources", 3-"Few resources"). The frequency of homework time was measured by students' response to one item "How often teacher give you homework

in science?" on a 1–5 Likert scale (1-"Every day", 5-"Never"). All the items were reverse coded so that a higher value means more resources and a higher frequency.

3.3 Data analysis

3.3.1 Preliminary analysis

Due to the complex sampling designs of the TIMSS, we employed several statistical techniques to increase the precision of the measures of the variables (Rutkowski et al. 2010; Rutkowski et al. 2014). Weights were first applied to all the variables in IBM SPSS Statistics (Version 24). Variables at the student level were weighted by total student weight and variables at the teacher level were weighted by science teacher weight. Then missing values were examined and handled. The dataset had a missing rate ranging from 1.78 to 4.02% for student variables and ranging from 27.4 to 38.0% for teacher variables and no missing value was found on the plausible values of student science achievement. Multiple Imputation (MI) approach was used in combination with two-level multilevel structural equation modeling (MSEM) to handle missing values in Mplus 8.3 and five imputations were made for each variable (Enders et al. 2015).

3.3.2 Modelling approach

In order to test the hypotheses regarding the direct and indirect effects of the exogenous variables on endogenous variable at both teacher- and student-level, we applied traditional multilevel structural equation modelling (MSEM) approach. Like Hierarchical Linear Modeling (HLM), MSEM can account for the variances at varying hierarchical levels. However, MSEM has advantages over HLM in the following three aspects. First, the approach is a synthesis of multilevel modeling and structural equation modeling (Morin et al. 2014; Rabe-Hesketh et al. 2007). Second, the approach can control both measurement and sampling errors, thus providing more accurate estimates of structural relationships among variables (Muthén and Asparouhov 2009). Third, the approach can conduct mediation analysis and further illustrate the direct and indirect effects of the variables at different levels on the endogenous variables (Preacher et al. 2010; Trusz 2018; Wendorf 2002).

MSEM analyses were conducted using the multilevel structural equation module in Mplus 8.3 (Muthén and Muthén 2010). To perform MSEM analyses, we followed the procedure proposed by Hayduk (1987). Measurement model was first set up to examine if measures of the variables were consistent with the theoretical constructs (see Fig. 1). Then a path model was used to test the hypothesized relationships among variables at the two levels (see Fig. 2). At the student level, two control variables, student SES and homework frequency, and student self-efficacy were included in the model as predictors of student science achievement. At the teacher level, the same two control variables, student self-efficacy, and three types of teaching approaches were included in the model as predictors of students' science achievement. At the same time, student self-efficacy was used as a mediator. In the
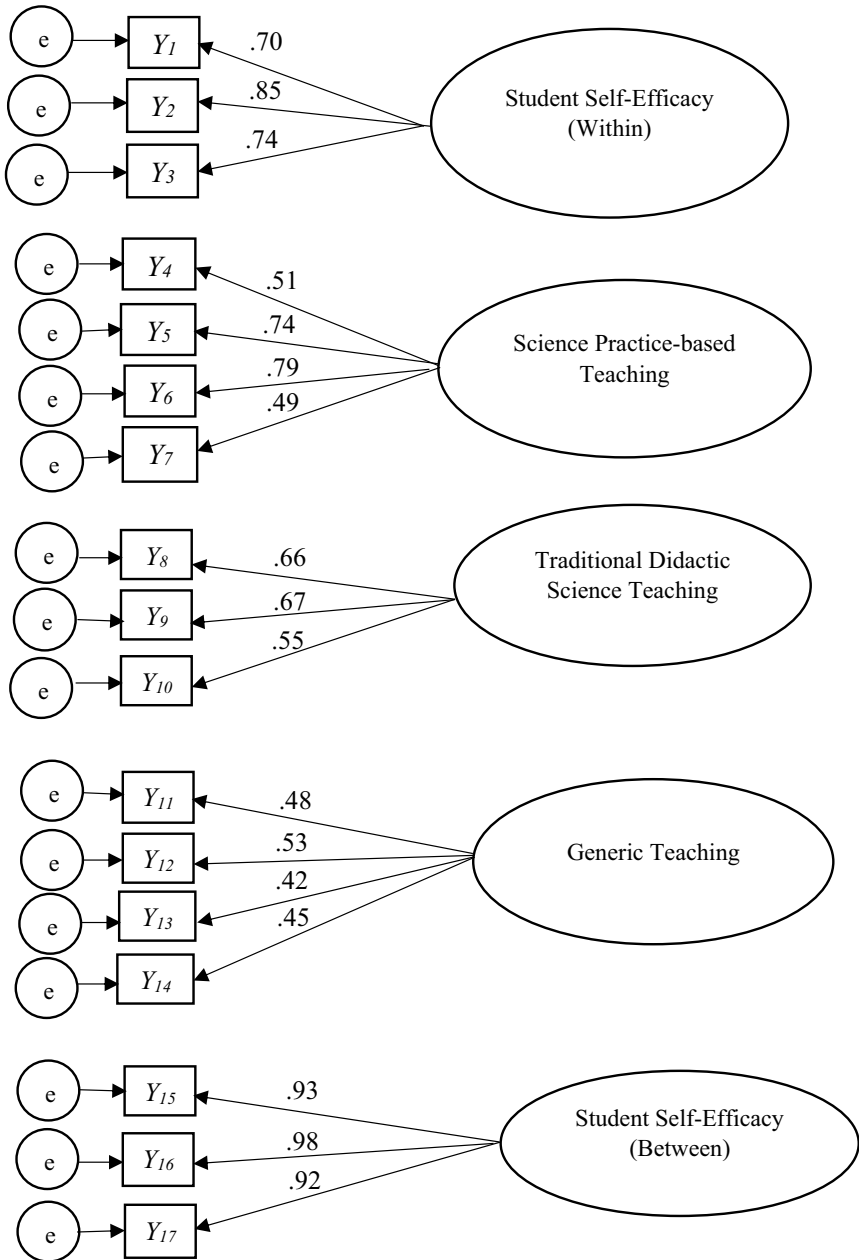
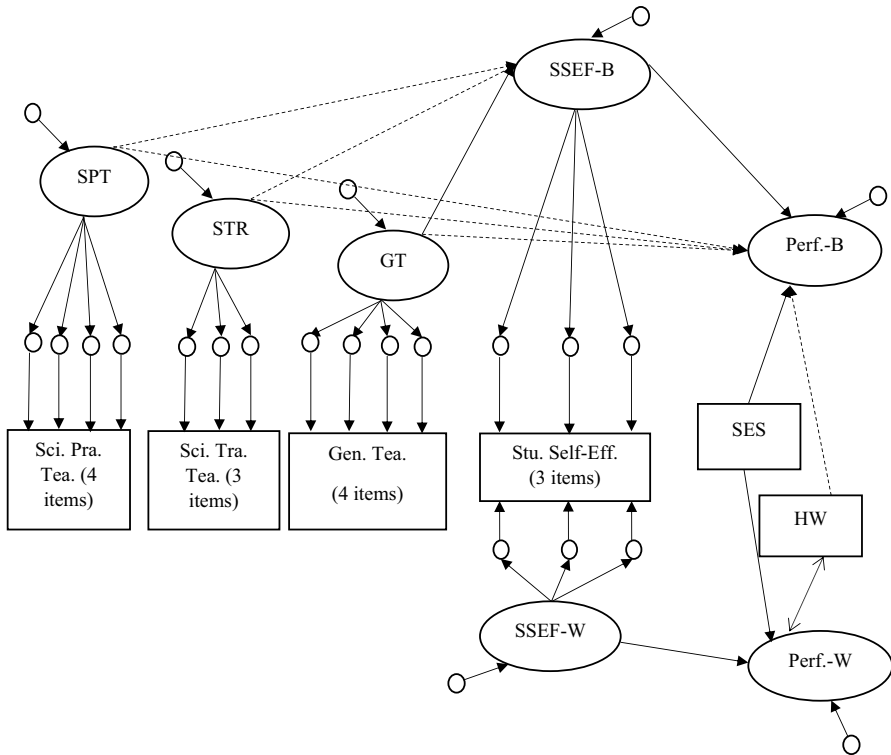**Fig. 1** Measurement model. All the factor loadings are above the line

Fig. 2 Multilevel structural equation model. Solid and single-arrow lines refer to significant and positive paths, solid and double-arrow lines refer to significant and negative paths, and dotted lines refer to insignificant paths. Perf.-B = Students' Science Performance at Between Level; Perf.-W = Students' Science Performance at Within Level; SSEF-B = Students' Self-Efficacy at Between Level; SSEF-W = Students' Self-Efficacy at Within Level; SPT = Teachers' Science Practice-based Teaching; STR = Teachers' Traditional Didactic Science Teaching; GT = Teachers' Generic Teaching; SES = Students' Socioeconomic Status; HW = Students' Frequency of Homework

hypothesized model, endogenous variables (i.e. three types of teaching approaches) are level-2 variables, mediator (i.e. student self-efficacy) is a level-1 variable and exogenous variable (i.e. student science achievement) is also a level-1 variable. Therefore, the model is a 2-1-1 MSEM model (Preacher et al. 2010).

Mplus uses the Chi square test, Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Standardized Root Mean Square Residual (SRMR) as the model fit indices (Muthén and Muthén 2010). According to previous studies (Hox and Bechger 1999; Hu and Bentler 1999; Schermelleh-Engel et al. 2003), CFI and TLI with a value of .90 or greater suggests an acceptable fit, and a value of .95 or greater suggests a good fit. RMSEA with a value of .06 or smaller suggests a good fit. SRMR with a value of .08 or smaller indicates acceptable fit. In the process of understanding the indirect effects, we focus on two points recommended by Rucker et al. (2011): the statistical significance of indirect effect and the size of the effect.

4 Results

4.1 Descriptive statistics and correlations

US students' average score in science assessment was 523.88 ($SD=77.11$), which was close to the score reported in the 2011 TIMSS (i.e. 525). Among the three items measuring students' self-efficacy, students rated the item "I usually do well in science" ($M=3.28$, $SD=.83$) higher than the other two items. This means that more students believe they do well in science than their ability of learning things quickly in science and solving difficult science problems. Among the three types of teaching, teachers reported more generic teaching than science practice-based teaching and traditional didactic science teaching. For the four science practice-based teaching approaches, teachers reported higher frequency of the practice "Relating what they are learning in science to their daily lives" ($M=3.37$, $SD=.74$) than the other three practices. The practice "Design or plan experiments or investigations" was reported with the least frequency ($M=2.47$, $SD=.73$). The three traditional didactic science teaching approaches were reported with similar frequency, with the approach of "Read their textbooks or other resources" having the most frequency ($M=2.70$, $SD=.85$). The teachers reported high frequency of four generic teaching approaches: "Use questioning to elicit reasons and explanations" ($M=3.75$, $SD=.53$), "Summarize what students should have learned from the lesson" ($M=3.69$, $SD=.60$), "Relating the lesson to students' daily lives" ($M=3.57$, $SD=.61$), and "Brining interesting materials to class" ($M=3.23$, $SD=.71$). The mean of students' SES indicator is 2.14 ($SD=.53$), indicating that American students' overall home educational resources were just above average. The average of students' report of the frequency of science homework was 3.04 ($SD=1.12$), meaning that American students' science homework frequency was about 3 or 4 times a week (see Table 1).

Overall, the correlations among all the items were not high, ranging from .00 to .65. The three items measuring student self-efficacy in science had high correlations (i.e. .63, .55, and .65). A few items measuring teachers' science practice-based teaching approaches had high correlations, such as the correlation between "Design or plan experiments or investigations" and "Conduct experiments or investigation" ($r=.65$). Two items measuring teachers' traditional didactic science teaching had a comparatively high correlation: "Read their textbooks or other resource materials" and "Have students memorize facts and principles" ($r=.48$). In addition, one item measuring teachers' generic teaching, "Relate the lesson to students' daily lives", and an item measuring teachers' science practice-based teaching, "Relate what they are learning in science to their daily lives" had a correlation of. 54 (see Table 2).

4.2 Measurement model

We conducted a Confirmatory Factor Analysis (CFA) in Mplus to test the measurement model before the test of MSEM. The results showed that $\chi^2 (71)=582.84$, RMSEA was .03, CFI was .96, TLI was .94, SRMR at the within level was .01,

Table 2 Zero-order correlations among variables

VA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
SSEF																
Y_1	–															
Y_2	.63**	–														
Y_3	.55**	.65**	–													
SPT																
Y_4	.02	.01	.01	–												
Y_5	−.01	−.03*	−.02	.37**	–											
Y_6	.05**	.02	.00	.39**	.65**	–										
Y_7	.02	.00	.01	.36**	.28**	.34**	–									
STR																
Y_8	−.01	−.02	−.03*	.05**	.09**	.08**	.14**	–								
Y_9	−.01	−.02	−.02	.10**	.10**	−.00	.14**	.48**	–							
Y_{10}	.01	−.02	−.03*	.11**	.18**	.18**	.19**	.42**	.38**	–						
GT																
Y_{11}	.02	−.00	−.01	.22**	.15**	.21**	.19**	.07**	.05**	.07**	–					
Y_{12}	.04**	.03**	.02	.23**	.08**	.21**	.54**	.05**	.01	.04**	.19**	–				
Y_{13}	.05**	.02	.05**	.19**	.10**	.19**	.24**	.03**	.02	.01	.27**	.25**	–			
Y_{14}	.06**	.03**	.03**	.19**	.23**	.24**	.24**	−.01	−.08**	.05**	.24**	.23**	.17**	–		
CON																
SES	.18**	.17**	.16**	.06**	.05**	.09**	.02	−.04**	−.07**	.−05**	.05**	.03*	.08**	.07**	–	
HWF	−.03**	.01	.04**	−.03*	−.05**	−.04**	−.01	.14**	.08**	−.00	−.01	−.01	.01	−.02	.02**	–

The specific item indicated by Yn refers to Table 1

VA = Variables; SSEF = Students' Self-Efficacy; SPT = Teachers' Science Practice-Based Teaching; STR = Teachers' Traditional Didactic Science Teaching; GT = Teachers' Generic Teaching; CON = Control Variables; SES = Students' Socioeconomic Status; HWF = Students' Homework Frequency

** $p < .001$; * $p < .05$

and SRMR at the between level was .07. These indicate an excellent fit between the model and the data. All the factor loadings in the measurement model were statistically significant at $p < .001$ and most of the loadings ranged from .50 to .98. The factor loadings for the three items measuring students' self-efficacy at the within level were .70 ("I usually do well in science"), .85 ("I learn things quickly in science"), and .74 ("I am good at working out difficult science problems"), respectively. Two items measuring teachers' science practice-based teaching had high factor loadings ($\lambda = .74$ for "Design or plan experiments or investigation" and $\lambda = .79$ for "Conduct experiments or investigations"). Another two items measuring teachers' traditional didactic science teaching had similar factor loadings ($\lambda = .66$ for "Read their textbooks or other resource materials" and $\lambda = .67$ for "Have students memorize facts and principles"). One item measuring teachers' generic teaching had high factor loading ($\lambda = .53$ for "Relate the lesson to students' daily lives") and the factor loadings of the other three items were close to .50. All the three items measuring student self-efficacy at the between level loaded strongly on the latent variable and all the factor loadings were higher than .90, with the factor loading of "I learn things quickly in science" being .98 (see Fig. 1).

4.3 Multilevel structural equation modeling

The results of the final multi-level structural equation modeling (MSEM) analysis showed that χ^2 (117) = 1073.53, RMSEA was .03, CFI was .93, TLI was .91, SRMR at the within level was .05, and SRMR at the between level was .08. All these values indicate an adequate fit between the model and the data.

4.3.1 Relationships at the within level

At the within level, student SES and science homework frequency were statistically significant predictors of student science achievement ($p < .001$). Student SES was a positive predictor, whereas student homework frequency was negative. This suggests that students who have more home educational resources are more likely to perform better in science. But students who have more science homework are less likely to have higher science achievement. After controlling student SES and homework frequency, student self-efficacy was still a significant and positive predictor of science achievement ($p = .00$) (see Fig. 2). This supports our hypothesis 1 at the within level. As shown in Table 3, student self-efficacy in science had a higher coefficient ($\beta = .20$, $SE = .01$) than SES ($\beta = .08$, $SE = .01$) and homework frequency ($\beta = -.04$, $SE = .01$). This means that student self-efficacy in science was a better predictor of student science achievement than the two control variables at the student level.

4.3.2 Relationships at the between level

At the between level, student SES was a significant and positive predictor of student science achievement ($p = .00$), but homework frequency was nonsignificant ($p = .61$). This suggests that students with more home educational resources taught

Table 3 Path standardized coefficients and standard error of MSEM model

Paths	Coefficient	SE
Within level		
SES-> Science Performance	.08**	.01
Homework Frequency-> Science Performance	−.04**	.01
Self-Efficacy-> Science Performance	.20**	.01
Between level		
SES-> Science Performance	.72**	.03
Homework Frequency-> Science Performance	−.02	.03
Self-Efficacy-> Science Performance	.31**	.05
Science Practice-Based Teaching-> Science Performance	.07	.08
Science Traditional Didactic Teaching-> Science Performance	−.03	.04
Generic Teaching-> Science Performance	−.01	.08
Science Practice-Based Teaching-> Self-Efficacy	−.12	.10
Science Traditional Didactic Teaching-> Self-efficacy	−.05	.08
Generic Teaching-> Self-efficacy	.25*	.10
Indirect effect		
Science Practice-Based Teaching-> Self-efficacy-> Science Performance	−.04	.04
Science Traditional Didactic Teaching-> Self-efficacy-> Science Performance	.01	.02
Generic Teaching-> Self-efficacy-> Science Performance	.08*	.04

Multiple imputations were used to replace missing data, parameter estimates are based on the average of imputated datasets

** $p < .001$; * $p < .05$

by the same teacher also performed better in science assessment than those with less resources. However, student homework frequency did not affect their science achievement at the group level. After controlling the two variables, student self-efficacy in science was still a significant and positive predictor ($p = .00$). This shows that students with higher self-efficacy in science still outperformed those with lower self-efficacy in science achievement even though they were taught by the same teacher. It also supports our hypothesis 1 at the between level. Comparatively speaking, students' SES ($\beta = .72$, $SE = .03$) had a higher coefficient than student self-efficacy in science ($\beta = .31$, $SE = .05$) (see Table 3). This suggests that student SES accounted for a larger variance of science achievement than student self-efficacy in science at the between level.

However, the three types of teaching approaches did not have significant direct effects on student science achievement ($p = .42$ for science practice-based teaching, $p = .50$ for traditional didactic science teaching, and $p = .97$ for generic teaching). This indicates that students with a teacher using any of these three types of teaching approaches did not directly affect student science achievement. These results did not support our hypothesis 2.

Interestingly, among the three types of teaching, only generic teaching was a significant predictor of student self-efficacy in science ($p = .02$). This indicates that students with a teacher using more generic teaching are more likely to have a higher

self-efficacy, but frequencies of using science-domain specific teaching do not affect student self-efficacy in science. The standardized coefficient of generic teaching was .25 ($SE = .10$). The results partially supported our hypothesis 3.

Further, no significant indirect effects were found for two types of science-domain specific teaching approaches (science practice-based teaching and traditional didactic science teaching) on student science achievement through student self-efficacy in science ($p = .25$ for science inquiry-based practices, $p = .58$ for science traditional practices). But a significant indirect effect was found for the generic teaching on student science achievement through the self-efficacy in science ($p = .04$). This shows that students with a teacher using more generic teaching approaches are more likely to have a higher science achievement through the effect of their self-efficacy in science. But the coefficient of the indirect effect is comparatively small ($\beta = .08$, $SE = .04$). These results partially supported our hypothesis 4.

5 Discussion

5.1 Contributions

This study sought to examine the direct and indirect relationships between teaching approaches, student self-efficacy, and science achievement with US eighth-grade students' data collected from the 2011 TIMSS assessment. It makes the following contributions to the current literature. First, it examines the relationships among the three variables by using a large-scale, nationally representative sample, through a multi-level structural modeling approach that takes into account the hierarchical relations between teachers and students. This provides the literature with valid and generalizable study results. Second, it found the internal mechanism of how generic teaching approaches affect student science achievement through the mediating effect of student self-efficacy in science. Third, it provides empirical evidence to support the differences between generic teaching approaches and science-domain teaching approaches and their dissimilar effects on students' self-efficacy in science and science achievement. Fourth, it shows that self-efficacy in science is a stronger predictor of student science achievement than other variables, corroborating the importance of students' self-theories in learning in a general sense.

5.2 Discussion

This study found that none of the three types of teacher-reported teaching practices (generic teaching, science practice-based teaching, and traditional didactic teaching) has significant direct effects on science achievement. This result is not consistent with some of previous empirical studies that identified a positive relationship between generic teaching approaches, such as questioning and providing or reinforcing objectives, and student academic achievement (Ames 1992; Dole et al. 1991; Hattie et al. 1996; Schroeder et al. 2007). However, this result is consistent with other recent empirical studies using the data of TIMSS and PISA, another

international large data set, that reported no significant or even negative relationship between practice-based science teaching approaches with student science achievement (Cairns and Areepattamannil 2019; Jerrim et al. 2019; Jiang and McComas 2015; Long 2016). Therefore, this finding empirically challenges a major assumption of science education reform again, which proposes to improve student science achievement by shifting the focus of science teaching from traditional didactic teaching to engaging students in science practices (NGSS Lead States 2013; NRC 1996).

This lack of significant effects also suggests that implementing a complex science activity to achieve meaningful goals in the classroom might require supports from resources beyond and above teaching. The literature showing positive relationships between practice-based science teaching and student achievement (Wilson et al. 2010; Taraban et al. 2007) usually involved effective professional development, curriculum materials, and specific practice guides to support teacher and student learning (Windschitl and Calabrese Barton 2016). Therefore, more factors that potentially support teachers to implement science practice-based teaching need to be further examined in future studies.

The current study also showed that only generic teaching was a significant predictor of student self-efficacy in science, but both science-domain specific science teaching (practice-based and didactic) were not. This finding contributes to our understanding of how regular teaching approaches impact on student self-efficacy in science by providing empirical evidence based on a large-scale, nationally representative dataset (Cheung 2015). It also adds more complexity to the current literature by agreeing with some scholars (Britner and Pajares 2001) but disagreeing with others (Sungur and Tekkaya 2006). It may also suggest that science-specific teaching approaches are related to students' science self-efficacy in science through other student variables, such as peer pressure within science practices (Gao and Wang 2014; Wang and Lin 2005) and students' cultural values. These variables can be examined in the future study.

Another important finding in this study is that student self-efficacy in science was significantly and positively associated with science achievement. This finding further supports theories from Bandura (1997) and Greene (2015) and adds to existent empirical studies in science education field (Aurah 2017; Boz et al. 2016; Britner 2008; Chen and Pajares 2010). It also encourages science teachers to continue to improve students' self-efficacy in science and further help students improve their science achievement (Britner 2008; Britner and Pajares 2001, 2006; House 2008; Lavonen and Laaksonen 2009). However, further research is needed to scrutinize the mechanisms behind this relationship and identify what other factors mediate the relationship, such as the time or efforts individual students spent in science learning (Schmidt et al. 2018).

This study also reported a surprising and significant indirect association between teaching practices and science achievement through student self-efficacy in science, even though generic teaching was not significantly, directly associated with achievement. This finding empirically corroborated the mediating role of self-efficacy in science between teaching approaches and science achievement and supported that student self-efficacy in science is influenced by generic teaching approaches, such as encouraging students to do their work (Azevedo 2015). This finding also suggested

that the reason science teaching approaches did not increase student science achievement was because these approaches suppress the improvement of students' academic self-efficacy (Chen and Pajares 2010). Further research, such as qualitative studies, could be conducted to explore how and why this suppression happens.

This study further found a positive effect of one control variable, student SES, on science achievement and a negative effect of another control variable, science homework frequency, at level 1. It is in line with previous studies that have consistently identified SES as a significant factor affecting student science achievement, including studies using large-scale datasets (Long and Pang 2017; Maltese et al. 2012). Researchers have also indicated that the association between homework frequency and achievement vary with the measures and the level of analysis chosen in the study, which was referred to as chameleon effects (Trautwein et al. 2009). However, it remains unclear whether teaching approaches interact with student SES and how. Therefore, in the future, it is important to examine how teaching approaches are related with non-teaching factors, such as social, economic, cultural, and historical contexts, in which teaching and curriculum practices are situated (Berliner 2009; Sykes et al. 2010).

5.3 Limitations

This study provides interesting insights into the current literature. However, it is not without limitations. First, the 2011 TMISS used in the study is a secondary database. Teacher and student participants in this dataset were asked to self-report the frequencies of different teaching activities in science classroom through questionnaires. The lack of observations in teaching activities could affect interpretation about how three type of teaching approaches were regularly used in the classrooms. Therefore, the findings of this study need to be verified and further examined in future studies based on systematic observations and other approaches, such as interviews. Second, only some components of generic teaching, science practice-based teaching, and traditional didactic science teaching approaches were surveyed in the instruments. Other components of the teaching approaches may be unrepresented and need to be explored using a more complete instruments in the future. Third, although the hierarchical relations between teachers and students were considered in this study, the causal inferences about the relationships among the three variables should be interpreted with caution. Finally, this study only examined direct and indirect relationships among three variables. In the future, the relationships of more variables and moderation or mediated moderation effects will be examined.

6 Conclusion

Using US eighth-grade data obtained from the 2011 TIMSS, this study focuses on examining two important theoretical assumptions that guided science education reform during the past two decades. It provides empirical evidence for not only the relationships between teaching approaches, student self-efficacy in science, and

science achievement, but also the mediation effect of student self-efficacy in science between teaching approaches and science achievement using a MSEM approach. Our finding indicated a significant mediation effect of student self-efficacy on the relationship between generic teaching practices and student science achievement even though none of the teaching approaches identified in this study were directly associated with student science achievement. This suggests that simply changing science teaching and curriculum in the reform of science education may not improve student science achievement. Rather, the reform needs to pay substantial attention to the complex relationships between teaching approaches and student science achievement, especially how teaching approaches influence student self-beliefs.

Compliance with ethical standards

Conflict of interest The authors declares that they have no conflict of interest.

Ethical approval Public data from 'Trends in International Mathematics and Science Study (TIMSS) 2011' is used in the manuscript.

References

Alt, D. (2015). Assessing the contribution of a constructivist learning environment to academic self-efficacy in higher education. *Learning Environments Research, 18*(1), 47–67.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271.

Andrew, S. (1998). Self-efficacy as a predictor of academic performance in science. *Journal of Advanced Nursing, 27,* 596–603.

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

Artino, A. R., Jr. (2012). Academic self-efficacy: From educational theory to instructional practice. *Perspectives in Medical Education, 1,* 76–85. https://doi.org/10.1007/s40037-012-0012-5.

Aurah, C. (2017). Investigating the relationship between science self-efficacy beliefs, gender, and academic achievement, among high school students in Kenya. *Journal of Education and Practice, 8*(8), 146–153.

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

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

Bandura, A., Caprara, G. V., Barbaranelli, C., Pastorelli, C., & Regali, C. (2001). Sociocognitive self-regulatory mechanisms governing transgressive behavior. *Journal of Personality and Social Psychology, 80,* 125–135.

Bartimote-Aufflick, K., Bridgeman, A., Walker, R., Sharma, M., & Smith, L. (2016). The study, evaluation, and improvement of university student self-efficacy. *Studies in Higher Education, 41*(11), 1918–1942.

Berliner, D. C. (2009). *Poverty and potential: Out-of-school factors and school success*. Boulder and Tempe: Education and the public interest center & education policy research unit. Retrieved November 9, 2019, from http://epicpolicy.org/publication/poverty-and-potential.

Bilgin, I., Karakuyu, Y., & Ay, Y. (2015). The effects of project based learning on undergraduate students' achievement and self-efficacy beliefs towards science teaching. *Eurasia Journal of Mathematics, Science & Technology Education, 11*(3), 469–477.

Bornstein, M. H., & Bradley, R. H. (Eds.). (2003). *Socioeconomic status, parenting, and child development*. Mahwah, NJ: Lawrence Erlbaum.

Boz, Y., Yerdelen-Damar, S., Aydemir, N., & Aydemir, M. (2016). Investigating the relationships among students' self-efficacy beliefs, their perceptions of classroom learning environment, gender, and chemistry achievement through structural equation modeling. *Research in Science & Technological Education, 34*(3), 307–324.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53,* 371–399.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, minds, expeirence, and school*. Washington, DC: National Academy Press.

Britner, S. L. (2008). Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes. *Journal of Research in Science Teaching, 45*(8), 955–970.

Britner, S. L., & Pajares, F. (2001). Self-efficacy beliefs, race, and gender in middle school science. *Journal of Women and Minorities in Science and Engineering, 7,* 271–285.

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.

Broughton, S. H., Sinatra, G. M., & Nussbaum, E. M. (2011). "Pluto has been a planet my whole life!" Emotions, attitudes, and conceptual change in elementary students' learning about Pluto's reclassification. *Research in Science Education, 43*(2), 529–550.

Cairns, D., & Areepattamannil, S. (2019). Exploring the relations of inquiry-based teaching to science achievement and dispositions in 54 countries. *Research in Science Education, 49*(1), 1–23.

Chen, J. A., & Pajares, F. (2010). Implicit theories of ability of Grade 6 science students: Relation to epistemological beliefs and academic motivation and achievement in science. *Contemporary Educational Psychology, 35*(1), 75–87.

Chen, J. A., & Usher, E. (2013). Profiles of the sources of science self-efficacy. *Learning and Individual Differences, 24,* 11–21. https://doi.org/10.1016/j.lindif.2012.11.002.

Cheung, D. (2015). The combined effects of classroom teaching and learning strategy use on students' chemistry self-efficacy. *Research in Science Education, 45*(1), 101–116.

Coley, R. J. (2002). *An uneven start: Indicators of inequality in school readiness*. Princeton, NJ: ETS.

Cooper, H., & Robinson, J. C. (2006). Does homework improve academic achievement? A synthesis of research, 1987-2003. *Review of Educational Research, 76,* 1–62.

Crawford, B., Zembal-Saul, C., Munford, D., & Friedrichsen, P. (2005). Confronting prospective teachers' ideas of evolution and scientific inquiry using technology and inquiry based tasks. *Journal of Research in Science Teaching, 42*(6), 613–637.

Diseth, A. (2011). Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. *Learning and Individual Differences, 21,* 191–195. https://doi.org/10.1016/j.lindif.2011.01.003.

Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research, 61*(2), 239–264.

Enders, C. K., Mistler, S. A., & Keller, B. T. (2015). Multilevel multiple imputation: A review and evaluation of joint modeling and chained equations imputation. *Psychological Methods, 21*(2), 222–240.

Fogleman, J., McNeill, K. L., & Krajcik, J. (2011). Examining the effect of teachers' adaptations of a middle school science inquiry-oriented curriculum unit on student learning. *Journal of Research in Science Teaching, 48*(2), 149–169.

Foy, P., Arora, A., & Stanco, G. M. (2013). *TIMSS 2011 user guide for the international database: Supplement 4 TIMSS 2011 sampling stratification information*. Boston, MA: International Association for the Evaluation of Education.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of educational research, 82*(3), 300–329.

Gao, S., & Wang, J. (2014). Teaching transformation under centralized curriculum and teacher learning community: Two Chinese chemistry teachers' experiences in developing inquiry-based instruction. *Teaching and Teacher Education, 44*(1), 1–11.

Glasson, G. E. (1989). The effects of hands-on and teacher demonstration laboratory methods on science achievement in relation to reasoning ability and prior knowledge. *Journal of Research in Science Teaching, 26*(2), 121–131.

Gormally, C., Brickman, P., Hallar, B., & Armstrong, N. (2009). Effects of inquiry-based learning on students' science literacy skills and confidence. *International Journal Scholarship of Teaching and Learning, 3,* 16.

Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist, 50*(1), 14–30.

Greene, B. A., & Miller, R. B. (1996). Influences on course achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology, 21,* 181–192.

Griggs, M. S., Rimm-Kaufman, S. E., Merritt, E. G., & Patton, C. L. (2013). The responsive classroom approach and fifth grade students' math and science anxiety and self-efficacy. *School Psychology Quarterly, 28*(4), 360–373. https://doi.org/10.1037/spq0000026.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*(2), 99–136.

Hayduk, L. A. (1987). Structural equation modeling with LISREL: Essentials and advances. Baltimore and London: The Johns Hopkins University Press.

Høigaard, R., Kovač, V. B., Øverby, N. C., & Haugen, T. (2015). Academic self-efficacy mediates the effects of school psychological climate on academic achievement. *School Psychology Quarterly, 30*(1), 64.

House, J. D. (2008). Effects of classroom instructional strategies and self-beliefs on science achievement of elementary-school students in Japan: Results from the TIMSS 2003 assessment. *Education, 129*(2), 259–266.

Hox, J. J., & Bechger, T. M. (1999). An introduction to structural equation modeling. *Family Science Review, 11,* 354–373.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Jackson, A., & Davis, G. (2000). *Turning points 2000: Educating adolescents in the 21st century*. New York: Teachers College Press.

Jansen, M., Scherer, R., & Schroeders, U. (2015). Students' self-concept and self-efficacy in the science: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology, 41,* 13–24. https://doi.org/10.1016/j.cedpsych.2014.11.002.

Jerrim, J., Oliver, M., & Sims, S. (2019). The relationship between inquiry-based teaching and students' achievement. New evidence from a longitudinal PISA study in England. *Learning and Instruction, 61,* 35–44.

Jiang, F., & McComas, W. F. (2015). The effects of inquiry teaching on student science achievement and attitudes: Evidence from propensity analysis of PISA data. *International Journal of Science Education, 37,* 554–576. https://doi.org/10.1080/09500693.2014.1000426.

Kastberg, D., Chan, J. Y., & Murray, G. (2016). *Performance of U.S. 15-year-old students in science, reading, and mathematics literacy in an international context: First look at PISA 2015 (NCES 2017-048). U.S. Department of Education*. Washington, DC: National Center for Education Statistics.

Kupermintz, H. (2002). Affective and conative factors as aptitude resources in high school science achievement. *Educational Assessment, 8,* 123–137.

Kwon, Y., & Lawson, A. (2000). Linking brain growth with the development of scientific reasoning ability and conceptual change during adolescence. *Journal of Research in Science Teaching, 37*(1), 44–62.

Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education, 36*(1), 143–152.

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

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research, 86*(3), 681–718.

Leone, C. M., & Richards, H. (1989). Classwork and homework in early adolescence: The ecology of achievement. *Journal of Youth and Adolescence, 18,* 531–548.

Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs in student engagement and learning in classroom. *Reading and Writing Quarterly, 19,* 119–137.

Liu, M., Horton, L., Olmanson, J., & Toprac, P. (2011). A study of learning and motivation in a new media enriched environment for middle school science. *Educational Technology Research Development, 59,* 249–265. https://doi.org/10.1007/s11423-011-9192-7.

Long, H. (2016). The suppression role of positive affect on students' science achievement in East Asia: The example of Taipei. *Social Psychology of Education, 19,* 815–842. https://doi.org/10.1007/s11218-016-9358-6.

Long, H., & Pang, W. (2017). Family socioeconomic status, parental expectations, and adolescents' academic achievements: A case of China. *Educational Research and Evaluation, 22,* 283–304.

Luzzo, D. A., Hasper, P., Albert, K. A., Bibby, M. A., & Martinelli, E. A., Jr. (1999). Effects of self-efficacy-enhancing interventions on the math/science self-efficacy and career interests, goals, and actionsn of career undecided college students. *Journal of Counseling Psychology, 46*(2), 233–243.

Maltese, A. V., Tai, R. H., & Fan, X. (2012). When is homework worth the time?: Evaluating the association between homework and achievement in high school science and math. *The High School Journal, 96,* 52–72.

Martin, M., & Mullis, I. (2012a). *Creating and interpreting the TIMSS and PIRLS 2011 context questionnaire scales*. Retrieved November 8, 2019, from http://www.timssandpirls.bc.edu/methods/t-context-q-scales.html#.

Martin, M. O., & Mullis, I. V. S. (Eds.). (2012b). *Methods and procedures in TIMSS and PIRLS*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Meluso, A., Zheng, M., Spires, H. A., & Lester, J. (2012). Enhancing 5th graders' science content knowledge and self-efficacy through game-based learning. *Computers & Education*, *59*(2), 497–504.

Miller, R. B., Greene, B. A., Montalvo, G. P., Ravindran, B., & Nichols, J. D. (1996). Engagement in academic work: The role of learning goals, future consequences, pleasing others and perceived ability. *Contemporary Educational Psychology, 21,* 388–422. https://doi.org/10.1006/ceps.1996.0028.

Milne, A., & Plourde, L. A. (2006). Factors of a low-SES household: What aids academic achievement? *Journal of Instructional Psychology, 33,* 183–193.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—What is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching, 47*(4), 474–496.

Moriarty, B., Douglas, G., Punch, K., & Hattie, J. (1995). The importance o self-efficacy as a mediating variable between learning and environments and achievement. *British Journal of Educational Psychology, 65,* 73–84.

Morin, A. J. S., Marsh, H. W., Nagengast, B., & Scalas, F. L. (2014). Doubly latent multilevel analyses of classroom climate: An illustration. *The Journal of Experimental Education, 82,* 143–167.

Muthén, B. O., & Asparouhov, T. (2009). Beyond multilevel regression modeling: Multilevel analysis in a general latent variable framework. In J. Hox & J. K. Roberts (Eds.), *Handbook of advanced multilevel analysis*. New York, NY: Routledge.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles: Muthén & Muthén.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

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

National Science Board (2018). *2018 Science and engineering indicators*. Retrieved November 8, 2019, from https://nsf.gov/statistics/2018/nsb20181/assets/nsb20181.pdf.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

Pajares, F., Britner, S. L., & Valiante, G. (2000). Relation between achievement goals and self-beliefs of middle school students in writing and science. *Contemporary Educational Psychology, 25*(4), 406–422.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233.

Provasnik, S., Malley, L., Stephens, M., Landeros, K., Perkins, R., & Tang, J. H. (2016). *Highlights from TIMSS and TIMSS advanced 2015: Mathematics and science achievement of U.S. students in grades 4 and 8 and in advanced courses at the end of high school in an international context (NCES 2017-002)*. Washington, DC: NCES, IES, U.S. Department of Education.

Quinn, D. M., & Cooc, N. (2015). Science achievement gaps by gender and race/ethnicity in elementary and middle school: Trends and predictors. *Educational Researcher, 44*(6), 336–346.

Rabe-Hesketh, S., Skrondal, A., & Zheng, X. (2007). Multilevel structural equation modeling. In S. Lee (Ed.), *Handbook of latent variable and related models* (pp. 209–227). Netherlands: North-Holland.

Rönnebeck, S., Bernholt, S., & Ropohl, M. (2016). Searching for a common ground—A literature review of empirical research on scientific inquiry activities. *Studies in Science Education, 52*(2), 161–197.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass, 5*(6), 359–371.

Russ, R. S., Sherin, B. L., & Sherin, M. G. (2016). What constitutes teacher learning. In D. Gitomer & C. Bell (Eds.), *Handbook of research on teaching* (pp. 391–438). Washington, DC: American Educational Research Association.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher, 39,* 142–151. https://doi.org/10.3102/0013189X10363170.

Rutkowski, L., von Davier, M., & Rutkowski, D. (Eds.). (2014). *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. Boca Raton, FL: CRC Press.

Ryan, A. M., Gheen, M. H., & Midgley, C. (1998). Why do some students avoid asking for help? An examination of the interplay among students' academic efficacy, teachers' social–emotional role, and the classroom goal structure. *Journal of Educational Psychology, 90*(3), 528–535.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23–74.

Schmidt, J. A., Rosenberg, J. M., & Beymer, P. N. (2018). A person-in-context approach to student engagement in science: Examining learning activities and choice. *Journal of Research in Science Teaching, 55*(1), 19–43.

Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The role of schooling in perpetuating educational inequality: An international perspective. *Educational Researcher*, *44*(7), 371–386.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education, 36*(1–2), 111–139.

Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T., & Lee, Y. (2007). A meta-analysis of national research : Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching, 44*(10), 1436–1460.

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.

Smerdon, B. A., Burkam, D. T., & Lee, V. E. (1999). Access to constructivist and didactic teaching: Who gets it? Where is it practiced? *Teachers College Record, 101*(1), 5–34.

Southerland, S. A., Smith, L. K., Sowell, S., & Kittleson, J. (2007). Resisting unlearning: Understanding science education's response to the United State's national accountability movement. *Review of Research in Education, 31*(1), 45–77.

Sungur, S., & Tekkaya, C. (2006). Effects of problem-based learning and traditional instruction on self-regulated learning. *The Journal of Educational Research, 99*(5), 307–320.

Sweller, J. (2009). What human cognitive architecture tells us about constructivism. In S. Tobias & T. M. Duffy (Eds.), *Constructivist instruction: Success or failure?* (pp. 127–143). New York, NY: Routledge.

Sykes, G., Bird, T., & Kennedy, M. (2010). Teacher education: Its problems and some prospects. *Journal of Teacher Education*, *61*(5), 464–476.

Taraban, R., Box, C., Myers, R., Pollard, R., & Bowen, C. W. (2007). Effects of active-learning experiences on achievement, attitudes, and behaviors in high school biology. *Journal of Research in Science Teaching, 44*(7), 960–979.

Trautwein, U. (2007). The homework-achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort. *Learning and Instruction, 17,* 372–388.

Trautwein, U., Schnyder, I., Niggli, A., Neumann, M., & Ludtke, O. (2009). Chameleon effects in homework research: The homework-achievement association depends on the measures used

and the level of analysis chosen. *Contemporary Educational Psychology, 34,* 77–88. https://doi.org/10.1016/j.cedpsych.2008.09.001.

Trusz, S. (2018). Four mediation models of teacher expectancy effects on students' outcomes in mathematics and literacy. *Social Psychology of Education, 21*(2), 257–287.

Tsai, C.-C., Ho, H. N. J., Liang, J.-C., & Lin, H.-M. (2011). Scientific epistemic beliefs, conceptions of learning science and self-efficacy of learning science among high school students. *Learning and Instruction, 21,* 757–769. https://doi.org/10.1016/j.learninstruc.2011.05.002.

van der Meij, H., van der Meij, J., & Harmsen, R. (2015). Animated pedagogical agents effects on enhancing student motivation and learning in a science inquiry learning environment. *Education Technology Research Development, 63,* 381–403. https://doi.org/10.1007/s11423-015-9378-5.

Wang, J., & Lin, E. (2005). Comparative studies on US and Chinese mathematics learning and the implications for standards-based mathematics teaching reform. *Educational Researcher, 34*(5), 3–13.

Wendorf, C. A. (2002). Comparison of structural equation modeling and hierarchical linear modeling approaches to couples' data. *Structural Equation Modeling, 9,* 126–140. https://doi.org/10.1207/S15328007SEM0901_7.

Wilson, C. D., Taylor, Ja, Kowalski, S. M., & Carlson, J. (2010). The relative effects and equity of inquiry-based and commonplace science teaching on students' knowledge, reasoning, and argumentation. *Journal of Research in Science Teaching, 47*(3), 279–301.

Windschitl, M., & Calabrese Barton, A. (2016). Rigor and equity by design: Locating a set of core teaching practices for the science education community. In D. Gitomer & C. Bell (Eds.), *Handbook of research on teaching* (pp. 1099–1158). Washington, DC: American Educational Research Association.

Wolf, S. J., & Fraser, B. J. (2007). Learning environment, attitudes and achievement among middle-school science students using inquiry-based laboratory activities. *Research in Science Education, 38*(3), 321–341.

Yazici, H., Seyis, S., & Altun, F. (2011). Emotional intelligence and self-efficacy beliefs as predictors of academic achievement among high school students. *Procedia-Social and Behavioral Sciences, 15,* 2319–2323.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17.

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

Dr. Su Gao is an assistant professor of School of Teacher Education at the Universitiy of Central Florida. She is experienced in science teaching and science teacher education at the national and international level. Dr. Gao's major research interests include: (1) exploring teaching strategies and science performance of diverse students in an international context; (2) preparing and supporting preservice and inservice science teachers to teach all students.

Dr. Haiying Long is an associate professor in the Department of Counseling, Recreation, and School Psychology in the School of Education and Human Development at Florida International University, U.S.A. Her research interests include quantitative methodology, international large-scale assessment, cross-cultural comparison, STEM education, motivation, creativity, and academic achievement.

Dr. Dan Li is an Education Program Consultant at Iowa Department of Education. She does research in primary education and secondary education, institutional research, concurrent enrollment, and Career and Technical Education.

Luxi Yang is an Associate Professor in the School of Foreign Languages and Literature at Chongqing Normal University in China. Her research interests include intercultural studies, STEM education, and TESOL.

Affiliations

Su Gao[1] · Haiying Long[2] · Dan Li[3] · Luxi Yang[4]

Haiying Long
hylong@fiu.edu

Dan Li
lidan0111@gmail.com

Luxi Yang
ylx1024@gmail.com

[1]   School of Teacher Education, College of Community Innovation and Education, P.O. Box 161250, Orlando, FL 32816-1250, USA

[2]   Department of Counseling, Recreation, and School Psychology, School of Education and Human Development, College of Arts, Sciences and Education, Florida International University, 11200 SW 8th St., Miami, FL 33199, USA

[3]   Iowa Department of Education, Grimes Building, 400 E 14th Street, Des Moines, IA 50319, USA

[4]   School of Foreign Languages and Literature, Chongqing Normal University, No. 12 Tianchen Rd., Shapingba District, Chongqing 400047, China