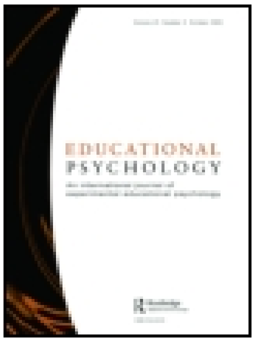

Educational Psychology

An International Journal of Experimental Educational Psychology

ISSN: 0144-3410 (Print) 1469-5820 (Online) Journal homepage: https://www.tandfonline.com/loi/cedp20

Student learning emotions in middle school mathematics classrooms: investigating associations with dialogic instructional practices

Alyssa Parr, Jamie Amemiya & Ming-Te Wang

To cite this article: Alyssa Parr, Jamie Amemiya & Ming-Te Wang (2019): Student learning emotions in middle school mathematics classrooms: investigating associations with dialogic instructional practices, Educational Psychology, DOI: 10.1080/01443410.2018.1560395

To link to this article: https://doi.org/10.1080/01443410.2018.1560395

Published online: 18 Feb 2019.

EDUCATIONAL PSYCHOLOGY
https://doi.org/10.1080/01443410.2018.1560395

Student learning emotions in middle school mathematics classrooms: investigating associations with dialogic instructional practices

Alyssa Parr[a], Jamie Amemiya[b] and Ming-Te Wang[a,b,c]

[a]School of Education, University of Pittsburgh, Pittsburgh, USA; [b]Department of Psychology, University of Pittsburgh, Pittsburgh, USA; [c]Learning Research and Development Center, University of Pittsburgh, Pittsburgh, USA

ABSTRACT

Emotions are central to how students experience mathematics, yet we know little about how specific instructional practices relate to students' emotions in mathematics learning. We examined how dialogic instruction, a socially dynamic form of instruction, was associated with four learning emotions in mathematics: enjoyment, pride, anger, and boredom. We also examined whether these associations differed by student gender and prior mathematics achievement. The sample consisted of 1307 sixth through eighth grade students (51.6% female, 59.0% White, 30.8% African American, and 10.3% other race; 42.3% receive free/reduced price lunch) from 70 mathematics classrooms. Results indicated that teachers who used more dialogic mathematics instruction had students who reported more enjoyment and pride, and less anger and boredom. Males and low-achieving students reported more positive and fewer negative emotions with greater dialogic instruction compared to their female and high-achieving counterparts.

ARTICLE HISTORY
Received 27 March 2018
Accepted 9 December 2018

KEYWORDS
Adolescence; learning emotions; control-value theory; mathematics instruction

For decades, education researchers and practitioners have sought to determine the best teaching strategies for promoting students' positive learning experiences in mathematics. This endeavour has resulted in an ongoing debate over the use of direct versus dialogic instructional practices (Tyack & Cuban, 1995). Though not necessarily opposite ends of a spectrum, these practices are fundamentally different approaches to mathematics instruction (Munter, Stein, & Smith, 2015). Direct instructional practices involve lecture-style instruction that focuses on memorising and mastering specific step-by-step mathematical procedures (Kirschner, Sweller, & Clark, 2006), while dialogic instructional practices seek to give students greater intellectual authority in the classroom and emphasise conceptual understanding and mastery of content through discourse (Fulmer & Turner, 2014; Grady, Watkins, & Montalvo, 2012). Although math teachers have traditionally used direct instructional practices, education researchers

CONTACT Alyssa Parr akp41@pitt.edu Department of Psychology in Education, University of Pittsburgh, 4200 5th Ave., Pittsburgh, PA 15260, USA

and practitioners have argued about the drawbacks of viewing the student as a passive absorber of information (i.e. direct approach) and the benefits of treating the student as an independent agent who actively constructs their knowledge (i.e. dialogic approach; Munter et al., 2015).

Understanding the effects of dialogic instructional practices has become of recent import as prominent aspects of mathematics teaching reform in the United States emphasise dialogic instructional techniques (NGAC & CCSSO, 2010). Literature examining the effectiveness of dialogic instruction in promoting students' positive learning experiences in mathematics exists (e.g. Hanze & Berger, 2007; Sawyer, 2004); however, this work is limited in scope as it largely focuses on academic achievement and motivation outcomes. To date, little research has explored the implications of dialogic instructional practices for students' learning emotions, which are key precursors to student motivation and achievement outcomes in science, technology, engineering, and mathematics (STEM) fields (Bailey, Taasooobshirazi, & Carr, 2014; Wang & Degol, 2013). Given that studies have shown gender and achievement differences in the effect of instruction on student learning outcomes (Fredricks, Hofkens, Wang, Mortenson, & Scott, 2017; Fulmer & Turner, 2014; Wang, 2012; Wang & Degol, 2016a), it would be informative to examine whether the effect of instructional practices on student learning emotions is consistent across male and female students as well as high and low achievers.

In this study, we seek to better understand the implications of dialogic instructional practices for adolescents' learning experiences in mathematics by examining how these instructional practices are associated with four learning emotions: enjoyment, pride, anger, and boredom. We also examine how the relationships between dialogic instruction and learning emotions differ by students' gender and prior mathematics achievement.

Control-value theory

We used Pekrun's (2000) control-value theory of learning emotions to conceptualise the links between instructional practices and student learning emotions. This theory posits that a student's classroom environment—including teacher instructional practices—shapes students' achievement-related control and value appraisals and these appraisals are antecedents of students' emotions in learning contexts or *learning emotions*. Students' *control appraisals* consist of competence-related judgments, such as expectancies and attributions for success or failure. A student may make stronger control appraisals when they have opportunities to check their work and comprehension by collaborating with peers or working in small groups with a teacher. Students' *value appraisals* refer to the perceived importance of achieving success or avoiding failure. A student may make stronger value appraisals in mathematics when their teacher highlights the practical value of a lesson or when the student is intrinsically interested in a topic. Together, students' control and value appraisals result in a variety of learning emotions, including enjoyment, pride, anger, and boredom (Pekrun, 2000).

Pekrun (2006) further described four instructional practices that shape the control and value appraisals contributing to students' learning emotions: cooperative goal structures, autonomy and collaboration, instructional quality, and value induction.

Cooperative goal structures—learning environments that promote positive outcomes for the whole class rather than only certain individuals—foster positive learning emotions by supporting students' control appraisals. If all students are expected to succeed, then students are more confident that they will actually succeed in such an environment. *Autonomy and collaboration* involves supporting students' independent thinking and decision-making while also providing opportunities for collaboration with peers. This practice fosters positive learning emotions by supporting students' control appraisals, because students have more opportunities to check their understanding with others. *Instructional quality*, characterised by clear, structured, cognitively activating, and appropriately challenging instruction, fosters positive learning emotions by supporting both students' control and value appraisals. Quality instruction supports student mastery of concepts and procedures by presenting material in a clear and structured manner, which in turn helps students organise the information they are learning. In addition, quality instruction relies on appropriately challenging learning activities that allow students to demonstrate their mastery of concepts and procedures. Together, these practices that support understanding and mastery experiences foster students' control appraisals. Furthermore, when the challenge of a learning activity matches the students' ability, students tend to see more value in the learning experience. Finally, *value induction* captures the use of authentic learning activities and discourse that foster positive learning emotions by emphasising the value of engaging in the learning activity. When used effectively, instructional practices with these characteristics foster positive and diminish negative learning emotions in the classroom by supporting students' control and value appraisals (Pekrun, 2006).

Preliminary research findings support the theorised relationship between instructional practices and student learning emotions in math (Becker, Goetz, Morger, & Ranellucci, 2014; Frenzel, Pekrun, & Goetz, 2007b; Schukajlo & Rakoczy, 2016). Operationalising instruction as 'high' or 'low' quality, Becker et al. (2014) found that high-quality math instructional behaviour was associated with more student enjoyment and less student anger, but this high-quality instruction was not associated with student anxiety. Similarly, Frenzel et al. (2007b) found that high-quality math instruction was associated with greater enjoyment and less anger, anxiety, and boredom. However, these studies did not attend to specific approaches to instruction, such as direct and dialogic instruction. Schukajlo and Rakoczy (2016) conducted a more targeted study, but they only compared two instructional conditions: multiple possible problem solutions versus a single possible problem solution. They found that telling students there was more than one solution to a math problem led students to feel greater enjoyment and less boredom (Schukajlo & Rakoczy, 2016). Although these studies provide supporting evidence for the theoretical relationship between instruction and learning emotions, they offer little contribution to current debates around mathematics instruction. We seek to fill this gap by focusing on one well-known approach: dialogic instruction.

Dialogic mathematics instruction and student learning emotions

Recent modifications of mathematics curriculum standards in the United States (e.g. NCTM, 2014; NGAC & CCSSO, 2010) call for a greater use of dialogic instructional practices that rely on students' active involvement in learning through inquiry, discussion, and collaboration (Kuhn, 2007). Dialogic instructional practices are rooted in constructivist approaches to learning (Munter et al., 2015). Constructivist approaches view learning as a process of active involvement whereby students master new knowledge by building on prior knowledge (Meece & Jones, 1996). For example, a student learning how to multiply fractions must apply their existing knowledge of multiplication and the relationships between parts and wholes in a new way. Students build their understanding of new concepts and procedures by working with others to test the viability of new ideas given what they already know to be true (Meece & Jones, 1996). Dialogic instructional practices provide students ample opportunities to connect new knowledge to old knowledge and question the viability of ideas through inquiry-based, discussion-oriented, and collaborative learning activities.

Dialogic instructional practices seek to give students greater intellectual authority in the classroom. In dialogic instruction, teachers expect students to generate their own ideas, as well as explore and share those ideas through discussion and cooperative learning activities (Fulmer & Turner, 2014). The dialogic approach to instruction may be particularly beneficial for adolescents who thrive in environments that provide challenging activities, support autonomy, and include opportunities for interaction with peers (Eccles et al., 1993; Wang & Degol, 2013).

Although little research has examined how dialogic instructional practices relate to learning emotions in math, characteristics of dialogic instruction may align well with the optimal teaching practices proposed by Pekrun's control-value theory. Ultimately, dialogic instructional practices may evoke positive learning emotions and lessen the likelihood of negative learning emotions among adolescents. The involvement of students in collaborative activities that offer autonomy, dialogue, and cooperation suggests that dialogic instructional practices might foster both stronger control and value appraisals (Pekrun, 2000, 2006). When students are able to share their thoughts and check their understanding, they likely feel more confident in their ability to succeed. Furthermore, when students have opportunities to engage in the learning process with peers and tackle meaningful problems, students often see more value in the learning activity. Stronger control and value appraisals lead to greater positive and lesser negative learning emotions (Pekrun 2000, 2006).

Indeed, researchers have found empirical evidence to suggest that dialogic instructional practices may foster positive learning emotions. Teachers in Fulmer and Turner's (2014) qualitative study stated that their students enjoyed challenging instruction that emphasises conceptual and analytical thinking more than procedural instruction. Teachers in this study also felt that challenging instructional practices provided students with more opportunities to feel proud of themselves and their abilities (Fulmer & Turner, 2014). Likewise, Marks (2000) found that students were more emotionally engaged during authentic instruction where students had the opportunity to take an active role in constructing knowledge with their peers and teacher. Similar studies have found that dialogic instructional practices are particularly effective in boosting student interest and engagement in learning activities (Hanze & Berger, 2007; Sawyer, 2004).

Existing research also suggests dialogic instructional practices may be associated with diminished negative emotions. Studies of learning emotions in secondary math classrooms have found that students report less anxiety, anger, and boredom when math instruction is focused on fostering in-depth understanding (Daschmann, Goetz, & Stupnisky, 2011; Frenzel et al., 2007b). Daschmann et al. (2011) also found students reported less boredom when math instruction incorporated activities rooted in practical applications.

In sum, existing theoretical and empirical work on instruction and learning emotions suggests dialogic instructional practices may foster positive learning emotions and diminish negative learning emotions in math. Accordingly, we hypothesise that higher ratings of dialogic instruction will be associated with greater enjoyment and pride and lesser anger and boredom when learning math. We specifically build on prior research by controlling for the direct instructional practices teachers use in their classrooms and exploring these relationships among a socioeconomically and racially diverse population of students.

Gender, mathematics achievement, and dialogic mathematics instruction

Prior literature has suggested that gender and achievement are important determinants of learning emotions (Frenzel, Pekrun, & Goetz, 2007a; Pekrun, 2006). However, it is unclear how the relationship between dialogic instruction and learning emotions may vary by students' gender or prior achievement. In the next two sections, we discuss how students' gender and prior achievement may moderate the associations between dialogic instruction and learning emotions.

Gender differences

Empirical evidence suggests that male and female students respond differently to instructional styles. Compared to female students, male students show preferences for instruction that relies on independent work and memorisation of mathematical procedures, which best aligns with more traditional direct instruction (Geist & King, 2008). Female students, however, prefer instruction that is personally relevant, process-oriented, and embedded in context and relationships (Fredricks et al., 2017; Geist & King, 2008; Meece & Jones, 1996). Extant empirical evidence suggests that female students perform better, are more motivated and have more positive affective experiences in collaborative and constructivist instructional environments compared to their male counterparts (Fredricks et al., 2017; Hossain & Tarmizi, 2012; Timmermans, Van Lieshout, & Verhoeven, 2007; Wang, 2012).

Instructional styles beneficial for female students appear to align with dialogic instructional approaches that may support positive learning experiences. Collaborative and constructivist approaches may strengthen control appraisals because teachers connect new knowledge to what students already know, engage students in finding multiple ways to solve a problem, and encourage students to work on projects with others (Pekrun, 2006). Furthermore, teachers' collaborative and constructivist approaches may strengthen value appraisals by engaging students in discussing the practical value of the lesson and providing opportunities for students to work with their peers (Pekrun, 2006). Since female students show greater preferences for collaborating with their peers and connecting knowledge to what they know (Fredricks et al., 2017; Geist & King, 2008; Meece & Jones, 1996), we anticipate that dialogic instruction may particularly support control and value appraisals for female students. As such, we expect female students to show more positive and less negative learning emotions in dialogic math instruction compared to their male counterparts.

Achievement differences

Theoretical and empirical evidence also contends that students at different achievement levels have different affective experiences in mathematics. Pekrun (2000) theorised that students with low achievement have weaker control and value appraisals, which result in more negative learning emotions than students with higher levels of achievement. Supporting these theoretical postulations, Goetz, Preckel, Pekrun, and Hall (2007) found that students with low levels of reasoning ability experienced less enjoyment and more anger when solving math problems than students with higher levels of reasoning ability. Similarly, Ahmed, van der Werf, Kuyper, and Minnaert (2013) found positive correlations between student achievement and enjoyment and pride in math. Therefore, due to poor control and value appraisals, we expect low-achieving students to generally have more negative experiences in math, than their high-achieving peers.

We also speculate that students with varying levels of achievement may have different affective experiences during dialogic instruction. Opportunities to work collaboratively with peers and engage with authentic learning activities through dialogic instruction may heighten both low control and value appraisals among low-achieving students (Hmelo-Silver, Duncan, & Chinn, 2007; Pekrun, 2006; Roseth, Johnson, & Johnson, 2008). In fact, teachers in Fulmer and Turner's (2014) study described their low-achieving students as feeling supported and motivated by scaffolding and encouragement of discussion through dialogic instructional practices. Along these lines, we speculate that low-achieving students may experience greater positive emotions and fewer negative emotions when their teachers use more dialogic instructional practices.

The current study

Theoretical and empirical evidence suggests that instructional practices are associated with learning emotions in mathematics (Frenzel et al., 2007b; Fulmer & Turner, 2014; Pekrun, 2006; Schukajlo & Rakoczy, 2016). However, few studies specifically examine whether and how dialogic instructional practices relate to learning emotions. Filling this gap in the literature is important as students' learning emotions in mathematics classes play an important role in their motivation and achievement in mathematics (Bailey et al., 2014). Dialogic instruction holds the promise of being especially beneficial for promoting positive learning emotions among adolescents (Eccles et al., 1993;

Wang & Degol, 2013). In this study, we examine relationships between dialogic instruction and learning emotions in a diverse sample of middle school students using multilevel modelling techniques. Moreover, we examine potential moderating effects of gender and prior mathematics achievement on the relationships between dialogic instruction and learning emotions.

Based on Pekrun's (2000, 2006) control-value theory of learning emotions and the reviewed empirical work, we hypothesise that: (a) greater use of dialogic instruction will be associated with higher levels of enjoyment and pride and lower levels of anger and boredom; (b) male students, compared to female students, will feel fewer positive emotions and greater negative emotions when teachers use more dialogic instruction; and (c) low-achieving students, compared to high-achieving students, will feel greater positive emotions along with fewer negative emotions when teachers use more dialogic instruction.

Method

Sample

The sample for the current study included middle school students and their mathematics teachers from nine public schools in the Mid-Atlantic United States. The sample of students included 1307 6^{th} through 8^{th} graders (28.7% 6^{th} grade, 34.6% 7^{th} grade, and 36.7% 8^{th} grade) and was 51.6% female, 59.0% White, 30.8% African American, and 10.3% other race. Approximately 42.3% of the sample qualified for free/reduced price lunch. Students were nested within 70 mathematics classrooms (average student participation per mathematics class = 19 students, SD = 5 students; range = 10 to 40 students).

The sample of teachers included 22 middle school mathematics teachers; 72.7% were female, and 93.2% were White. Twenty of the teachers (90.9%) taught only mathematics courses and two teachers (9.1%) taught both mathematics and science courses. The teachers had on average 9.3 years of teaching experience and ranged from 1 to 32 years of teaching experience. 64% of teachers had their master's degree and 36% had their bachelor's degree. There were no statistically significant differences

Table 1. Mean of instructional practices by teacher demographics.

	Dialogic Instruction (Mean)	Direct Instruction (Mean)
Gender		
Male ($N = 6$)	4.02	3.51
Female ($N = 16$)	3.73	3.62
Race		
White ($N = 20$)	3.81	3.59
African American ($N = 2$)	3.83	3.83
Level of education		
Bachelor's degree ($N = 8$)	3.83	3.42
Master's degree ($N = 14$)	3.79	3.69
Years of experience		
Less than 10 years of experience ($N = 15$)	3.94	3.85
10 years or more of experience ($N = 7$)	3.53	3.02

Note. There were no statistically significant mean differences in instructional practices by teacher demographics.

in the means for dialogic and direct instruction by teacher demographics (see Table 1).

Procedure

To recruit teachers and students, we first described the study to the mathematics teachers and obtained their consent to participate and recruit students in their classroom. Teachers provided letters describing the study and opt-out permission slips to students' families, of which 98% agreed to participate (the participation rate was similar across grades). We then provided students who agreed to participate in the study with a computer-based survey that was administered by the researchers and was completed during regular instructional time. Research staff was available during the survey administration to answer any questions that students had about the survey's purpose or content. Following survey administration, research staff provided students with a small gift for participating in the study. All scales included in the student survey were validated through cognitive testing procedures with middle school students to ensure that items were comprehended and interpreted as we intended. We sent teachers a link to the teacher survey via email and they completed it on their own time. Over 90% of teachers who were sent the link completed the survey. Student demographic data and course grades were gathered from school records of the prior year and the fall and spring of the year in which the survey data were collected.

Measures

Learning emotions

We measured learning emotions with an adapted version of the subscale targeting student learning emotions during math class from the *Academic Emotions Questionnaire-Mathematics*, a well-validated self-report scale that has been used with middle and high school students (Peixoto, Mata, Monteiro, Sanches, & Pekrun, 2015; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). The response scale for pride, anger, and boredom items ranged from 1 (*strongly disagree*) to 5 (*strongly agree*). The response scale for the enjoyment items ranged from 1 (*not at all like me*) to 5 (*very much like me*). Scores on each item were averaged to obtain a composite score for each emotion. Higher values indicated stronger endorsements, while lower values indicate weaker endorsements of that emotion.

The Enjoyment scale included nine items (e.g. 'I look forward to mathematics class') that measured the extent to which students enjoyed their math class (Cronbach's α = .88). The Pride scale included four items assessing student pride in mathematics (e.g. 'I think I can be proud of my knowledge in mathematics class', Cronbach's α = .83). The Anger scale consisted of three items measuring anger and frustration in mathematics class (e.g. 'I am annoyed during mathematics class', Cronbach's α = .81). Finally, the Boredom scale included three items assessing boredom in mathematics class (e.g. 'I cannot concentrate in mathematics class because I am so bored', Cronbach's α =.85).

Classroom-level instructional practices

We measured classroom-level instructional practices with a well-validated teacher-report instruction scale from Kaufman and Junker (2011). Teachers reported on each class they taught. Responses ranged from 1 (*never*) to 5 (*always*). We measured dialogic instructional practices with three items: 'When students construct their own ways of doing a problem, I have the students themselves share their approaches with the rest of the class using their own ways of expressing themselves.'; 'After students have worked on a particularly challenging assignment, I provide opportunities for them to see how others have approached the assignment.'; and 'I use students' responses to problems as the fodder for class discussions.' (Cronbach's α = .63).

To account for the effects of direct instructional practices that teachers tend to use in combination with dialogic instructional practices (Langer-Osuna & Avalos, 2015), we also measured direct instructional practices. Teachers reported on each class they taught and responses ranged from 1 (*never*) to 5 (*always*). We used three items from Kaufman and Junker's (2011) scale to assess direct instruction: 'When a student is unable to complete a task on his/her own, I give him/her a set of steps to follow.'; 'When students get stuck on a multi-step problem, I walk them through the steps they need to perform.'; and 'When students are uncertain about how to get started on an open-ended project, I tell them how to do the first step.' (Cronbach's α = .88).

Scores were tallied and averaged to obtain a composite score for the direct and dialogic instruction scales for each class taught by each teacher. Higher values indicated greater frequency of those practices, while lower values indicated weaker endorsements of those practices.

Covariates

We included several covariates closely related to student mathematics learning outcomes (Frenzel et al., 2007a, Frenzel et al., 2007b; Pekrun, 2000). Students self-reported their gender, grade level, and race. We divided the students into three racial categories: White, African American, and all other racial groups. Student qualification for free/reduced lunch was collected through school record data and was used as a proxy for socioeconomic status. White, male, and students not eligible for free/reduced lunch were the reference groups for race, gender, and free lunch status, respectively, in all data analyses. Students' prior grades in mathematics were gathered through school record data. Students' mathematics grade from the prior year was represented on a scale from 0% to 100%. Student gender, race, free/reduced lunch status, and prior mathematics achievement were included as student-level covariates. The grade level of the class and the mean prior mathematics achievement for each class were included as classroom-level covariates.

Data analysis

To investigate associations between dialogic instruction and learning emotions, as well as gender and prior mathematics achievement as moderators of these relationships, we used multilevel modelling techniques. We used this approach for two reasons. First, our data is composed of students nested within classrooms, thus a multilevel approach accounts for the clustered structure of our data. Second, we measured instructional practices at the classroom level and key moderators of interest—student gender and prior achievement—at the student-level. Because multilevel modelling allows for simultaneous estimation of student- and classroom-level effects and cross-level moderation, we were able to examine the direct effects of the classroom-level variables as well as the moderating effects of student-level variables on classroom-level relationships.

We conducted our analyses in a series of three steps to justify the use of multilevel methods. First, we estimated the fully unconditional models, which yielded the proportion of within- and between-class variance and intraclass correlation of each emotion outcome (i.e. enjoyment, pride, anger, and boredom). This step allowed us to determine that multilevel modelling was appropriate, as the proportion of variance in each emotion attributed to the classroom—the between-class variance—was significant. Second, we added the student-level (i.e. free/reduced lunch, race, gender, and prior mathematics achievement) and classroom-level (i.e. grade level and classroom mean mathematics achievement) background characteristics. This step ensured that the between-class variance was not merely explained by these covariates. In the third and final step, we added our classroom-level instructional predictors and cross-level moderation effects of student gender and prior mathematics achievement. Accordingly, the full level 1 and level 2 equations were as follows:

Level 1:

$$\begin{aligned} \textit{Emotion (of student i in class j)} = B_{0j} + B_{1j}\left(\textit{Black}_{ij}\right) + \\ B_{2j}\left(\textit{OtherRace}_{ij}\right) + B_{3j}\left(\textit{Free Lunch}_{ij}\right) + \\ B_{4j}\left(\textit{Female}_{ij}\right) + B_{5j}\left(\textit{Prior Math Achievement}_{ij}\right) + r_{ij} \end{aligned} \tag{1}$$

Level 2:

$$\begin{aligned} B_{0j} = \gamma_{00} + \gamma_{01}\left(\textit{Grade Level}_j\right) + \gamma_{02}\left(\textit{Classroom Mean Math Achievement}_j\right) + \\ \gamma_{03}\left(\textit{Direct Instruction}_j\right) + \gamma_{04}\left(\textit{Dia}\log\textit{ic Instruction}_j\right) + u_{0j} \end{aligned} \tag{2}$$

$$B_{1j} = \gamma_{10} \tag{3}$$

$$B_{2j} = \gamma_{20} \tag{4}$$

$$B_{3j} = \gamma_{30} \tag{5}$$

$$B_{4j} = \gamma_{40} + \gamma_{41}\left(\textit{Direct Instruction}_j\right) + \gamma_{42}\left(\textit{Dia}\log\textit{ic Instruction}_j\right) + u_{4j} \tag{6}$$

$$B_{5j} = \gamma_{50} + \gamma_{51}\left(\textit{Direct Instruction}_j\right) + \gamma_{52}\left(\textit{Dia}\log\textit{ic Instruction}_j\right) + u_{5j} \tag{7}$$

We conducted all analyses using M*plus* version 7 (Muthén & Muthén, 2012). We used the 'TYPE = TWOLEVEL RANDOM' option for the multilevel modelling and the 'ALGORITHM = INTERGRATION' option to estimate cross-level moderation effects. Due to these procedures, all coefficients presented are unstandardised coefficients. The percentage of missing data ranged from 2% to 36% on any given variable. The largest percentage of data was missing on prior mathematics achievement (26%) and the pride (33%), anger (36%), and boredom (36%) composites. Prior mathematics achievement was collected through school records from the prior year, meaning students who attended a different school in the prior year would not have this data in their record. The pride, anger, and boredom scales were positioned towards the end of the survey, meaning it is likely students were missing data on these measures because they did not finish the survey. Students in the 'other race' category (N = 134) were more likely to have missing data than African American or White students (χ^2 (1) = 6.78, $p < .01$). However, there were no other patterns in the missing data. Little's Missing Completely at Random (MCAR) test was statistically non-significant (χ^2 (17) = 12.26, p = .78) suggesting that data were MCAR (Little, 1988; Schafer & Graham, 2002). Therefore, we used full information maximum likelihood estimation (FIML) to handle missing data (Baraldi & Enders, 2010). FIML uses all available data to identify parameter values that are most likely to produce the sample data and draws on both coefficients and variance values when calculating deviance scores, thus allowing for model comparison (Aguinis, Gottfredson, & Culpepper, 2013; Peugh, 2010). We chose FIML because it can be done directly in M*plus*, it yields unbiased parameter estimates when missing data are MCAR (Baraldi & Enders, 2010), and unlike other techniques (e.g. multiple imputation), it is more efficient, easier to replicate, and does not introduce randomness into the model (Dong & Peng, 2013). In each multilevel model, we grand mean centred prior mathematics achievement, instructional practices, and the learning emotion outcomes.

Results

Preliminary analyses

Table 2 provides descriptive statistics, student-level and intraclass correlations for the learning emotions, and classroom-level correlations between the two types of instructional practices. Skewness and kurtosis were assessed for each variable and suggested that all key variables were normally distributed (i.e. values were $<|1|$). In addition, our variables were correlated in the directions we expected. Dialogic and direct instruction were moderately negatively correlated ($r = -.460$), suggesting that teachers who use more dialogic instruction tend to use less direct instruction. However, because these constructs were not highly correlated, teachers are assumed to use a mix of these instructional practices. Therefore, we controlled for direct instruction to account for the mix of instructional practices teachers use. This allowed us to more accurately assess the usefulness of dialogic instruction above and beyond direct instruction.

Multilevel modelling

Steps 1—3: intraclass correlations and explained variances

In Step 1, we ran the fully unconditional models to obtain baseline estimates of between-class variance and intraclass correlation for each learning emotion. The between-class variances significantly differed from zero for all four learning emotions ($ps < .05$), suggesting that classrooms systematically differed in the levels of learning emotions. The intraclass correlations for enjoyment, pride, anger, and boredom were .09, .05, .21, and .13, respectively. These intraclass correlations suggest that significant amounts of students' emotional experience were due to between-classroom differences.

Table 2. Student- and classroom-level correlations and ICCs.

Student-level predictors and outcomes	1	2	3	4	5	6	7	8	9	10	Mean	SD
1. Receives free/reduced lunch	–	–	–	–	–	–	–	–	–	–	–	–
2. White	−.552**	–	–	–	–	–	–	–	–	–	–	–
3. African American	.553**	−.799**	–	–	–	–	–	–	–	–	–	–
4. Other race	.056*	−.405**	−.225**	–	–	–	–	–	–	–	–	–
5. Female	−.022	.032	−.011	−.036	–	–	–	–	–	–	–	–
6. Prior mathematics achievement	−.289**	.258**	−.269**	−.001	.126**	–	–	–	–	–	82.94	11.12
7. Enjoyment	−.114**	.062*	−.061*	−.008	−.059*	.261**	–	–	–	–	3.79	0.88
8. Pride	−.030	−.033	.022	.022	−.068*	.284**	.515**	–	–	–	3.99	0.79
9. Anger	.307**	−.219**	.247**	−.021	.059	−.306**	−.644**	−.319**	–	–	2.36	1.07
10. Boredom	.186**	−.126**	.145**	−.017	.029	−.245**	−.621**	−.362**	.760**	–	2.43	1.12
ICC	–	–	–	–	–	–	*0.094*	*0.052*	*0.214*	*0.133*	–	–

Classroom-level predictors	1	2	3	4	Mean	SD	–	–	–	–	–	–
1. Grade level	–	–	–	–	–	–	–	–	–	–	–	–
2. Class mean mathematics achievement	−.205	–	–	–	83.24	7.43	–	–	–	–	–	–
3. Direct instruction	−.207	−.001	–	–	3.52	1.16	–	–	–	–	–	–
4. Dialogic instruction	−.080	.316*	−.460**	–	3.85	0.77	–	–	–	–	–	–

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

In steps 2 and 3, we added student- and classroom-level covariates and the instructional predictors, respectively. In each step, the within- and between-class residual variances decreased, suggesting that the added predictors explained a portion of the variance in the learning emotion outcomes. One exception was the residual variance for pride, which slightly increased between steps 2 and 3. This finding does not necessarily mean that the model with instructional predictors fits worse than the model with covariates. In fact, in multilevel modelling, adding significant predictors does not always lead to a reduction in residual variance (Nezlek, 2011).

In the final models, the between-class residual variance was no longer significantly different from zero for pride (*ps* $> .05$), but not for enjoyment, anger, and boredom (see Table 3). These results suggest that the instructional practices were excellent predictors of classroom-level experiences of pride, but only adequate predictors of classroom-level experiences of enjoyment, anger and boredom. Yet, between-class residual variances for enjoyment, anger, and boredom approached non-significant values as the covariates and instructional predictors were added, suggesting that the instructional practices did explain a portion of the variance.

Effects of instructional practices on learning emotions

The multilevel modelling results showed significant main effects of dialogic instruction on each learning emotion (see Table 3). Teachers who used more dialogic instruction tended to have students who reported feeling greater enjoyment ($B = 0.14$, $p < .05$) and pride ($B = 0.15$, $p < .05$), along with less anger ($B = -0.20$, $p < .05$) and boredom ($B = -0.22$, $p < .05$).

Effects of gender and prior mathematics achievement on the dialogic instruction-emotion relationship

Our findings revealed significant cross-level moderation effects of gender on three occasions. Gender significantly moderated the relationship between dialogic instruction and enjoyment ($B = -0.12$, $p < .05$), anger ($B = -0.11$, $p < .05$), and boredom ($B = 0.21$, $p < .05$). To better interpret these findings, we graphed these effects (see Figures 1–3) and conducted simple slopes analyses in M*plus* (Muthén & Muthén, 2012). These findings suggest that male students experienced more enjoyment ($B = 0.15$, $p < .05$), less anger ($B = -0.22$, $p < .05$), and less boredom ($B = -0.22$, $p < .05$) when teachers used more dialogic instruction. Female students' enjoyment ($B = 0.04$, $p = ns$), anger ($B = -0.12$, $p = ns$), and boredom ($B = -0.02$, $p = ns$), on the other hand, did not vary as a function of instructional practices.

In addition, our findings demonstrated a significant cross-level moderation effect of prior achievement. Prior achievement significantly moderated the relationship between dialogic instruction and student pride in mathematics ($B = -0.01$, $p < .05$). As shown in Figure 4, low-achieving students experienced greater boosts in pride when their teachers used more dialogic instructional practices ($B = 0.24$, $p < .05$). On the other hand, no difference was found in high achieving students' pride as a function of the level of dialogic instruction ($B = 0.05$, $p = ns$).

Table 3. Full multilevel model of student- and classroom-level covariates and instructional practices on learning emotions

	Enjoyment	*Pride*	*Anger*	*Boredom*
Student-level variables	Coefficient (s.e.)	Coefficient (s.e.)	Coefficient (s.e.)	Coefficient (s.e.)
Receives free/reduced lunch (γ_{10})	−0.057 (0.070)	−0.013 (0.062)	0.331 (0.090)**	0.143 (0.103)
African American (γ_{20})	0.091 (0.072)	0.223 (0.072)*	0.095 (0.095)	0.028 (0.105)
Other race (γ_{30})	−0.008 (0.081)	0.177 (0.095)	−0.152 (0.124)	−0.110 (0.138)
Female (γ_{40})	−0.173 (0.033)**	−0.144 (0.053)*	0.212 (0.048)**	0.134 (0.075)
Prior mathematics achievement (γ_{50})	0.025 (0.006)**	0.022 (0.007)*	−0.022 (0.004)**	−0.023 (0.005)**
Classroom-level variables				
Grade level (γ_{01})	0.031 (0.012)*	−0.008 (0.006)	−0.011 (0.013)	0.020 (0.009)*
Class mean mathematics achievement (γ_{02})	−0.011 (0.007)	−0.007 (0.010)	0.005 (0.009)	0.008 (0.010)
Direct instruction (γ_{03})	0.001 (0.037)	0.081 (0.033)*	−0.017 (0.061)	0.054 (0.056)
Dialogic instruction (γ_{04})	0.154 (0.054)*	0.149 (0.053)*	−0.224 (0.081)*	−0.221 (0.089)*
Female slope				
Direct instruction (γ_{41})	−0.023 (0.039)	0.023 (0.049)	0.007 (0.070)	−0.010 (0.050)
Dialogic instruction (γ_{42})	−0.117 (0.056)*	−0.061 (0.054)	0.106 (0.044)*	0.206 (0.094)*
Mathematics achievement slope				
Direct instruction (γ_{51})	0.004 (0.005)	−0.002 (0.005)	−0.002 (0.004)	0.000 (0.004)
Dialogic instruction (γ_{52})	0.004 (0.007)	−0.012 (0.006)*	−0.004 (0.007)	0.000 (0.006)
Residual variances				
Between-classroom residual variance in intercept ($\tau_{\beta 0}$)	0.032 (0.015)*	0.010 (0.008)	0.088 (0.039)*	0.079 (0.030)*
Between-classroom residual variance in Female slope ($\tau_{\beta 01}$)	0.028 (0.017)	0.049 (0.019)*	0.005 (0.007)	0.033 (0.021)
Between-classroom residual variance in mathematics achievement slope ($\tau_{\beta 05}$)	0.000 (0.000)	0.000 (0.000)	0.000 (0.001)	0.000 (0.001)
Within-classroom residual variance in outcome (σ^2)	0.631 (0.029)**	0.522 (0.027)**	0.836 (0.044)**	1.024 (0.053)**

*$p<.05$;
** $p<.001$.
Note. All coefficients are unstandardised.

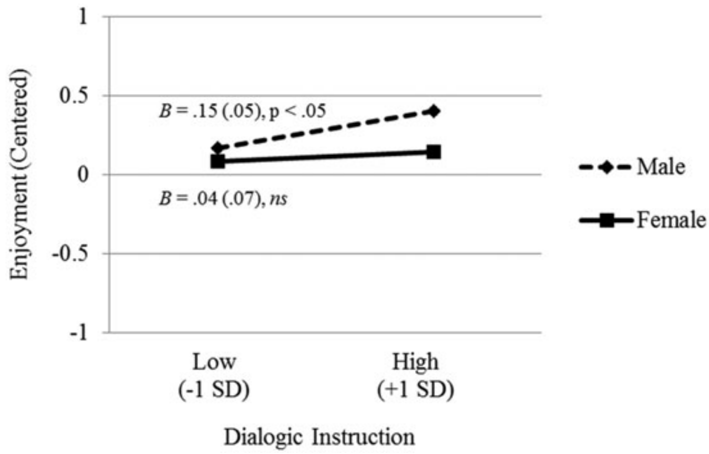

Figure 1. Gender as a moderator of the dialogic instruction–enjoyment slope. Coefficients are unstandardised and enjoyment is centred at the grand mean.

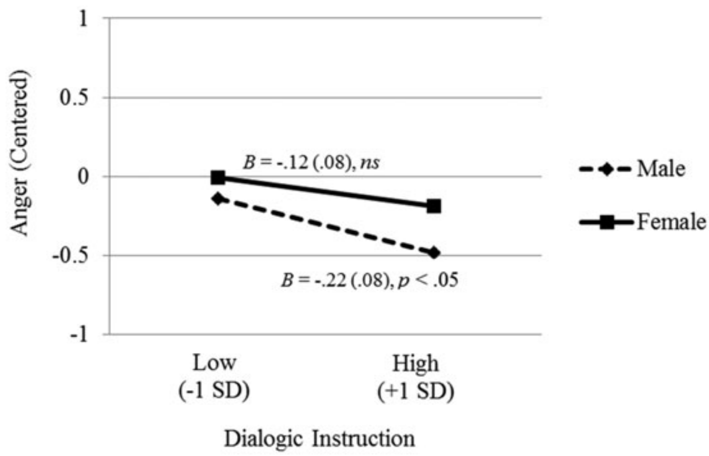

Figure 2. Gender as a moderator of the dialogic instruction-anger slope. Coefficients are unstandardised and anger is centred at the grand mean.

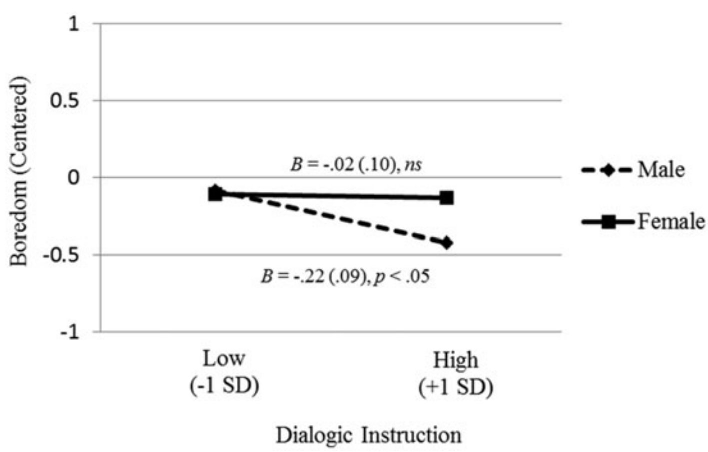

Figure 3. Gender as a moderator of the dialogic instruction-boredom slope. Coefficients are unstandardised and boredom is centred at the grand mean.

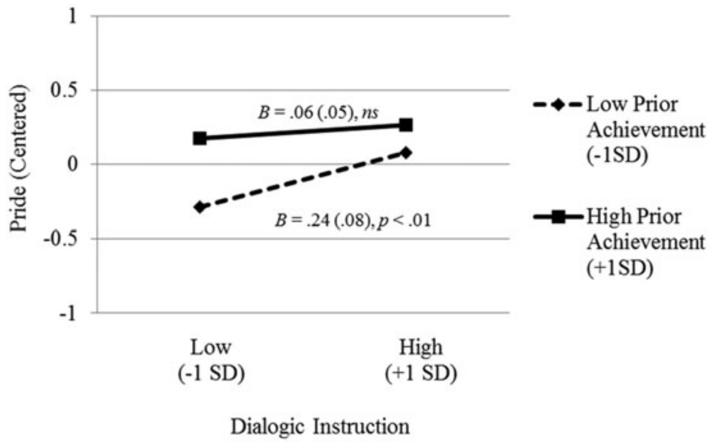

Figure 4. Prior achievement as a moderator of the dialogic instruction-pride slope. Coefficients are unstandardised and pride is centred at the grand mean.

Discussion

Adolescents who experience more positive emotions when learning mathematics are more motivated, use more cognitive strategies when learning, have higher levels of achievement, and are more likely to pursue STEM careers (Bailey et al., 2014; Pekrun, Goetz, Titz, & Perry, 2002; Wang & Degol, 2013). To foster positive learning experiences in mathematics, researchers and practitioners have explored the effectiveness of dialogic instructional practices in boosting achievement and engagement outcomes (Cain, 2002; Hanze & Berger, 2007; Kuhn, 2007; Sawyer, 2004), but they have paid little attention to the implications of dialogic instructional practices for learning emotions. As such, the overarching goal of this study was to examine associations between dialogic instruction and learning emotions in mathematics. Using a diverse middle school sample, our findings partly supported our hypotheses. Higher endorsements of dialogic instruction were associated with more enjoyment and pride, and less anger and boredom on average. However, patterns were not consistent across all groups. Males and low-achieving students showed more benefits when their teachers used more dialogic instruction than their female and high-achieving counterparts.

Dialogic instructional practices and learning emotions

To our knowledge, this is the first empirical study of the relation between dialogic instruction and student learning emotions in math. Our findings provide evidence supporting a general benefit of dialogic instructional practices in fostering positive and reducing negative learning emotions in math. Teachers who reported using more dialogic instruction had students who reported feeling greater enjoyment and pride, and lesser anger and boredom in math class. These findings align with extant studies demonstrating that dialogic instructional practices are fun and enjoyable for students (Fulmer & Turner, 2014; Marks, 2000) as well as those indicating that dialogic instruction fosters student motivation and engagement (Hanze & Berger, 2007; Sawyer, 2004). Given that learning emotions can facilitate or undermine positive learning outcomes, dialogic instruction may be a promising means through which teachers can create more positive affective learning experiences for adolescents in math class.

We speculate that dialogic instruction is associated with favourable learning emotions in math because many of its characteristics align with instructional qualities theorised to be beneficial for student learning emotions in the control–value model (Pekrun, 2006). Dialogic instruction is characterised by an emergent and developmental approach to learning that places emphasis on flexible trajectories for learning and incremental mastery of content material (Munter et al., 2015). This approach aligns with Pekrun's (2006) cooperative goal structures that promote whole class mastery of content material rather than relying on individualised or comparative performance measures. In addition, dialogic instruction positions students as co-participants, even drawing on student thinking and behaviours as a resource when lesson planning. Through co-participation, students can be autonomous in their learning and collaborate with their peers, both of which are known to fulfil psychological needs for autonomy and relatedness, particularly in adolescence (Eccles et al., 1993; Pekrun, 2006; Wang & Degol, 2013). Discussion, the primary mode for learning in dialogic instruction, further cognitively activates and involves students in the learning process by positioning them as having intellectual authority. Finally, productive time spent wrestling with complex ideas in dialogic instruction presents students with appropriate challenges and opportunities to reflect on what they are learning, which are expected to lead to more positive learning emotions (Munter et al., 2015; Pekrun, 2006).

Math educators already promote dialogic instructional practices as they are known to encourage deep understanding of mathematical concepts and procedures (NCTM, 2014). Our findings suggest dialogic instruction may also offer affective benefits for students by creating more positive learning emotions. Ultimately, more positive affective experiences may lead students to be more motivated and engaged in math courses, and more likely to pursue STEM college majors and careers that rely on mathematics (Maltese & Tai, 2011; Wang & Degol, 2013). In fact, some researchers suggest that affective experiences while learning a challenging subject, such as mathematics, may be a stronger predictor of pursuing a career in STEM than other factors such as course enrolment or achievement (Maltese & Tai, 2011).

Dialogic instruction: an asset for males and low-achieving students?

Although the significant direct effects in our study suggest dialogic instructional practices contribute to more favourable learning emotions across all students, moderating effects imply that the benefits of dialogic instruction may be concentrated among male and low-achieving students.

Male students

Contrary to our hypothesis, teachers' use of dialogic instructional practices was associated with greater positive and fewer negative learning emotions for male students than female students. Dialogic instruction is a multidimensional construct that incorporates collaborative inquiry- and discussion-based activities that seek to give students more intellectual authority in the classroom (Fulmer & Turner, 2014; Grady et al., 2012;

Munter et al., 2015). The items in this study asked teachers how often they had students demonstrate their unique approach to a problem, had students share how they solved a particularly challenging assignment, and used student responses to prompt discussion. These items captured certain aspects of dialogic instructional practices, such as encouraging discussion and handing intellectual authority over to students. However, they did not assess other aspects of dialogic instructional practices, such as inquiry-based and collaborative activities. Coincidentally, these inquiry-based and collaborative activities have been associated with more positive academic outcomes among female students (Brotman & Moore, 2008; Fredricks et al., 2017; Hossain & Tarmizi, 2012); therefore, the omission of these practices in our measure of dialogic instruction might explain why we saw little benefit to learning emotions for female students.

Furthermore, gendered variation in classroom participation levels may have exerted an unexpected influence on our results. Researchers have found that female students participate less in mathematics classes (Samuelsson & Samuelsson, 2016) and shy away from getting involved in STEM contexts because they feel unwelcome or less competent despite requisite ability and performance (Walton, Logel, Peach, Spencer, & Zanna, 2015; Wang & Degol, 2013). Dialogic instruction relies heavily on participation as a means for learning. Therefore, female students may have participated less than male students leading to the unexpected results. In particular, we speculate that the content of our items may have also heightened the likelihood of these unexpected results because the items focus on psychologically risky activities of sharing one's own and others' ways of solving a problem. Placing the spotlight on students in these ways may particularly reduce the likelihood that female students would participate in math class. As such, it is possible that male students would experience more enjoyment/less anger and boredom, and female students would show less positive learning emotions when teachers used more dialogic instruction. The graphical representation of the findings in Figures 1–3 fit well with these explanations.

Low-achieving students

Our findings also show that dialogic instruction may be particularly beneficial for low-achieving students. Compared to their higher achieving peers, low-achieving students in the current study reported feeling significantly greater pride when teachers used more dialogic instructional practices. We postulate that this finding is due to the control appraisal-supportive practices teachers use in dialogic instruction, as pride is largely determined by control appraisals rather than value appraisals (Pekrun, 2006). The strengths of dialogic instructional practices in providing opportunities to think critically and share one's thoughts (Fulmer & Turner, 2014) may help students develop a deeper understanding of the material. Although our findings show these practices boost pride among all students, there seems to be a particularly beneficial effect on pride for low-achieving students. It is plausible that low-achieving students are not used to feeling strong control appraisals in math because they have had few experiences of success to support them in feeling competent in math. Thus, compared to their higher-achieving peers, low-achieving students may feel particularly proud when they put effort into learning and deeply understand the material through dialogic instruction (Ahmed et al., 2013).

Limitations

The present study contributed to the literature on instruction and learning emotions using a diverse sample of students; yet, future research could build on limitations in this work. First, although teacher reports of instruction are commonly used in educational research (Wang & Degol, 2016b), teachers often vary in the accuracy of their self-assessments of instructional practices (Kaufman, Stein, & Junker, 2016). Teacher self-report measures of their instruction across multiple classes may be biased by social desirability or limitations in their ability to detect and report subtle differences in instruction across their classrooms. Future research should consider using classroom observational assessment to assess teaching practices as well as examine how teacher-reports compare to observation and student-report assessments of teacher instructional practices.

In addition, although we rely on theory in our assumption that instruction leads to learning emotions, it is likely that there may be reciprocal relationships between teachers' instructional practices and learning emotions over time. If students are consistently frustrated or bored with particular instructional strategies, a teacher may decide to change practices in order to better meet the needs of their students. As a result, students' learning emotions may change over time as a function of their teachers' instructional practices and vice versa. Future studies should collect multiple waves of data to explore the potential reciprocal relationship between instructional practices and learning emotions.

Furthermore, we focused on enjoyment, pride, anger, and boredom in our study and did not include anxiety in this study because we sought to explore learning emotions that have been less commonly studied in existing literature. However, we acknowledge the importance of studying math anxiety, which has shown to be linked to academic outcomes (e.g. motivation and performance) and is especially pertinent to gender differences in students' math experiences (Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013; Ma & Xu, 2004; Meece, Wigfield, & Eccles, 1990). Future research should also consider exploring how dialogic instruction may relate to anxiety in mathematics and how gender may moderate that relationship.

It is also notable that some learning emotions in our study were correlated at greater than .6, namely anger and boredom as well as enjoyment with pride, anger and boredom. These high correlations suggest that these learning emotions are closely related to each other. While we ran the models separately by learning emotion, meaning multicollinearity does not affect our model results, more research should be done to tease apart these emotions and better understand how they are interrelated.

Finally, dialogic and direct instruction are multifaceted constructs characterised by a wide variety of teaching practices (Hmelo-Silver et al., 2007; Kirschner et al., 2006; Kuhn, 2007). While practical constraints on surveys limit how many questions can be included, it is important to recognise that the operational definitions of direct and dialogic instruction in this study, generated by three items in each scale, only focus on certain conceptual components of these instructional approaches. For example, our dialogic instruction measure represented sharing students' approaches to solving a problem and discussion. Future measures may seek to capture other aspects of dialogic instruction, such as inquiry and collaboration. In addition, we examined differential effects of dialogic instruction on the learning emotions of males and low-achieving students compared to their female and high achieving counterparts. It is plausible that certain aspects of dialogic instruction could drive the differential effects for males and low-achieving students that we observed in this study. Future research should consider relying on more items to capture additional components of dialogic instruction and explore how various practices characteristic to dialogic instruction foster positive learning experiences in general and for particular groups of students.

Conclusion

Despite various efforts to examine the effectiveness of dialogic instruction in promoting motivation and achievement outcomes in math, the implications of this type of instruction for learning emotions have received little attention in educational research. Our findings provide timely insight into associations between instructional practices and learning emotions, which are relevant to recent curriculum reform efforts encouraging teachers' use of more dialogic mathematics instructional practices (NGAC & CCSSO, 2010). Building on prior literature, our findings suggest that beyond boosting engagement and critical thinking skills (Fulmer & Turner, 2014; Marks, 2000), dialogic mathematics instructional practices create more positive learning emotions for middle school students. Furthermore, these practices may particularly benefit males and low-achieving students. Therefore, our findings provide empirical support for curriculum reforms stressing the use of dialogic instructional practices. Yet, future research needs to examine the mechanisms by which dialogic instructional practices lead to more positive learning emotions for all students and for students with certain demographic characteristics. With greater insight into effective dialogic instructional practices and mechanisms by which these practices foster positive learning experiences, we may better understand how instruction can be designed to improve students' learning experiences in math.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by the National Science Foundation Grant 1315943.

ORCID

Jamie Amemiya http://orcid.org/0000-0001-6946-4717

References

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, *39*(6), 1490–1528.

Ahmed, W., van der Werf, G., Kuyper, H., & Minnaert, A. (2013). Emotions, self-regulated learning, and achievement in mathematics: A growth curve analysis. *Journal of Educational Psychology*, *105*(1), 150–161.

Bailey, M., Taasooobshirazi, G., & Carr, M. (2014). A multivariate model of achievement in geometry. *The Journal of Educational Research*, *107*(6), 440–461.

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, *48*(1), 5–37.

Becker, E. S., Goetz, T., Morger, V., & Ranellucci, J. (2014). The importance of teachers' emotions and instructional behavior for their students' emotions – An experience sampling analysis. *Teaching and Teacher Education*, *43*, 15–26.

Brotman, J. S., & Moore, F. M. (2008). Girls and science: A review of four themes in the science education literature. *Journal of Research in Science Teaching*, *45*(9), 971–1002.

Cain, J. S. (2002). An evaluation of the connected mathematics project. *The Journal of Educational Research*, *95*(4), 224–233.

Daschmann, E. C., Goetz, T., & Stupnisky, R. H. (2011). Testing the predictors of boredom at school: Development and validation of the precursors to boredom scales. *British Journal of Educational Psychology*, *81*(3), 421–440.

Dong, Y., & Peng, C. Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*(1), 1–17.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Iver, D. M. (1993). Development during adolescence. *American Psychologist*, *48*(2), 90–101.

Fredricks, J. A., Hofkens, T., Wang, M.-T., Mortenson, E., & Scott, P. (2017). Supporting girls' and boys' engagement in math and science learning: A mixed methods study. *Journal of Research in Science Teaching*, *55*(2), 271–298.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007a). Girls and mathematics – A 'hopeless' issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology in Education*, *22*(4), 497–514.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007b). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction*, *17*(5), 478–493.

Fulmer, S. M., & Turner, J. C. (2014). The perception and implementation of challenging instruction by middle school teachers. *The Elementary School Journal*, *114*(3), 303–326.

Geist, E. A., & King, M. (2008). Different, not better: Gender differences in mathematics learning and achievement. *Journal of Instructional Psychology*, *35*(1), 43–52.

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science*, *24*(10), 2079–2087.

Goetz, T., Preckel, F., Pekrun, R., & Hall, N. C. (2007). Emotional experiences during test taking: Does cognitive ability make a difference? *Learning and Individual Differences*, *17*(1), 3–16.

Grady, M., Watkins, S., & Montalvo, G. (2012). The effect of constructivist mathematics on achievement in rural schools. *The Rural Educator*, *33*(3), 37–46.

Hanze, M., & Berger, R. (2007). Cooperative learning, motivational effects, and student characteristics: An experimental study comparing cooperative learning and direct instruction in 12th grade physics classes. *Learning and Instruction*, *17*(1), 29–41.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107.

Hossain, A., & Tarmizi, R. A. (2012). Gender-related effects of group learning on mathematics achievement among the rural secondary students. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, *47*, 1–22.

Kaufman, J., & Junker, B. (2011, September). What drives alignment between teachers' survey self-reports and classroom observations of standards-based mathematics instruction? Poster presented at the meeting of the Society for Research on Educational Effectiveness, Washington, D.C.

Kaufman, J. H., Stein, M. K., & Junker, B. (2016). Factors associated with alignment between teacher survey reports and classroom observation ratings of mathematics instruction. *The Elementary School Journal*, *116*(3), 339–364.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.

Kuhn, D. (2007). Is direct instruction an answer to the right question? *Educational Psychologist*, *42*(2), 109–113.

Langer-Osuna, J. M., & Avalos, M. A. (2015). 'I'm *trying* to figure this out. Why don't *you* come up here?': Heterogeneous talk and dialogic space in a mathematics discussion. *ZDM Mathematics Education*, *47*(7), 1313–1322. doi:10.1007/s11858-015-0735-y

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202.

Ma, X., & Xu, J. (2004). The causal ordering of mathematics anxiety and mathematics achievement: A longitudinal panel analysis. *Journal of Adolescence*, *27*(2), 165–179.

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education Policy*, *95*(5), 877–907. doi:10.1002/sce.20441

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, *37*(1), 153–184.

Meece, J., & Jones, G. (1996). Girls in mathematics and science: Constructivism as a feminist perspective. *The High School Journal*, *79*(3), 242–248.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology*, *82*(1), 60–70.

Munter, C., Stein, M. K., & Smith, M. S. (2015). Dialogic and direct instruction: Two distinct models of math instruction and the debate(s) surrounding them. *Teachers College Record*, *117*(11), 1–32.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

National Council of Teachers of Mathematics (NCTM). (2014). *Principles to Actions: Executive Summary*. Retrieved from http://www.nctm.org/uploadedFiles/Standards_and_Focal_Points/Principles_to_Action/PtAExecutiveSummary.pdf

National Governors Association Center for Best Practices (NGAC) & Council of Chief State School Officers (CCSSO). (2010). *Common core state standards for mathematics*. Washington, DC: Author.

Nezlek, J. B. (2011). *Multilevel modeling for social and personality psychology*. London: SAGE Publications Ltd.

Peixoto, F., Mata, L., Monteiro, V., Sanches, C., & Pekrun, R. (2015). The achievement emotions questionnaire: Validation for pre-adolescent students. *European Journal of Developmental Psychology*, *12*(4), 472–481.

Pekrun, R. (2000). A social cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development: Developing motivation and motivating development* (pp. 143–163). Oxford, England: Elsevier.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational Psychology*, *36*(1), 36–48.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*(2), 91–106.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, *48*(1), 85–112.

Roseth, C. J., Johnson, D. W., & Johnson, R. T. (2008). Promoting early adolescents' achievement and peer relationships: The effects of cooperative, competitive, and individualistic goal structures. *Psychological Bulletin*, *134*(2), 223–246.

Samuelsson, M., & Samuelsson, J. (2016). Gender differences in boys' and girls' perception of teaching and learning mathematics. *Open Review of Educational Research*, *3*(1), 18–34.

Sawyer, R. K. (2004). Creative teaching: Collaborative discussion as disciplined improvisation. *Educational Researcher*, *33*(2), 12–20.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177.

Schukajlo, S., & Rakoczy, K. (2016). The power of emotions: Can enjoyment and boredom explain the impact of individual preconditions and teaching methods on interest and performance in mathematics? *Learning and Instruction*, *44*, 117–127. doi:10.1016/j.learninstruc.2016.05.001

Timmermans, R. E., Van Lieshout, E. C. D. M., & Verhoeven, L. (2007). Gender-related effects of contemporary math instruction for low performers on problem-solving behavior. *Learning and Instruction*, *17*(1), 42–54.

Tyack, D. B., & Cuban, L. (1995). *Tinkering toward utopia*. Cambridge, MA: Harvard University Press.

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a 'chilly climate' transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology*, *107*(2), 468–485.

Wang, M.-T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental Psychology*, *48*(6), 1643–1657.

Wang, M.-T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*(4), 1–49.

Wang, M.-T., & Degol, J. S. (2016a). Gender gap in STEM: Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review*, *29*(1), 119–140.

Wang, M.-T., & Degol, J. L. (2016b). School climate: A review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review*, *28*(2), 315–352.