Achievement emotions and elementary school children's academic performance : Longitudinal models of developmental ordering

Lichtenfeld, Stephanie, Pekrun, Reinhard Herrmann, Marsh, Herbert Warren, Nett, Ulrike and Reiss, Kristina

Achievement Emotions and Elementary School Children's Academic Performance: Longitudinal Models of Developmental Ordering

Stephanie Lichtenfeld[1], Reinhard Pekrun[2,3,4], Herbert W. Marsh[3,5], Ulrike E. Nett[6], and Kristina Reiss[7]

Journal of Educational Psychology, in press

(accepted February 22, 2022)

[1] School of Education, University of Durham

[2] Department of Psychology, University of Essex

[3] Institute for Positive Psychology and Education, Australian Catholic University

[4] Department of Psychology, University of Munich

[5] Department of Education, University of Oxford

[6] Faculty of Philosophy and Social Sciences, University of Augsburg

[7] School of Social Sciences and Technology, Technical University of Munich

Author Note

Stephanie Lichtenfeld https://orcid.org/0000-0003-3485-9078

Reinhard Pekrun https://orcid.org/0000-0003-4489-3827

Herbert W. Marsh https://orcid.org/0000-0002-1078-9717

Ulrike E. Nett https://orcid.org/**0000-0002-2767-1993**

Kristina Reiss https://orcid.org/**0000-0002-8954-5036**

Correspondence concering this article should be addressed to Stephanie Lichtenfeld, School of Education, University of Durham, Durham, UK; Phone: +44 191 334 8394; Email: stephanie.lichtenfeld@durham.ac.uk.

The study received Institutional Review Board approval from the Bavarian State Ministry for Education, Science, and the Arts, and we treated participants in accordance with APA ethical guidelines. The data, Mplus analysis code, and research materials for this study are available at https://osf.io/ywt8q/ without any restrictions.

Abstract

Achievement emotions have received increasing attention in research on adolescence and young adulthood, but little is known about these emotions in the early years of schooling. Studies addressing the development of different achievement emotions and their linkages with achievement during these years are largely lacking. The present longitudinal study aimed to fill this gap by examining the development of enjoyment, boredom, and anxiety in mathematics across second to fourth grade (N = 670 German students; M_{age} = 8.45 years, 51.0 % female at baseline) as well as relations between these emotions and children's math achievement. Students' emotions during learning and when taking test and exams in math, school grades in math, and math achievement test scores were measured in annual assessments. Latent structural equation modeling showed that enjoyment decreased, whereas boredom and anxiety remained relatively stable across these years. Moreover, the findings from reciprocal effects models (REMs) show that emotions and achievement were reciprocally linked over time, controlling for autoregressive effects, gender, and family socioeconomic status. Enjoyment positively predicted subsequent achievement, and achievement positively predicted subsequent enjoyment. Boredom and anxiety negatively predicted subsequent achievement, and achievement negatively predicted subsequent boredom and anxiety. The results were consistent across waves and achievement indicators, and highlight the need to attend to students' achievement emotions already during the early years of schooling. Directions for future research and implications for educational practice are discussed.

Keywords: achievement emotion; enjoyment; boredom; math anxiety; math achievement; control-value theory

Educational Impact And Implications Statement

This longitudinal study investigated how elementary school students' emotions towards math develop from Year 2 to 4 and how they relate to students' math performance. Anxiety and boredom remained at the same levels during this time, but enjoyment of math decreased. Emotions and math performance were linked over the years; enjoyment predicted improved performance, and anxiety and boredom predicted reduced performance, from one year to the next. High performance, in turn, predicted more enjoyment, and poor performance predicted anxiety and boredom.Thus, the findings suggest that providing students with opportunities for success will improve their emotions towards the subject, which will ultimately enhance their performance.

Emotions are of pivotal importance for children's academic development. A considerable amount of research has shown that children's and adolescents' emotions are related to their academic achievement (see Goetz & Hall, 2013; Pekrun, 2017; Pekrun & Linnenbrink-Garcia, 2014; von der Embse et al., 2018). However, most of the findings have been obtained from cross-sectional studies, which cannot test the variables' temporal and causal ordering. Only a few studies have investigated linkages between students' emotions and their academic achievement over time (see Pekrun et al., 2017; Pinxten et al., 2014; Putwain, Becker, et al., 2018). Furthermore, none of these studies have investigated the development of different achievement emotions and their links with achievement in the early school years, despite the importance of these years for students' emotional and academic development. As such, empirical evidence on longitudinal linkages between emotions other than anxiety and achievement is lacking for the early school years.

To fill this gap in the literature, we investigated the development of children's achievement emotions in mathematics during the early elementary school years (Years 2 – 4) and examined the effects of these emotions on children's math achievement, as well as reciprocal effects of achievement on emotions, across these years. Thus, with this study we seek to further our knowledge as to when and how achievement emotions change and how they are linked with achievement. This knowledge is needed, for instance, to devise interventions aiming to promote adaptive and reduce maladaptive emotions at school, and to decide when to implement those interventions.

In the following sections, we address the concept of achievement emotion, introduce theoretical underpinnings of our study, and summarize prior empirical evidence. Using control-value theory (CVT) of achievement emotions (Pekrun, 2006, 2018) as a conceptual framework, we discuss the development of achievement emotions and related constructs as well as their linkages with achievement. We then address the aims and hypotheses of our study.

Achievement Emotions

Emotions are commonly viewed as sets of interrelated affective, cognitive, motivational, and physiological processes that are triggered and synchronized in response to important events (see, e.g., Scherer & Moors, 2019). In line with this view, we conceptualize emotions as multifaceted reactions. For instance, enjoyment can consist of happy and elated feelings (affective component), positive thoughts about an event (cognitive component), impulses to approach the stimulus/situation (motivational component), and peripheral physiological activation (physiological component). Achievement emotions refer to emotions that are directly linked to achievement activities or outcomes (Pekrun et al., 2002).

Achievement emotions can be experienced in different situational contexts. In educational research, research on these emotions has traditionally focused on tests and examinations (e.g., test anxiety; von der Embse et al., 2018; Zeidner, 1998). In contrast, we consider both situations of assessing achievement outcomes (such as tests and exams) and situations involving learning-related activities (such as classwork and homework). In both types of situations, behavioral performance is typically judged according to competence-based standards of quality (successful learning in learning situations; successful test performance in tests and exams). As such, both types of situations can trigger achievement emotions. Thus, we assessed both learning-related and test-related achievement emotions separately in the present study.

Different learning- and test-related emotions can be distinguished along the dimensions of valence (pleasant vs. unpleasant), activation (activating vs. deactivating), and object focus (activity vs. outcome; Pekrun, 2006; Pekrun et al., 2007; Pekrun et al., 2002; Pekrun & Perry, 2014). In the present study we focused on students' learning- and test-related enjoyment, boredom, and anxiety in mathematics. These three emotions have been reported to be among the most intensely and frequently experienced emotions in achievement contexts, and to show substantial correlational links with students' academic achievement (see, e.g., Barroso et al., 2020; Camacho-Morles et al., 2021; Larson & Richards, 1991; Pekrun et al., 2002; Pekrun et al., 2010). Thus, they are likely to play a more important role in academic contexts and impact students' achievement more than other emotions.

Enjoyment is defined as a positive (i.e., pleasant) activating emotion. Both boredom and anxiety are negative (i.e., unpleasant) emotions, but they differ in respect to their level of activation, with boredom being a deactivating emotion and anxiety an activating emotion. Furthermore, they differ in object focus. In the achievement context, enjoyment and boredom are activity emotions that relate to current achievement activities, which are either pleasurable (enjoyment) or not sufficiently stimulating and unpleasant (boredom). In contrast, anxiety is an outcome emotion that relates to the outcomes of an activity in terms of a possible future failure.

Moreover, constructs of emotions can be categorized based on temporal stability. *Trait emotions* are dispositions to repeatedly experience an emotion over a longer period of time, whereas *state emotions* refer to "the momentary feeling of an individual" (Eid et al., 1999, p. 284). While state emotions can vary considerably over short periods of time, trait emotions are more stable and longer-lasting. Nevertheless, trait emotions can still be subject to change and may fluctuate over longer periods of time, such as the course of one school year (e.g., Pekrun et al., 2017). In the current study, we focused on students' trait-like emotions in mathematics, that is, their propensities to experience enjoyment, boredom, and anxiety in mathematics over a longer period of time. By investigating them as trait-like emotions, our study considered students' emotions in mathematics more generally, rather than emotions only related to single achievement situations in math.

Control-Value Theory of Achievement Emotions

The present study was based on CVT and the reciprocal effects model of emotion and achievement that is part of the theory (Pekrun, 2006, 2018, 2021; Pekrun et al., 2017). We derived hypotheses on the development of students' achievement emotions and their linkages with achievement from this theory.

Development of Achievement Emotions

According to CVT (Pekrun, 2006), emotions are elicited by control and value appraisals. Control appraisals refer to one's perceived competence to successfully act, such as academic self-concept and self-efficacy expectations, as well as one's outcome expectations to achieve a goal successfully. Value appraisals relate to individuals' perceptions of the importance of achievement activities and outcomes. Succinctly stated, the theory proposes that individuals' achievement emotions are aroused when individuals feel in control over, or out of control of achievement activities and outcomes that are subjectively important to them.

Activity emotions such as enjoyment and boredom depend on appraisals of achievement activities (Putwain, Pekrun, et al., 2018). Specifically, when students feel competent to master the learning material and are interested in the material, they can enjoy learning. In contrast, if they are not interested in the material and do not value learning, they may feel bored. Outcome emotions such as anxiety depend on the perceived controllability and value of success and failure outcomes (Shao et al., 2020). Specifically, anxiety is aroused when perceived control over possible failure is low, and the value of achievement is high. For example, if students feel out of control over their performance on an important exam, they will feel fearful before the exam.

These propositions imply that the development of students' achievement emotions should be driven by their control and value beliefs (Pekrun, 2006). Accordingly, hypotheses on the development of these emotions can be derived from knowledge about the development of these beliefs, as detailed in the next section. Specifically, enjoyment should increase over time if control and value increase; boredom should increase if value declines; and anxiety should increase when control decreases or value increases. Below we specify hypotheses on the development of the three emotions in elementary school based on these principles.

Reciprocal Relations between Achievement Emotions and Achievement

The reciprocal effects model that is part of CVT (see Pekrun, 2006, 2018; Pekrun et al., 2014; Pekrun et al., 2017) suggests that positive activating emotions such as enjoyment promote students' attention, motivation to learn, and use of flexible learning strategies, thus positively influencing academic achievement. In contrast, negative deactivating emotions such as boredom are supposed to deplete cognitive resources, reduce attention and motivation, and lead to shallow information processing, thus exerting detrimental effects on achievement (Goetz et al., 2019). The pattern for negative activating emotions is supposed to be more complex. Specifically, anxiety can reduce intrinsic motivation and the use of deep learning strategies; however, it can also prompt motivation to invest effort to avoid failure (such as in productive failure situations, which can lead to an increase in students' persistence; Kapur, 2010; McNamara, 2001; Schwartz & Martin, 2004) and facilitate the use of more rigid strategies such as rehearsal (see also Bohn-Gettler, 2019; Fiedler & Beier, 2014). However, in most students and under most conditions, the overall effects of anxiety on achievement are expected to be negative as well, in line with the existing evidence on correlations with achievement (Barroso et al., 2020; von der Embse et al., 2018).

These propositions imply that emotions influence learning and achievement. However, achievement outcomes also reciprocally influence students' emotions. Specifically, CVT proposes that success and failure shape students' perceptions of control over achievement, such as their competence beliefs and expectancies of success, and may also influence the perceived value of achievement (Forsblom et al. 2021). As outlined above, perceived control and value are proximal antecedents of achievement emotions. By implication, success and failure shape the development of students' emotions. Success is thought to promote students' enjoyment and reduce their boredom due to enhancing a sense of control and the perceived value of learning. Failure is supposed to increase students' anxiety due to reducing perceived competence and control.

Taken together, these propositions imply that students' achievement emotions and their academic achievement should be reciprocally linked over time. These effects should be positive for enjoyment, with positive effects of enjoyment on subsequent achievement, and positive effects of achievement on the development of enjoyment. In contrast, we expected effects to be negative for boredom and anxiety, with negative effects of boredom and anxiety on achievement, and negative effects of achievement on these negative emotions.

Prior Research

Development of Achievement Emotions

As noted, studies investigating the development of students' achievement emotions in the early elementary school years are lacking, with few exceptions. However, there are several studies of elementary school students' development of competence and value beliefs, most of which are grounded in Eccles's expectancy-value theory of motivation (EVT, Eccles et al., 1983;

Eccles & Wigfield, 2020). Similar to CVT studies that often used competence beliefs as a measure of control (e.g., Forsblom et al., 2021), EVT studies typically used measures of perceived competence to operationalize expectancy. EVT and CVT share structural similarities, with EVT using combinations of expectancy and value to explain motivation, and CVT combinations of control and value to explain emotions (for in-depth discussions of these constructs, see Marsh et al., 2019; Pekrun, 2018).

Control- and Value-Related Beliefs

On average, children's competence beliefs and achievement expectancies decline throughout elementary school and into the middle and high school years (for reviews, see Muenks et al., 2018; Wigfield & Eccles, 2020; Wigfield et al., 2015). Young children are quite optimistic regarding their competencies in various domains (Marsh et al., 1998). This optimistic view changes with increasing age to a more realistic or even pessimistic view, as shown by longitudinal evidence (e.g., Archambault et al., 2010; Fredricks & Eccles, 2002; Jacobs et al., 2002; Spinath & Steinmayr, 2008). Similarly, mean levels of students' interest and intrinsic task value decline during elementary school (for reviews see Wigfield & Eccles, 2020; Wigfield et al., 2015). However, perceptions of the value of achievement may stay stable or even further increase (e.g., Darensbourg & Blake, 2013; Li et al., 2021). One plausible explanation for this development is the increasing emphasis on evaluation over the school years. Evaluative pressure may enhance the importance of success and failure but impair students' intrinsic value for learning (see Wigfield & Cambria, 2010; Wigfield et al., 1998).

In sum, these findings reveal that important antecedents (control- and value-related beliefs) of achievement emotions change across the elementary school years. As a consequence, achievement emotions should likewise change over the elementary school years. Specifically, the decline of both competence beliefs and intrinsic value beliefs suggests that students' enjoyment should decrease and their boredom increase. In addition, the decline of competence beliefs, combined with stable or increasing beliefs in the value of achievement, suggests that anxiety should increase.

Achievement Emotions

Only very few studies have investigated the development of students' emotions during the elementary school years. A cross-sequential study by Vierhaus et al. (2016) found that both enjoyment and boredom during school lessons showed similar levels across Years 2 to 5 and across Years 4 to 7. However, the authors assessed domain-general emotions rather than domain-specific ones, even though achievement emotions are organized in domain-specific ways (Goetz et al., 2007). In contrast, the cross-sectional study by Raccanello et al. (2018) assessed children's enjoyment, boredom, and anxiety towards mathematics and their native language. Second graders reported higher enjoyment and lower boredom and anxiety compared to fourth-graders in both domains. As such, the findings of this study suggest that positive emotions decline and negative emotions increase during the elementary school years. Raccanello et al.'s findings are in line with our hypotheses, but limited given the cross-sectional study design. Longitudinal research is needed to investigate the developmental trajectories of different achievement emotions in relation to specific academic domains and settings during the early years of schooling.

Relations between Achievement Emotions and Academic Achievement

Achievement emotions are linked to students' academic achievement, with enjoyment showing positive relations, and anxiety and boredom negative relations with achievement (for recent meta-analyses, see Barroso et al., 2020; Camacho-Morles et al., 2021). However, only a few studies have examined the development of these relations over time. Supporting the CVT propositions summarized earlier, longitudinal research has provided evidence that achievement emotions are dynamically related to academic achievement in adolescence and young adulthood, with positive emotions typically showing positive reciprocal linkages and negative emotions negative reciprocal linkages with achievement (Forsblom et al., 2021; Ma & Xu, 2004; Meece et al., 1990; Pekrun et al., 2014, 2017; Pinxten et al., 2014; Putwain, Becker, et al., 2018). Related evidence for the early school years is largely lacking, except for three studies investigating longitudinal relations between children's anxiety and achievement in mathematics.

Cargnelutti et al. (2017) assessed students' math anxiety and math achievement in second and third grade. Anxiety and achievement showed negative correlations over time. However, cross-lagged effects derived from two-wave path analysis were not significant, likely due to the small size of the longitudinal sample ($N = 80$). Vukovic et al. (2013; $N = 113$) reported that children's math anxiety in second grade was significantly negatively related to their computation and mathematical applications scores in third grade, but not to their geometric reasoning scores. Gunderson et al. (2018) used a larger sample of first and second graders ($N = 634$) and found significant reciprocal relations between math anxiety and math achievement in first and second graders over the course of six months. Prior math anxiety negatively predicted subsequent math achievement, and prior math achievement negatively predicted subsequent math anxiety.

These findings suggest that elementary students' math anxiety can relate negatively to their math achievement over time, but are limited by small sample size in two of the three studies. For emotions other than anxiety, longitudinal evidence for the early school years is lacking. In sum, existing longitudinal research has focused on students' emotions during adolescence and young adulthood but has neglected the early school years. Thus, in the present study we aim to fill this gap.

Developmental Equilibrium

In addition to analyzing if emotions and achievement are longitudinally linked in elementary school, we also sought to investigate if these links are stable over time. Based on previous research (e.g., Arens et al., 2017; Marsh et al., 2016; Marsh et al., 2018; Pekrun et al., 2017), we refer to a state of balance that renders relations between constructs to be invariant across time as developmental equilibrium. Previous studies on reciprocal relations between emotions and achievement indicate that the direction and the size of the effects of emotions on achievement and vice versa remain stable across the secondary school years (Pekrun et al., 2017). The present study seeks to determine if this state of developmental equilibrium is already reached in the early years of schooling or if reciprocal effects between emotions and achievement diminish over time, or actually become more pronounced as students progress through elementary school.

While teachers may have changed during the time of the study, there was no school transition until the end of elementary school (Year 4), and students usually remained in the same classroom. Thus, the broad academic context for students' affective development remained stable across the three waves of data collection. While we expected levels of emotions to change, the stability of context suggests that the relations between emotions and achievement remained stable and equivalent across the one-year intervals between waves.

Support for developmental equilibrium would offer several important advantages. In statistical models that are based on assumptions of developmental equilibrium, parameters for autoregressive and crosslagged effects are constrained to be equal across waves. From a methodological perspective, these models are more parsimonious and provide more robust and precise estimates, which facilitates the interpretation of results (for a similar line of argument, see Marsh et al., 2018). From a substantive perspective, support for developmental equilibrium suggests that findings are generalizable across critical developmental periods, such as the early years of schooling.

Aims and Hypotheses of the Present Research

We designed the current study to investigate the developmental trajectories of achievement emotions and their linkages with academic achievement in a sample of German elementary school students over the course of three school years (Years 2 to 4). We tested the models using data from a longitudinal study of children's development in mathematics (Quality of Elementary School Education in Math [BIGMATH]; see Reiss, & Winkelmann, 2009). Given that emotions are organized in domain-specific ways (Goetz et al., 2007), both emotions and achievement were measured in relation to a specific subject, namely mathematics. We decided to focus particularly on mathematics as knowledge and skills in this domain and the STEM subjects, in general, are crucially important to prepare students for the challenges of the global economy of the 21st century (Rask, 2010; Tseng et al., 2013).

We focused on math-related enjoyment, boredom, and anxiety as these are frequently experienced achievement emotions in mathematics (Barroso et al., 2020; Frenzel et al., 2007; Goetz et al., 2007). As noted earlier, we conceptualized these emotions as habitual, re-occurring, trait-like emotions. In contrast to momentary emotional episodes, habitual emotions can influence learning and achievement over a longer period of time. Specifically, we assessed learning-related and test-related variants of students' math emotions. We included enjoyment, boredom, and anxiety typically experienced when attending class and doing homework for learning-related emotions. Regarding test emotions, given that boredom is typically not experienced during tests, we assessed only test enjoyment and anxiety. As detailed below, the learning-related emotions were assessed from Year 2 to 4. Test emotions were assessed in Years 3 and 4 only, as Year 2 students did not take tests during the first two years of schooling.

We measured students' achievement by their end-of-the-year grades attained over the academic year, representing students' cumulative curriculum-related performance during the year. Thus, these grades are particularly well suited to investigate the impact of emotions on students' long-term development of achievement. In addition, to investigate the generalizability of the findings across different achievement outcomes, we included test scores from a standardized achievement test that reflects generic mathematical competencies (Reiss et al., 2008). Grades represent students' achievement in the classroom and the feedback students receive on their performance in a domain. Grades have been found to be more closely related to self-concept (Marsh et al., 2005) and emotions (Pekrun et al., 2017) in secondary school students. Thus, we likewise expected students' emotions to be more closely related to their school grades than their test scores in our study.

The study spanned the time from Years 2 to 4. The analyses were guided by hypotheses derived from CVT as outlined earlier. Succinctly stated, we tested the following propositions:

Hypothesis 1. Enjoyment decreases over the school years; boredom and anxiety increase over the years.

Hypothesis 2. Enjoyment and achievement are linked by positive reciprocal effects. Specifically, enjoyment has positive effects on subsequent achievement, and achievement has positive effects on subsequent enjoyment.

Hypothesis 3. Boredom and anxiety are linked with achievement by negative reciprocal effects. Boredom and anxiety have negative effects on subsequent achievement, and achievement has negative effects on subsequent boredom and anxiety.

Hypothesis 4. Reciprocal relations between students' emotions and their achievement remain stable, with effects of emotions on achievement and vice versa showing equivalence (i.e., developmental equilibrium) across each of the one-year intervals between waves.

Method

Participants and Procedure

The sample comprised 670 German elementary school students (51.0 % female) from 30 elementary schools (98.7% of all students in these classrooms). In each school one class was tested. The sample represented a wide range of students in terms of ability, socioeconomic background (see Table 1), and rural (70%) versus urban (30%) areas. For 71.2% of the students, both parents were born in Germany, 9.9% had one parent born in Germany, and 13.7% had parents not born in Germany (4.3% did not report on parents' origin). Students' native langue was German for 86.7% of the students, and 12.5% had another native langue (0.7% did not report on native language). At the first assessment (Year 2), the sample included 593 students (50.9% female, mean age = 8.45 years, *SD* = .40). In Years 3 and 4, we included the students who had participated in the previous assessments and additionally included students from the same classes who had not taken part before but volunteered to join the study. In Years 3 and 4, the sample consisted of 534 students (51.7% female) and 519 students (51.2% female), respectively. 84.0% of the total sample participated in all three assessments, and 8.1% and 7.9% completed two or one assessment(s), respectively.

Towards the end of each of the three school years, participants answered an emotion questionnaire to assess students' reports of their emotions during the year and completed the Mathematics Competency Test-Elementary School (MCT-ES; for a more detailed description, see Reiss et al., 2008). All instruments were administered in children's classrooms by trained external test administrators. We obtained end of year grades in mathematics from school records. We received ethical approval and treated participants in accordance with APA ethical guidelines. We obtained written informed consent from parents and oral consent from students before the data collection.

Transparancy and Openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study, and we follow Journal Article Reporting Standards (JARS; Kazak, 2018). All data underlying the findings, Mplus analysis code, and research materials are available at https://osf.io/ywt8q/. Data were analyzed using MPlus 8 (Muthén & Muthén, 1998-2017). This study's design and its analysis were not pre-registered.

Measures

All variables that were analyzed for this research are reported. The project included various assessments of children, teachers, and parents. For the purpose of investigating the developmental trajectories of achievement emotions and their linkages with academic achievement, the current study focused on the following measures:

Emotions

We measured emotions using the *Achievement Emotions Questionnaire–Elementary School* (AEQ-ES, Lichtenfeld et al., 2012). The AEQ-ES consists of 28 items (nine enjoyment, seven boredom, and 12 anxiety items) that assess trait-like math emotions and are organized in five scales. Students are asked to report on their general feelings towards learning and test-taking in math (e.g., "Please indicate how you feel during math class"). Enjoyment is measured with two scales assessing learning-related and test enjoyment, respectively. The learning-related enjoyment scale measures enjoyment when attending class or doing homework (e.g., "I enjoy math class"); the test enjoyment scale measures enjoyment related to taking tests and exams (e.g., "I look forward to math tests"). Similarly, anxiety is measured with two scales assessing learning-related anxiety (during class or doing homework; e.g., "When I do math homework, I worry if I will ever understand it") and test anxiety (e.g., "I get very nervous during math tests"), respectively. For boredom, the AEQ-ES includes only one scale measuring learning-related boredom (during class and doing homework; e.g., "I find doing math boring"), as boredom is not typically experienced during tests and exams. Students respond on a 5-point Likert scale anchored by five graphical displays of faces showing increasing emotional intensity. To ensure that both boys and girls could identify themselves with the faces, there are different versions for male and female students using male and female faces, respectively (see Lichtenfeld et al., 2012).

In Year 2, we only administered the three learning-related emotion scales as German students lack experience with taking tests and receiving grades during the first two years of schooling. Third- and fourth-graders answered all five emotion scales.

Grades

The end-of-the-year grades in mathematics used in the study were summative scores based on multiple exams within each school year; they represent students' cumulative performance throughout the school year relative to the implemented curriculum in math, thus supporting their validity; In fact, end of year grades represent a broader and fuller account of students' achievement (Harlen, 2008) and therefore are more valid relative to the curriculum than standardized achievement tests (for a similar line of argument see also, Pekrun et al., 2017). Grades ranged from 1 (*excellent*) to 6 (*poor*). Grade scores were reversed prior to the analyses to ease interpretation.

Test Scores

The test scores were derived from the standardized achievement test for elementary school students (MCT-ES), which focuses on measuring students' competencies in arithmetic and geometry. The test includes different test forms for different age levels. The obtained scores were scaled within each wave using one-parameter logistic item–response theory (Rasch scaling; Linacre, 2018), yielding good person reliabilities (see Table 1).

Background Variables

Gender (1 = female, 2 = male) and socioeconomic variables were included as covariates in the analysis. Parental education level is more strongly linked with children's abilities as compared to parental wealth indicators (Zhang & Lee, 2011). Thus, we included parental education to assess SES. In addition, we included one of the most widely used indicators of family capital, namely, the number of books in students' homes. This single-item measure has been used in several large-scale international student assessments (e.g., Organization for Economic Co-operation and Development [OECD], 2011); Trends in International Mathematics and Science Study [TIMSS]; Martin et al., 2020; for evidence supporting reliability and validity, see Sieben & Lechner, 2019). Thus, the highest level of parental education (ranging from 1 = none to 4 = Abitur [highest degree in the German school system]) and the number of books at home (ranging from 1 = none or only a few to 5 = more than 200 books) served as indicators for the socioeconomic background of students.

Strategy of Data Analysis

In preliminary analyses, we examined descriptive statistics as well as manifest and latent correlations. Structural equation modeling (MPlus 8; Muthén & Muthén, 1998-2017) was used to conduct confirmatory factor analyses and to estimate latent neighbor change and reciprocal effects models. All models were estimated using the <type=complex> option implemented in Mplus (Muthén & Muthén, 1998-2017) to correct for the clustering of the data as students were nested in classes.

We used latent neighbor change models to analyze the developmental trajectories for each emotion in both learning-related and test settings. Due to the different time frames for the assessment of learning-related and test emotions, and to obtain more fine-grained evidence about differential developmental trajectories, we estimated separate models for these two types of emotions. Specifically, we estimated models for learning-related enjoyment, boredom, and anxiety from Year 2 to 4, and for test enjoyment and test anxiety from Year 3 to 4. Following recommendations by Steyer et al. (1997), we estimated latent difference variables to capture true change between two neighboring measurement occasions (see Figure 1). Latent trait change in neighbor change models reflects the difference between latent trait levels at two neighboring measurement occasions while controlling for measurement error (Steyer et al., 1997; Steyer et al., 2000). To interpret latent variable change as trait change, the latent variables must have the same meaning across time points (strong measurement invariance; Meredith, 1993; Steyer et al., 2000). As such, we constrained the latent variable structure, factor loadings, and intercepts to be invariant across time. Using the resulting models, we were able to investigate if changes in students' emotions in a particular setting occurred from Year 2 to 3, and from Year 3 to 4.

To examine reciprocal linkages between emotions and achievement across years, we estimated latent cross-lagged models. Generally, to test reciprocal models, longitudinal designs are needed that assess each variable at multiple points in time (McArdle, 2009; Orth et al., 2021; Rosel & Plewis, 2008), as was the case in the present study. Even though alternative explanations cannot be ruled out entirely, these designs make it possible to control for prior levels of outcome variables, thus allowing for stronger inferences about the causal ordering of variables. In the cross-lagged models, we analyzed effects of emotions at each assessment on subsequent achievement one year later, and reciprocal effects of achievement at each assessment on subsequent emotion one year later (Figure 2).

Following recommendations by Forsblom et al. (2021) and Pekrun et al. (2017), we estimated separate models for the five different emotion constructs (i.e., learning-related and test enjoyment and anxiety; learning-related boredom) given the multicollinearity between these emotion variables in the dataset (see Table 1). Relations with grades and test scores were estimated in two separate sets of models. The models for the learning-related emotions included three waves of assessments, and the models for test-related emotions included two waves (see above for the procedure, and see Figure 2).

For the learning-related emotion models, in line with findings on how to model emotions using the AEQ-ES (Lichtenfeld et al., 2012), the emotion variables were modeled as hierarchical latent constructs, with two first-order factors representing the items related to attending class and doing homework, respectively, and one second-order learning-related emotion factor. Following recommendations by Marsh et al. (1992), a correlated uniqueness approach was used by including correlations between residuals for identical emotion items across measurement occasions to control for systematic measurement error. Students' gender and their family socioeconomic status (parental education, cultural capital) were controlled in the analysis to ensure that any observed relations were not mere artifacts of other plausible variables. We included directional paths of the covariates to all of the emotion and achievement variables. We expected the effects linking emotions and achievement to be consistent over time but modest in size due to controlling for autoregressive effects and demographic variables.

Measurement Equivalence across Waves

Preliminary analyses established measurement equivalence of the latent emotion constructs over time. Models of configural, metric, scalar, and residual invariance (Marsh et al., 2014; Meredith, 1993) were sequentially evaluated for each of the emotion variables. Configural invariance refers to equivalent patterns of factor loadings across time points. Metric or weak factorial invariance additionally requires factor loadings to be equivalent, scalar or strong factorial invariance requires factor loading and intercepts to be equivalent, and residual or strict factorial invariance requires equivalence of factor loadings, intercepts, and residual variances. Metric invariance is the minimum requirement to estimate correlations and path coefficients over time (Chen, 2007; Steenkamp & Baumgartner, 1998). Following recommendations by Chen (2007), provided adequate sample size, we considered a change of ΔCFI $\leq$ -.010, supplemented by a change of ΔRMSEA $\leq$ .015 or a change of ΔSRMR $\leq$ .030 to indicate metric invariance. We considered a change of ΔCFI $\leq$ -.010, supplemented by a change of ΔRMSEA $\leq$ .015 or a change of ΔSRMR $\leq$ .010 to indicate scalar or residual invariance. As recommended, the χ^2 difference test was not used in the model comparisons, because it is overly sensitive to sample size (Marsh et al., 1988). Nevertheless, we emphasize that these guidelines are merely rules of thumb and not "golden rules" (Marsh et al., 2004).

Developmental Equilibrium

The three learning-related emotion models included three waves, thus making it possible to investigate developmental equilibrium (for further theorizing on developmental equilibrium see, Marsh et al., 2016). To this end, we estimated two versions for each of these models. In the first version, we freely estimated autoregressive coefficients, cross-paths, and residual factor variances. In the second version, we constrained all three parameters to be invariant across time intervals (i.e., we constrained the effects of emotion in Year 2 on achievement in Year 3 to be the same as those from emotion in Year 3 on achievement in Year 4; similarly, we constrained the effects of achievement on emotions to be the same across intervals).

Estimator and Missing Values

We used the MLR estimator, which is robust against non-normality of observed variables (Muthén & Muthén, 1998-2017). Given the longitudinal design of the study, there is missing data for children who only participated in part of the investigation. To make full use of the data (Enders, 2010), we applied full information maximum likelihood (FIML) estimation, which has shown to be an adequate way of dealing with missing data in longitudinal designs even with high rates of missing data (Jelicic et al., 2009).

Goodness-of-fit Indexes to Evaluate Model Fit

To evaluate the fit of the models, both absolute and incremental fit indices were used. Specifically, we employed the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). Traditionally, values of CFI and TLI higher than .90 and close to .95, values of RMSEA lower than .06, and values of SRMR lower than .08 have been interpreted as indicating good fit (Browne & Cudeck, 1993; Hu & Bentler, 1999). While these fit indexes allow for comparisons with previous research, recommended cut-off values often are not met with datasets derived from more complex studies and should thus be used with caution (Heene et al., 2011; Marsh et al., 2004).

Results

Preliminary Analysis

Descriptive statistics and alpha coefficients for the emotion scales and manifest and latent correlations for emotions and achievement are reported in Table 1. In line with previous evidence (e.g., Pekrun et al., 2011, 2017), enjoyment was negatively related to boredom and anxiety, whereas boredom and anxiety were positively correlated. Enjoyment was positively linked with mathematics achievement each year, whereas boredom and anxiety correlated negatively with achievement. Correlations with covariates predominantly yielded relations between SES variables and anxiety (see Table 1). Specifically, both higher levels of parental education and higher levels of cultural capital were related to lower levels of anxiety. Moreover, parental education and cultural capital were positively related to students' achievement in mathematics.

Confirmatory Factor Analyses (CFAs)

For each wave, we conducted an item-based CFA to confirm that each of the learning-related and test emotions form distinct latent factors. In line with the original factor structure of the AEQ-ES (Lichtenfeld et al., 2012), the CFA models including learning-related emotions with one higher order learning-related emotion factor and two lower order factors representing emotions in the two contexts of learning (i.e., class-related and homework-related emotions). Each of the two test emotions were presented by one factor only. All three models showed a good fit to the data, supporting the hierarchical factor structure of the data as well as the measurement quality of the emotion variables, with CFIs > .93, TLIs > .91, RMSEAs < .05, and SRMRs < .05 (see Supporting Information, Table S2).

The pattern of latent correlations derived from these CFAs was similar to the pattern of manifest correlations (Table 1). Latent correlations were corrected for measurement error and thus represent the highest estimates possible. This indicates that different achievement emotions are closely related, but nevertheless represent distinct constructs (for similar findings with secondary school and university students, see Pekrun et al., 2017; Pekrun et al., 2011). The present findings are also in line with previous findings showing that students' trait-like emotions are more strongly related to each other than are state emotions (see, e.g., Bieg et al., 2013 for similar findings on anxiety and boredom).

Measurement Invariance

Measurement invariance across years was assessed separately for each of the five emotion constructs (see Supporting Information, Table S3). All configural invariance models yielded a good fit, with CFIs > .96, TLIs > .94, RMSEAs < .04, and SRMRs < .05. Reduction in fit for the metric invariance models compared to the configural invariance models was ΔCFIs ≤ -.002, ΔRMSEAs ≤ .006, and ΔSRMRs ≤ .006 for all models, supporting metric invariance for all emotion scales. Reduction in fit for the scalar invariance models was ΔCFIs ≤ -.005, ΔRMSEAs ≤ .005, and ΔSRMRs ≤ .002 for all emotion scales, also indicating clear support for the scalar invariance. Reduction in fit for the residual invariance model for learning-related enjoyment, test enjoyment, and test anxiety was ΔCFI ≤ -.010, ΔRMSEA ≤ .008, and ΔSRMRs ≤ .009, indicating support for invariance. Reduction in fit for the residual invariance models for learning-related boredom and learning-related anxiety was ΔCFIs ≥ -.012, ΔRMSEAs ≥ .005, and ΔSRMRs ≥ .008, indicating non-invariance. In sum, the results supported strong longitudinal invariance of the scales, suggesting that the latent emotion variables can be used in longitudinal analysis. In subsequent models, we used constraints for scalar invariance to keep constraints consistent across the models for all emotions.

Latent Neighbor Change Models

The fit indexes provided support for the latent neighbor change models for all five emotion constructs (CFIs ≥ .95, TLIs > .94, RSMEAs < .04, and SRMRs ≤ .05; see Table 2). In support of Hypothesis 1, students' enjoyment declined both in learning and test settings (see Table 3). Specifically, both learning-related and test enjoyment declined from Year 3 to 4. Boredom remained low and stable over the three years. Counter to our expectations, anxiety related to learning decreased from Year 2 to 3 and remained stable thereafter. Similarly, test anxiety remained relatively stable from Year 3 to 4.

Reciprocal Effects Models of Emotions and Achievement

For both learning-related and test emotions, the cross-lagged structural equation models freely estimating autoregressive effects, cross-lagged effects, and factor residual invariances showed a good fit to the data (see Table 2; for results on test scores, see Supporting Information, Table S4), with CFIs > .94, TLIs > .93, RSMEAs < .05, and SRMRs < .05. When constraining autoregressive effects, cross-lagged effects, and factor residual variances to be equal across the one-year intervals for the learning-related emotions, the reductions in fit were small (ΔCFI ≤ .007s, ΔRMSEA ≤ .002, and ΔSRMR ≤ .009; see Table 2; for results on test scores, see Supporting Information, Table S4). In line with our expectations (Hypothesis 4), these results showed that both autoregressive paths and linkages between emotions and achievement remained stable across the elementary school years, providing evidence for developmental equilibrium. Thus, we used the more parsimonious, constrained models, which provide more robust and precise parameter estimates, for further interpretation (note that unstandardized coefficients are fixed to be equal and standardized coefficients, which are reported herein, can still differ due to the standardization procedure).

Emotions and Grades

Table 4 reports the standardized factor loadings, standardized path coefficients, and residual variances for each of the reciprocal effects models for emotions and grades. In all models, both the emotion variables and students' grades showed considerable autoregressive stability over time (*ß*s $\geq$ .47 and *ß*s $\geq$ .52, respectively).

Learning-related enjoyment and achievement were significantly linked to each other at the first measurement occasion (Year 2), $r = .26$, $p < .001$. Nevertheless, and above and beyond autoregressive effects and effects of gender and the SES variables, there were significant incremental reciprocal effects linking enjoyment and achievement (Table 4). In support of Hypothesis 2, enjoyment significantly predicted subsequent grades, and grades significantly predicted subsequent enjoyment, in each of the two time intervals.

The models for learning-related boredom and anxiety showed that both emotions were significantly linked with achievement in Year 2, latent *r*s = -.28 and -.38, respectively, *p*s < .001. Over and above these initial links, and despite considerable autoregressive stability of the emotion and achievement variables, there were significant reciprocal effects between the two negative emotions and achievement over time while controlling for gender and SES (Table 4). In line with Hypothesis 3, boredom and anxiety were significant negative predictors of subsequent grades, and grades were significant negative predictors of subsequent anxiety and boredom.

Concerning test enjoyment, initial relations with grades at the first measurement occasion (Year 3) were significant as well, latent $r = .35$, $p < .001$. Year 3 test enjoyment had a positive predictive effect on subsequent Year 4 grades; this effect was not significant. In contrast, Year 3 grades had a substantial and significant effect on subsequent Year 4 test enjoyment. Test anxiety and grades were significantly and negatively related in Year 3, latent $r = -.38$, $p < .001$. There was a non-significant negative effect of Year 3 test anxiety on Year 4 grades. Year 3 grades yielded a significant negative effect on subsequent test anxiety, indicating that the lower students' performance, the more anxiety they experienced in the following year.

Emotions and Test Scores

The results for reciprocal relations between emotions and test scores yielded an acceptable fit to the data. They replicated the findings for grades, which supports the generalizability of the results across different achievement measures (Supporting Information, Tables S4 and S5). There were substantial initial positive links between learning-related enjoyment and test enjoyment with test scores, latent *r*s = .15 and .23, respectively, *p*s < .001. Likewise, the models for learning-related boredom and learning-related anxiety and test anxiety showed significant links with achievement in Year 2, latent *r*s = -.34, -.12 and -.12, respectively, *p*s < .001. Despite these links, and over and above autoregressive effects and effects of covariates, learning-related enjoyment and test enjoyment showed positive reciprocal relations with test scores over time. Learning-related boredom, learning-related anxiety, and test anxiety showed negative reciprocal relations with test scores. The direction of effects was entirely consistent across time intervals and with the findings for emotions and grades. However, as expected, the effects were weaker than for grades and did not always reach statistical significance (see Table S5).

Effects of the Covariates

Gender had significant effects on the emotions, indicating that boys reported more math enjoyment (*ß* range .10 to .13; all *p*s < .01) and less math anxiety (*ß* range -.13 to -.15; all *p*s < .01) in both learning and test settings. However, boredom did not differ across genders (*p* > .05). Furthermore, gender had significant effects on math grades, with boys showing better grades than girls (*ß* range .07 to .10, *p*s range .123 to .007). However, gender differences were more pronounced for test scores than for grades (*ß* range .15 to .17, *p*s < .001), confirming that boys performed better in math than girls.

Parents' education level had only weak predictive effects on students' emotions (Table 4). However, in line with expectations, level of education positively predicted students' math grades (*ß* range .23 to .29; all *p*s < .001), indicating that children of parents with higher levels of education performed better in math. In line with these findings, parents' level of education also had predictive effects on test scores (Supporting Information, Table S5; *ß* range .24 to .26; all *p*s < .001). Cultural capital also had positive, albeit weaker, effects on grades (*ß* range .11 to .13; all *p*s < .05) and test scores (ß = .08 to .11). Cultural capital did not have any significant effects on students' emotions.

Discussion

In the present research we sought to investigate developmental changes in students' emotions within different achievement settings at school as well as reciprocal relations between students' emotions and their achievement across the early elementary school years. Specifically, we examined developmental trajectories in elementary school students' enjoyment, boredom, and anxiety and their relations with achievement across Years 2 to 4.

Developmental Trajectories

There was a significant decrease in student's enjoyment both in learning-related and testing situations, showing that students experience considerably less enjoyment across the elementary school years. Interestingly, students' learning-related anxiety decreased from Year 2 to 3, and remained stable thereafter. A potential explanation may be that students in Year 2 are taught various new constructs in mathematics and the content in mathematics is seen as rather challenging for students, while there is a considerable amount of repetition in Year 3. Thus, there may be fewer students who feel overly challenged in Year 3, which may reduce average anxiety levels in class and homework situations. Moreover, no changes in students' boredom or test-related anxiety were detected.

Previous cross-sectional studies point to a maladaptive pattern of emotions emerging with more negative and less positive emotions experienced as students grow older (Raccanello et al., 2018). However, the current longitudinal study replicates this pattern for positive emotions only. A possible explanation for the stability and low levels of boredom may be that teachers emphasize hands-on activities in elementary school more (Zahorik, 1996), which may be less boring than math instruction in the secondary school years. However, the present findings for boredom and anxiety seem at odds with the changes in mean levels of students' control- and value-related beliefs found in previous research. Given that studies based on Eccles's EVT found that both competence beliefs and intrinsic task value beliefs decline during the elementary school years, one would expect that boredom would change as well. Specifically, boredom is likely to increase given that interest and intrinsic value decline during the elementary school years (Fredricks & Eccles, 2002; Jacobs et al., 2002).

As such, the present findings on the development of emotions only partially mirror prior findings for motivation and for competence- and value-related beliefs. Whereas the decline of enjoyment in the present study is congruent with the decline of intrinsic value in previous studies, the lack of change in boredom is not. Based on EVT and CVT, future research should simultaneously include motivation, beliefs, and emotions in the same study design to analyze their joint developmental trajectories in-depth.

Reciprocal Relations with Academic Achievement

The findings support our hypotheses on relations between children's math emotions and their achievement. Enjoyment and achievement were positively linked over time; enjoyment positively predicted subsequent achievement, and achievement, in turn, positively predicted enjoyment. For boredom and anxiety, we found the opposite pattern of effects. Both boredom and anxiety resulted in lower levels of achievement the following year, and lower levels of achievement, in turn, led to higher subsequent boredom and anxiety. While these predictive effects may be considered small in size (but see Gignac & Szodorai, 2016), it is important to bear in mind that they are incremental effects explaining changes in achievement and emotions over and above prior links between emotions and achievement, autoregressive effects, and effects of the background variables. Given that emotions and achievement were already substantially related at the start of the study and both emotions and achievement remained relatively stable over time, only a limited amount of variance was left to be explained. Nevertheless, prior emotions consistently and significantly contributed to explaining subsequent achievement, and prior achievement consistently and significantly contributed to explaining subsequent emotional development.

We found reciprocal links with emotions both for students' math grades and their standardized achievement test scores. Strong correlations between math grades and test scores (r range .61 to .67; Table 1) suggest that the grades were valid indicators of students' compentencies in maths. These end of year grades reflect students' cumulative performance throughout the school year relative to the curriculum in math, thus further supporting their validity. The consistency of the relations between emotions and students' grades and test scores documents the generalizability of the relations between emotions and math performance. However, the reciprocal relations between emotions and achievement were more pronounced for grades as compared with test scores. This finding is in line with previous research and theory suggesting that school grades should be more strongly related to emotional and motivational factors as compared with test scores, given that they represent a more salient source of feedback to students (see also Hattie, 1992; Marsh et al., 2005; Pekrun et al., 2017).

The results support previous cross-sectional evidence that emotions and academic achievement are related during the elementary school years (Lichtenfeld et al., 2012; Raccanello et al., 2018), but extend them by examining the directionality of these relations. Counter to traditional unidirectional views focusing on the effects of emotions on performance (e.g., research investigating mediating mechanisms such as cognitive interference explaining the effects of test anxiety on students' achievement; Zeidner, 1998, 2014; or research on the influence of mood states on cognitive performance; Clore & Huntsinger, 2009), the findings show that reciprocal effects of achievement on the development of emotions are no less important.

Effects of Gender and SES

The findings on gender differences are in line with previous research indicating that girls report less enjoyment and more anxiety in mathematics. The results suggest that girls show a more maladaptive pattern of math emotions as compared with boys. Lower competence beliefs and perceptions of the value of mathematics may explain these differences (Goetz et al., 2013). In addition, boys showed slightly better math achievement than girls, which may also help to explain the differences in emotions. Previous findings on gender differences in math achievement are inconsistent. A cross-cultural study across 69 countries found that mean effect sizes were very small, supporting the gender similarity hypothesis (Else-Quest et al., 2010). Likewise, PISA data from 2018 shows that the gender gap is generally small and did not significantly change between 2009 and 2018 (OECD, 2019; Schleicher, 2019). Given the small size of gender effects on achievement in the present study, our results are in line with these findings.

Parental education and cultural capital showed only few effects on students' math emotions. However, both SES variables positively predicted their math grades and test scores. Specifically, parental education yielded moderately positive effects on students' achievement. Over and above this effect, cultural capital also had a small positive effect. These effects are in line with previous findings and may be explained by parental expectations (Davis-Kean, 2005) and parental involvement (Zhang, 2020). Parents with higher parental education levels and cultural capital may hold higher expectations and show more school involvement, which may promote students' achievement. In sum, all three covariates had predictive effects, either on students' emotions or their achievement or both.

Limitations and Directions for Future Research

The present study yielded robust findings on the links between students' emotions and their achievement while controlling for initial links, autoregressive stability, and critical background variables. It significantly advances research in the field by focusing on a critically important age group that has not received much attention in research on achievement emotions, and by generating a more fine-grained picture of the development of these emotions by distinguishing between different emotions as well as different academic settings in which these emotions occur. Nevertheless, the current findings should be interpreted in the context of several limitations, which can also serve as a basis to identify directions for future research.

First, the study produced strong evidence supporting reciprocal predictive effects of emotions on achievement and vice versa, but it was based on a longitudinal rather than an experimental design. As such, although the power of non-experimental designs to derive causal conclusions should not be underestimated (Grosz et al., 2020), firm conclusions about cause-effect relations cannot be derived. While the results are based on multi-wave longitudinal SEMs controlling for initial correlations, autoregressive effects, and relevant related constructs, there could still be other variables that may have contributed to the observed links between emotions and achievement. Nevertheless, longitudinal studies in real-life contexts such as the present one are considered to be more ecologically valid, include more representative samples as compared with small-scale experimental studies, have higher statistical power due to large sample sizes, and bear fewer ethical concerns than experimental emotion studies which rely on manipulation procedures. Future research would do well to expand upon the present findings by using both longitudinal and experimental designs.

Second, the current research examined three different achievement emotions as experienced by children in one specific academic domain. Given limited attention span in young children (Betts et al., 2006; Lewis et al., 2017), it was necessary to keep administration time to a minimum. As such, it was not possible to assess a more exhaustive range of emotions and academic domains. Thus, it is open to question whether the observed reciprocal effects would also occur for other emotions not assessed herein. Specifically, future research on other positive achievement emotions such as pride and relief, as well as negative achievement emotions such as anger and hopelessness, in elementary school would further our knowledge of how both activating and deactivating positive and negative emotions develop across the early years of schooling.

Third, emotions were assessed with self-report measures in the current study. Self-report has been considered as indispensable to assess students' emotions (e.g., Pekrun, 2020), but is susceptible to response biases such as social desirability and acquiescent responding (Kreitchmann et al., 2019; Paulhus, 1991) and to memory biases (Levine et al., 2006; Levine & Safer, 2002; Wilson et al., 2003). Specifically, acquiescent responding has been found to be more pronounced in children and adolescents, which may reduce the validity of self-report (Soto et al., 2008). Due to controlling for autoregressive effects for emotions, the influence of such biases on the links with achievement may be reduced in the present longitudinal analyses. Nevertheless, future research would do well to include physiological indicators, facial expression analysis, and implicit measures of emotions to gain a more fine grained picture of emotional processes and their linkages with academic achievement during the early years of schooling. Specifically, adding physiological and behavioral measures may shed light on the interplay and relative importance of different component processes of emotion in explaining emotion-performance links in elementary school students.

Furthermore, the current study focused on changes over the course of several school years. Data collections took place towards the end of each school year to gauge the linkages between students' reports of their emotions during the year and their cumulative achievement during the year (for a similar approach, see, e.g., Pekrun et al., 2017). However, future studies could amend this approach by using ambulatory assessment methods, such as experience sampling methodology, which has been successfully used in studies of achievement emotions (Ahmed et al., 2010; Goetz et al., 2014; Nett et al., 2017). The use of these methodologies may be suited to capture the situational dynamics and immediate performance effects of situational emotional experiences in elementary school.

In addition, recent research found that epistemic emotions such as confusion and frustration are frequently experienced by elementary school students in math (Munzar et al., 2021). While several studies investigated antecedents and outcomes of students' state epistemic emotions (e.g., Chevrier et al., 2019; see also Vogl et al., 2020), an interesting avenue for future research would be to investigate elementary school students' trait epistemic emotions and their interrelations with achievement. Moreover, future research would do well to investigate whether the present findings generalize to domains other than mathematics and cultural contexts other than Germany.

Finally, future studies could explore effects of achievement emotions in the context of productive failure. Much resdevlopmentearch indicates that failure is linked to enhanced levels of negative emotions such as anxiety (for meta-analyses, see von der Embse et al., 2018; Zhang et al., 2019). On the other hand, studies on productive failure have shown that making mistakes during problem-solving can support learners in their strivings to persist even in the face of failure, which may promote learning and performance (Kapur, 2010; McNamara, 2001; Schwartz & Martin, 2004). However, a recent meta-analysis (Sinha & Kapur, 2021) has suggested that failures in complex learning prior to a memory consolidation phase can be detrimental rather than beneficial for young students' learning (2nd to 5th graders). Thus, it would be interesting to explore if negative emotions during initial phases of complex learning (such as frustration and anxiety) vary between younger and older students, and to investigate if these differences translate into different effects on motivation and performance.

Implications for Practice

Important implications for educational practice can be derived from the present work. First, the results suggest that children's enjoyment of mathematics is increasingly undermined as students move through elementary school. This finding underlines that teachers and parents should be encouraged to invest an effort to foster students' positive emotions towards mathematics. Promising findings on how this can be achieved come, for instance, from research on utility value interventions which indicate that college students partaking in such an intervention developed more positive expectancy and value appraisals (Hulleman et al., 2017), which are deemed to be important antecedents of positive achievement emotions (Pekrun, 2006). Moreover, initial evidence indicates that utility value interventions may already be beneficial for elementary school students in fostering their interest and willingness to engage (Shin et al., 2019), which may be beneficial for the development of their emotions.

Second, the findings from our cross-lagged analyses provide evidence that emotions may influence elementary school students' academic achievement, over and above the impact of prior performance. This further emphasizes the necessity to focus on students' emotional experiences and highlights that fostering positive and reducing negative achievement emotions can help to promote students' learning and performance.

Finally, the results suggest that achievement, in turn, had a substantial impact on students' emotions. This was especially true for the grades that students received in math. As such, the findings imply that achievement and related feedback likely are central determinants of students' emotional development at school. Fredricks and Eccles (2002) suggested that the downward trajectory of students' competence beliefs may be attributed to the fact that they tend to rely more and more on comparative standards to judge their abilities (Spinath & Spinath, 2005; Stipek & MacIver, 1989). This trend in competence beliefs across elementary school may be responsible for the maladaptive changes in students' achievement emotions. Thus, providing positive feedback on students' performance and establishing opportunities to experience success by implementing individual rather than comparative standards for feedback may be a critically important avenue to establish a beneficial pattern of emotions in the early years of schooling. Doing so may lead to a virtuous cycle of positive emotions fostering children's achievement, and children's success facilitating their positive emotional development.

References

Ahmed, W., van der Werf, G., Minnaert, A., & Kuyper, H. (2010). Students' daily emotions in the classroom: Intra-individual variability and appraisal correlates. *British Journal of Educational Psychology, 80*, 583–597. http://dx.doi.org/10.1348/000709910X498544

Archambault, I., Eccles, J. S., & Vida, M. N. (2010). Ability self-concepts and subjective value in literacy: Joint trajectories from grades 1 through 12. *Journal of Educational Psychology, 102*(4), 804-816. https://doi.org/10.1037/a0021075

Arens, A. K., Marsh, H. W., Pekrun, R., Lichtenfeld, S., Murayama, K., & vom Hofe, R. (2017). Math self-concept, grades, and achievement test scores: Long-term reciprocal effects across five waves and three achievement tracks. *Journal of Educational Psychology, 109*(5), 621-634. https://doi.org/10.1037/edu0000163

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2020). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*. Advance online publication. https://doi.org/10.1037/bul000030

Betts, J., McKay, J., Maruff, P., & Anderson, V. (2006). The development of sustained attention in children: The effect of age and task load. *Child Neuropsychology, 12*(3), 205-221. https://doi.org/10.1080/09297040500488522

Bieg, M., Goetz, T., & Hubbard, K. (2013). Can I master it and does it matter? An intraindividual analysis on control–value antecedents of trait and state academic emotions. *Learning and Individual Differences, 28*, 102-108. https://doi.org/https://doi.org/10.1016/j.lindif.2013.09.006

Bohn-Gettler, C. M. (2019). Getting a grip: The PET framework for studying how reader emotions influence comprehension. Discourse Processes, 56(5-6), 386-401. https://doi.org/10.1080/0163853X.2019.1611174

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Sage.

Camacho-Morles, J., Slemp, G. R., Pekrun, R., Loderer, K., Hou, H., & Oades, L. G. (2021). Activity achievement emotions and academic performance: A meta-analysis. *Educational Psychology Review.* Advance online publication. https://doi.org/10.1007/s10648-020-09585-3

Cargnelutti, E., Tomasetto, C., & Passolunghi, M. C. (2017). How is anxiety related to math performance in young students? A longitudinal study of Grade 2 to Grade 3 children. *Cognition and Emotion, 31*(4), 755-764. https://doi.org/10.1080/02699931.2016.1147421

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464-504. https://doi.org/10.1080/10705510701301834

Chevrier, M., Muis, K. R., Trevors, G. J., Pekrun, R., & Sinatra, G. M. (2019). Exploring the antecedents and consequences of epistemic emotions. *Learning and Instruction, 63*, 101209. https://doi.org/https://doi.org/10.1016/j.learninstruc.2019.05.006

Darensbourg, A. M., & Blake, J. J. (2013). Predictors of achievement in African American students at risk for academic failure: The roles of achievement values and behavioral engagement. *Psychology in the Schools, 50*(10), 1044-1059.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment.

Journal of Family Psychology, 19(2), 294-304. https://doi.org/10.1037/0893-3200.19.2.294

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75-146). Freeman.

Eccles, J., S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology.* Advance online publication. https://doi.org/10.1016/j.cedpsych.2020.101859

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103-127. https://doi.org/10.1037/a0018053

Enders, C. K. (2010). *Applied missing data analysis*. Guilford.

Fiedler, K., & Beier, S. (2014). Affect and cognitive processing in educational contexts. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), International handbook of emotions in education (pp. 36-55). Taylor & Francis.

Forsblom, L., Pekrun, R., Loderer, K., & Peixoto, F. (2021). Cognitive appraisals, achievement emotions, and students' math achievement: A longitudinal analysis. *Journal of Educational Psychology.* Advance online publication. https://doi.org/10.1037/edu0000671

Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology, 38*(4), 519-533. https://doi.org/10.1037//0012-1649.38.4.519

Frenzel, A. C., Thrash, T., Pekrun, R., & Goetz, T. (2007). Achievement emotions in Germany and China. A cross-cultural validation of the academic emotions questionnaire-mathematics. *Journal of Cross Cultural Psychology, 38*(3), 302-309. https://doi.org/10.1177/0022022107300276

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. Personality and Individual Differences, 102, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

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. https://doi.org/10.1177/0956797613486989

Goetz, T., Frenzel, A. C., Hall, N. C., Nett, U. E., Pekrun, R., & Lipnevich, A. A. (2014). Types of boredom: An experience sampling approach. *Motivation and Emotion, 38,* 401– 419. http://dx.doi.org/10.1007/s11031-013-9385-y

Goetz, T., Frenzel, A. C., Pekrun, R., Hall, N. C., & Lüdtke, O. (2007). Between- and within-domain relations of students' academic emotions. *Journal of Educational Psychology, 99*(4), 715-733. https://doi.org/10.1037/0022-0663.99.4.715

Goetz, T., & Hall, N. C. (2013). Emotion and achievement in the classroom. In J. Hattie & E. M. Anderman (Eds.), *International guide to student achievement* (pp. 192-195). Routledge.

Goetz, T., Hall, N. C., & Krannich, M. (2019). Boredom. In K. A. Renninger & S. E. Hidi (Eds.), The Cambridge handbook on motivation and learning (pp. 465-486). Cambridge University Press.

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The taboo against explicit causal inference in non-experimental psychology. *Perspectives on Psychological Science, 15*(5), 1243-1255. https://doi.org/10.1177/1745691620921521

Gunderson, E. A., Park, D., Maloney, E. A., Beilock, S. L., & Levine, S. C. (2018). Reciprocal relations among motivational frameworks, math anxiety, and math achievement in early elementary school. *Journal of Cognition and Development, 19*(1), 21-46. https://doi.org/10.1080/15248372.2017.1421538

Harlen, W. (2008). Trusting teachers' judgment. In S. Swaffield (Ed.), *Unlocking assessment: Understanding for reflection and application*, (pp. 138-153). Routledge.

Hattie, J. (1992). *Self-concept*. Erlbaum.

Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Buhner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: a cautionary note on the usefulness of cut-off values of fit indices. *Psychological Methods, 16*(3), 319-336. https://doi.org/10.1037/a0024917

Hu, L. t., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hulleman, C. S., Kosovich, J. J., Barron, K. E., & Daniel, D. B. (2017). Making connections: Replicating and extending the utility value intervention in the classroom. *Journal of Educational Psychology, 109*(3), 387-404. https://doi.org/10.1037/edu0000146

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one though twelve. *Child Development, 73*(2), 509-527. https://doi.org/10.1111/1467-8624.00421

Jelicic, H., Phelps, E., & Lerner, R. M. (2009). Use of missing data methods in longitudinal studies: the persistence of bad practices in developmental psychology. *Developmental Psychology, 45*(4), 1195-1199. https://doi.org/10.1037/a0015665

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science, 38*(6), 523-550. https://doi.org/10.1007/s11251-009-9093-x

Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist, 73*(1), 1-2. https://doi.org/10.1037/amp0000263

Kreitchmann, R. S., Abad, F. J., Ponsoda, V., Nieto, M. D., & Morillo, D. (2019). Controlling for response biases in self-report scales: Forced-choice vs. psychometric modeling of likert items. *Frontiers in Psychology, 10*(2309). https://doi.org/10.3389/fpsyg.2019.02309

Larson, R. W., & Richards, M. H. (1991). Boredom in the middle school years: Blaming schools versus blaming students. *American Journal of Education, 99*(4), 418-443. https://doi.org/10.1086/443992

Levine, L. J., & Safer, M. A. (2002). Sources of bias in memory for emotions. *Current Directions in Psychological Science, 11*(5), 169-173. https://doi.org/10.1111/1467-8721.00193

Levine, L., Safer, M., & Lench, H. C. (2006). Remembering and misremembering emotions. In L. J. Sanna & E. C. Chang (Eds.), *Judgments over time: The interplay of thoughts, feelings, and behaviors* (pp. 271-290). Oxford University Press.

Lewis, F. C., Reeve, R. A., Kelly, S. P., & Johnson, K. A. (2017). Sustained attention to a predictable, unengaging Go/No-Go task shows ongoing development between 6 and

11 years. *Attention, Perception, & Psychophysics, 79*(6), 1726-1741. https://doi.org/10.3758/s13414-017-1351-4

Li, X., Huebner, E. S., & Tian, L. (2021). Relations between achievement task values and academic achievement and depressive symptoms in Chinese elementary school students: Variable-centered and person-centered perspectives. *School Psychology, 36*(3), 167–180. https://doi.org/10.1037/spq0000384

Lichtenfeld, S., Pekrun, R., Stupnisky, R. H., Reiss, K., & Murayama, K. (2012). Measuring students' emotions in the early years: The Achievement Emotions Questionnaire-Elementary School (AEQ-ES). *Learning and Individual Differences, 22*(2), 190-201. https://doi.org/10.1016/j.lindif.2011.04.009

Linacre, J. M. (2018). *Winsteps® (Version 4.3.1) [Computer Software]*. In Winsteps.com.

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. https://doi.org/10.1016/j.adolescence.2003.11.003

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*(3), 391-410. https://doi.org/10.1037/0033-2909.103.3.391

Marsh, H. W., Byrne, B. M., & Craven, R. (1992). Overcoming problems in confirmatory factor analyses of MTMM data: The correlated uniqueness model and factorial invariance. *Multivariate Behavioral Research, 27*(4), 489-507. https://doi.org/10.1207/s15327906mbr2704_1

Marsh, H. W., Craven, R., & Debus, R. (1998). Structure, stability, and development of young children's self-concepts: A multicohort–multioccasion study. *Child Development*, *69*(4), 1030-1053. https://doi.org/10.2307/1132361

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cut-off values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320-341. https://doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*(1), 85-110. https://doi.org/10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Pekrun, R., Lichtenfeld, S., Guo, J. S., Arens, A. K., & Murayama, K. (2016). Breaking the double-edged sword of effort: Developmental equilibrium and longitudinal relations among effort, achievement, and academic self-concept. *Developmental Psychology, 52*(8), 1273-1290. https://doi.org/10.1037/dev0000146

Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology, 54*(2), 263-280. https://doi.org/10.1037/dev0000393

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology*, *111*(2), 331–353. https://doi.org/10.1037/edu0000281

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397-416. https://doi.org/http://www.jstor.org/stable/3696511

Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and procedures: TIMSS 2019 Technical Report*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls.bc.edu/timss2019/methods

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*(1), 577-605. https://doi.org/10.1146/annurev.psych.60.110707.163612

McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, *55*(1), 51–62. doi:10.1037/h0087352.

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. https://doi.org/10.1037/0022-0663.82.1.60

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance [journal article]. *Psychometrika, 58*(4), 525-543. https://doi.org/10.1007/bf02294825

Muenks, K., Wigfield, A., & Eccles, J. S. (2018). I can do this! The development and calibration of children’s expectations for success and competence beliefs. *Developmental Review*, *48*(1), 24-39. https://doi.org/10.1016/j.dr.2018.04.001

Munzar, B., Muis, K. R., Denton, C. A., & Losenno, K. (2021). Elementary students' cognitive and affective responses to impasses during mathematics problem solving. *Journal of Educational Psychology, 113*(1), 104-124. https://doi.org/10.1037/edu0000460

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nett, U. E., Bieg, M., & Keller, M. M. (2017). How much trait variance is captured by measures of academic state emotions. *European Journal of Psychological Assessment, 33,* 239 – 255. http://dx.doi.org/10.1027/1015-5759/a000416

Nicholls, J. G. (1979). Development of perception of own attainment and causal attributions for success and failure in reading. *Journal of Educational Psychology, 71*(1), 94-99. https://doi.org/10.1037/0022-0663.71.1.94

Nicholls, J. G ., & Miller, A. T. (1984). Reasoning about the ability of self and other: A developmental study. *Child Development, 55*(6), 1990–1999. https://doi.org/10.2307/1129774

OECD. (2011). PIAAC conceptual framework of the background questionnaire main survey. Retrieved from http://www.oecd.org/skills/piaac/PIAAC(2011_11)MS_BQ_Conceptual Framework_1%20Dec%202011.pdf

OECD. (2019). *PISA 2018 Results (Volume II): Where All Students Can Succeed*. PISA, OECD Publishing. https://doi.org/10.1787/b5fd1b8f-en

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2020). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology, 120*(4). https://doi.org/10.1037/pspp0000358

Parsons, J. E., & Ruble, D. N. (1977). The development of achievement-related expectancies. *Child Development, 48*(3), 1075-1079. https://doi.org/10.2307/1128364

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). Academic Press.

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. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. (2017). Emotion and achievement during adolescence. *Child Development Perspectives, 11*(3), 215-221. https://doi.org/10.1111/cdep.12237

Pekrun, R. (2018). Control-value theory: A social-cognitive approach to achievement emotions. In G. A. D. Liem & D. M. McInerney (Eds.), *Big theories revisited 2* (pp. 162-190). Information Age Publishing.

Pekrun, R. (2020). Self-report is indispensible to assess students' learning. *Frontline Learning Research, 8*(3), 185-193. https://doi.org/10.14786/flr.v8i3.637

Pekrun, R. (2021). Self-appraisals and emotions: A generalized control-value approach. In T. Dicke, F. Guay, H. W. Marsh, R. G. Craven, & D. M. McInerney (Eds). *Self – a multidisciplinary concept.* Information Age Publishing.

Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 13-36). Academic Press.

Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Control-value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, *102*(3), 531-549. https://doi.org/10.1037/a0019243

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. https://doi.org/10.1016/j.cedpsych.2010.10.002

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. https://doi.org/10.1207/S15326985EP3702_4

Pekrun, R., Hall, N. C., Goetz, T., & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology, 106*(3), 696-710. https://doi.org/10.1037/a0036006

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development, 88*(5), 1653-1670. https://doi.org/10.1111/cdev.12704

Pekrun, R., & Linnenbrink-Garcia, L. (Eds.). (2014). *International handbook of emotions in education*. Taylor & Francis.

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 120-141). Routledge.

Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., & Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology, 84*(1), 152-174. https://doi.org/10.1111/bjep.12028

Putwain, D. W., Becker, S., Symes, W., & Pekrun, R. (2018). Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. *Learning and Instruction, 54,* 73-81. https://doi.org/10.1016/j.learninstruc.2017.08.004

Putwain, D. W., Pekrun, R., Nicholson, L. J., Symes, W., Becker, S., & Marsh, H. W. (2018). Control-value appraisals, enjoyment, and boredom in mathematics: A longitudinal latent interaction analysis. *American Educational Research Journal, 55*(6), 1339-1368. https://doi.org/10.3102/0002831218786689

Raccanello, D., Brondino, M., Moè, A., Stupnisky, R. H., & Lichtenfeld, S. (2018). Enjoyment, boredom, anxiety in elementary schools in two domains: Relations with achievement. *The Journal of Experimental Education*, *87*(3), 449-469. https://doi.org/10.1080/00220973.2018.1448747

Rask, K. (2010). Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences. *Economics of Education Review, 29*(6), 892-900. https://doi.org/10.1016/j.econedurev.2010.06.013

Reiss, K., Heinze, A., & Pekrun, R. (2008). Mathematische Kompetenz und ihre Entwicklung in der Grundschule. In M. Prenzel, I. Gogolin, & H.-H. Krüger (Eds.), *Kompetenzdiagnostik: Zeitschrift für Erziehungswissenschaft* (pp. 107-127). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90865-6_7

Reiss, K., & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel Panhuizen, K. Reiss & G. Walther (Hrsg.), Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule (S. 120–141). Weinheim: Beltz.

Rosel, J., & Plewis, I. (2008). Longitudinal data analysis with structural equations. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *4*(1), 37-50. https://doi.org/10.1027/1614-2241.4.1.37

Sağkal, A. S., & Sönmez, M. T. (2021). The effects of perceived parental math support on middle school students' math engagement: the serial multiple mediation of math self-efficacy and math enjoyment. *European Journal of Psychology of Education*. Advance online publication. https://doi.org/10.1007/s10212-020-00518-w

Scherer, K. R., & Moors, A. (2019). The emotion process: Event appraisal and component differentiation. *Annual Review of Psychology*, *70*(1), 719-745. https://doi.org/10.1146/annurev-psych-122216-011854

Schleicher, A. (2019). *PISA 2018: Insights and Interpretations*. OECD Publishing. https://www.oecd.org/pisa/PISA%202018%20Insights%20and%20Interpretations%20FINAL%20PDF.pdf

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, *22*(2), 129–184. doi:10.1207/s1532690xci2202_1.

Shao, K., Pekrun, R., Marsh, H. W., & Loderer, K. (2020). Control-value appraisals, achievement emotions, and foreign language performance: A latent interaction analysis. *Learning and Instruction*, *69*, 101356.

Shin, D. D., Lee, M., Ha, J. E., Park, J. H., Ahn, H. S., Son, E., Chung, Y., & Bong, M. (2019). Science for all: Boosting the science motivation of elementary school students with utility value intervention. *Learning and Instruction*, *60*, 104-116. https://doi.org/https://doi.org/10.1016/j.learninstruc.2018.12.003

Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research, 91*(5), 761-798. https://doi.org/10.3102/00346543211019105

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology, 94*(4), 718-737. https://doi.org/10.1037/0022-3514.94.4.718

Spinath, B., & Spinath, F. M. (2005). Development of self-perceived ability in elementary school: The role of parents' perceptions, teacher evaluations, and intelligence. *Cognitive Development, 20*(2), 190-204. https://doi.org/10.1016/j.cogdev.2005.01.001

Spinath, B., & Steinmayr, R. (2008). Longitudinal analysis of intrinsic motivation and competence beliefs: Is there a relation over time? *Child Development*, *79*(5), 1555-1569. https://doi.org/10.1111/j.1467-8624.2008.01205.x

Steenkamp, J. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*(1), 78-90. https://doi.org/10.1086/209528

Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research*, *2*(1), 21-33.

Steyer, R., Partchev, I., & Shanahan, M. J. (2000). Modeling true intra-individual change in structural equation models: The case of poverty and children’s psychosocial adjustment. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 109–126). Erlbaum.

Stipek, D. J., & Mac Iver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development*, *60*(3), 521-538. https://doi.org/10.2307/1130719

Tseng, K.-H., Chang, C.-C., Lou, S.-J., & Chen, W.-P. (2013). Attitudes towards science, technology, engineering and mathematics (STEM) in a project-based learning (PjBL) environment. *International Journal of Technology and Design Education, 23*(1), 87-102. https://doi.org/10.1007/s10798-011-9160-x

Vierhaus, M., Lohaus, A., & Wild, E. (2016). The development of achievement emotions and coping/emotion regulation from primary to secondary school. *Learning and Instruction, 42*, 12-21. https://doi.org/10.1016/j.learninstruc.2015.11.002

Vogl, E., Pekrun, R., Murayama, K., & Loderer, K. (2020). Surprised–curious–confused: Epistemic emotions and knowledge exploration. *Emotion, 20*(4), 625-641. https://doi.org/10.1037/emo0000578

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders, 227*, 483-493. https://doi.org/10.1016/j.jad.2017.11.048

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, *30*(1), 1-35. https://doi.org/10.1016/j.dr.2009.12.001

Wigfield, A., & Eccles, J. S. (2020). 35 years of research on students' subjective task values and motivation: A look back and a look forward. In A. J. Elliot (Ed.), *Advances in Motivation Science* (Vol. 7, pp. 161-198). Elsevier.

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (pp. 657–700). Wiley.

Wigfield, A., Eccles, J. S., & Rodriguez, D. (1998). The development of children’s motivation in school contexts. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 23, pp. 73-118). AERA.

Wilson, T. D., Meyers, J., & Gilbert, D. T. (2003). “How happy was I, anyway?” A retrospective impact bias. *Social Cognition, 21*(6), 421-446. https://doi.org/10.1521/soco.21.6.421.28688

Zahorik, J. A. (1996). Elementary and secondary teachers' reports of how they make learning interesting. *The Elementary School Journal, 96*(5), 551-564. https://doi.org/10.1086/461844

Zhang, F., Jiang, Y., Ming, H., Ren, Y., Wang, L., & Huang, S. (2020). Family socioeconomic status and children’s academic achievement: The different roles of parental academic involvement and subjective social mobility. *British Journal of Educational Psychology, 90*(3), 561-579. https://doi.org/10.1111/bjep.12374

Zhang, L., & Lee, K. A. (2011). Decomposing achievement gaps among OECD countries. *Asia Pacific Education Review*, *12*(3), 463-474. https://doi.org/10.1007/s12564-011-9151-3

Zhang, J., Zhao, N., & Kong, Q. P. (2019). The relationship between math anxiety and math performance: A meta-analytic investigation. *Frontiers in Psychology, 10*(1613). https://doi.org/10.3389/fpsyg.2019.01613

Table 1

Descriptive Statistics and Pearson Product-Moment Correlations and Latent Correlations for Emotions and Achievement

	Year	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1 Learning enjoyment	2	3.83	1.12	*.91*		**-.85**	**-.65**		**.24**	**.15**	**.13**	.04	**-.08**
	3	3.62	1.08	*.91*	**.85**	**-.91**	**-.61**	**-.44**	**.30**	**.22**	**.11**	.03	-.03
	4	3.31	1.01	*.92*	**.85**	**-.87**	**-.54**	**-.46**	**.24**	**.18**	.10	.01	-.03
2 Test enjoyment	3	3.33	1.21	**.71**	*.86*	**-.64**	**-.58**	**-.60**	**.35**	**.29**	**.23**	.05	.04
	4	2.93	1.17	**.72**	*.85*	**-.51**	**-.46**	**-.61**	**.36**	**.33**	**.19**	.02	.09
3 Learning boredom	2	1.75	1.01	**-.75**		*.91*	**.73**		**-.25**	**-.11**	-.07	-.05	.05
	3	1.78	.93	**-.74**	**-.52**	*.90*	**.71**	**.48**	**-.20**	**-.15**	.03	-.03	.02
	4	1.88	.92	**-.69**	**-.43**	*.92*	**.48**	**.42**	**-.15**	**-.13**	.12	-.07	.02
4 Learning anxiety	2	1.59	.79	**-.52**		**.62**	*.83*		**-.38**	**-.36**	**-.13**	**-.13**	-.04
	3	1.50	.69	**-.49**	**-.50**	**.59**	*.85*	**.82**	**-.39**	**-.34**	**-.12**	**-.20**	**-.17**
	4	1.50	.71	**-.45**	**-.43**	**.40**	*.88*	**.78**	**-.39**	**-.39**	**-.14**	**-.15**	**-.17**
5 Test anxiety	3	1.90	.94	**-.39**	**-.57**	**.43**	**.70**	*.88*	**-.38**	**-.29**	**-.13**	**-.16**	**-.14**
	4	2.07	1.03	**-.39**	**-.58**	**.36**	**.70**	*.92*	**-.42**	**-.43**	**-.11**	-.13	**-.24**
6 Achievement	2	4.68	1.01	**.22**		**-.22**	**-.33**			**.61**	.07	**.20**	**.33**
(grades)	3	4.62	.90	**.28**	**.34**	**-.18**	**-.35**	**-.37**		**.60**	**.10**	**.20**	**.26**
	4	4.36	1.04	**.23**	**.37**	**-.17**	**-.40**	**-.44**		**.63**	.04	**.18**	**.30**
7 Achievement	2	4.68	1.01	**.13**		**-.12**	**-.33**		**.61**	*.74*	**.17**	**.16**	**.29**
(test scores)	3	4.62	.90	**.20**	**.28**	**-.13**	**-.30**	**-.28**	**.60**	*.75*	**.16**	**.18**	**.29**
	4	4.36	1.04	**.20**	**.34**	**-.14**	**-.36**	**-.42**	**.67**	*.77*	**.11**	**.20**	**.33**
8 Gender	2	1.49	.50	**.11**	**.23**	-.07	**-.12**	**-.13**	.07	**.18**		.08	-.03
9 Cultural capital	2	3.43	1.23	.02	.07	-.06	**-.10**	**-.14**	**.20**	**.16**	.08		**.21**
10 Parents‘ education	2	3.01	.98	-.07	.03	.04	-.04	**-.13**	**.32**	**.29**	.04	**.20**	

Note. [a] Correlations with demographic variables are reported for first measurement wave. Coefficients below main diagonal are manifest correlations. Coefficients above main diagonal are latent correlations. Coefficients in italic are Cronbach’s Alphas for the self-report scales and person reliabilities for the Rasch scores. **Bold** coefficients $p < .05$.

Table 2

Fit Indexes for Latent Neighbor Change and Reciprocal Effects Models

	χ^2	*df*	CFI	TLI	RMSEA	SRMR
Latent Neighbor Change Models						
Learning enjoyment	188.093***	127	.987	.985	.027	.033
Test enjoyment	7.883	9	1.000	1.001	.000	.022
Learning boredom	311.685***	181	.972	.967	.033	.040
Learning anxiety	302.362***	181	.951	.943	.032	.050
Test anxiety	69.902***	37	.979	.975	.039	.036
Reciprocal Effects Models with Grades						
	Cross-paths and autoregressive effects freely estimated					
Learning enjoyment	360.579***	221	.974	.968	.034	.041
Test enjoyment	71.237***	29	.976	.959	.047	.033
Learning boredom	492.292***	294	.965	.958	.035	.040
Learning anxiety	500.704***	294	.940	.930	.032	.047
Test anxiety	287.989***	141	.954	.946	.039	.042
	Cross-paths and autoregressive effects invariant across waves					
Learning enjoyment	367.767***	225	.973	.967	.034	.042
Learning boredom	503.437***	298	.964	.957	.032	.045
Learning anxiety	525.842***	299	.935	.924	.034	.055

** $p < .01$. *** $p < .001$.

Table 3

Latent Neighbor Change Models: Means and Standard Deviations for Latent Variables across the Three Years

	Year 2		Year 3		Year 4	
Variable	*M*	*SE*	*M*	*SE*	*M*	*SE*
Enjoyment						
Learning enjoyment	4.04	.06	$3.91^{\dagger}$	.07	3.64^{***}	.05
Test enjoyment			3.31	.09	2.88^{***}	.09
Boredom						
Learning boredom	1.72	.05	1.70	.07	1.79	.06
Anxiety						
Learning anxiety	1.34	.04	1.23^{**}	.05	1.22	.03
Test anxiety			2.50	.10	$2.63^{\dagger}$	.07

Note. Changes to previous wave: † $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4

Reciprocal Effects Models for Emotions and Grades: Standardized Factor Loadings, Path Coefficients, and Residual Variances

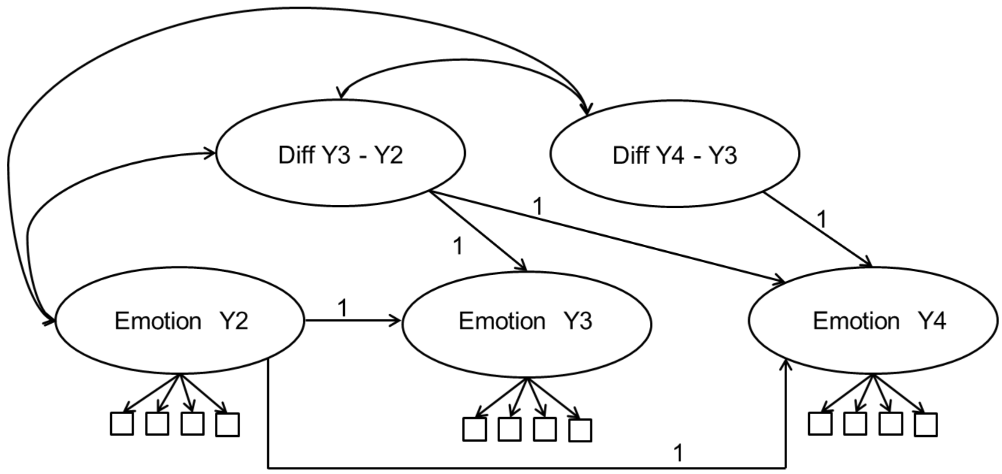
	Enjoyment				Boredom		Anxiety			
	Learning enjoyment	Grades	Test enjoyment	Grades	Learning boredom	Grades	Learning anxiety	Grades	Test anxiety	Grades
Factor loadings	.71-.92 [a]		.67-.91 [a]		.72-.88 [a]		.63-.80 [a]		.62-.82 [a]	
Autoregressive effects										
Y2 → Y3	**.59** (.04)	**.66** (.04)			**.49** (.05)	**.68** (.04)	**.57** (.06)	**.62** (.04)		
Y3 → Y4	**.62** (.05)	**.59** (.05)	**.54** (.05)	**.65** (.04)	**.47** (.07)	**.61** (.05)	**.59** (.07)	**.52** (.06)	**.57** (.04)	**.57** (.07)
Cross-lagged effects	*Grades → Learning enjoyment*	*Learning enjoyment → Grades*	*Grades → Test enjoyment*	*Test enjoyment → Grades*	*Grades → Learning boredom*	*Learning boredom → Grades*	*Grades → Learning anxiety*	*Learning anxiety → Grades*	*Grades → Test anxiety*	*Test anxiety → Grades*
Y2 → Y3	**.12** (.03)	**.09** (.04)			**-.10** (.04)	**-.09** (.04)	**-.14** (.04)	**-.19** (.05)		
Y3 → Y4	**.11** (.03)	**.09** (.04)	**.19** (.04)	.02 (.05)	**-.11** (.04)	**-.08** (.03)	**-.15** (.05)	**-.17** (.04)	**-.19** (.06)	-.05 (.05)
Effects of Covariates Y2(for learning-related emotions) / Y3 (for test emotions)										
Gender	**.13** (.05)	.07 (.04)	**.23** (.05)	**.10** (.04)	-.07 (.04)	.07 (.04)	**-.15** (.06)	.08 (.04)	**-.13** (.04)	**.10** (.04)
Cultural capital	.05 (.04)	**.11** (.05)	.02 (.07)	**.13** (.06)	-.06 (.04)	**.12** (.05)	-.11 (.06)	**.13** (.05)	-.12 (.08)	**.13** (.06)
Parental education	**-.10** (.04)	**.29** (.06)	.04 (.06)	**.23** (.05)	.06 (.03)	**.29** (.05)	-.03 (.06)	**.29** (.05)	**-.12** (.05)	**.23** (.05)
Residual Variances										
Y2	.97	.88			.99	.89	.99	.88		
Y3	.60	.48	.95	.91	.73	.47	.73	.46	.95	.91
Y4	.57	.50	.58	.46	.74	.57	.74	.47	.51	.46

Note. Y2 = Year 2, Y3 = Year 3, and Y4 = Year 4; Coefficients in parentheses are standard errors. Effects of covariates are reported for first measurement wave. **Bold** coefficients $p < .05$. All models were estimated with strong invariance of parameters across waves.

Figure 1

Basic structure of latent neighbor change models for learning-related emotions (panel A) and test emotions (panel B)

A



B

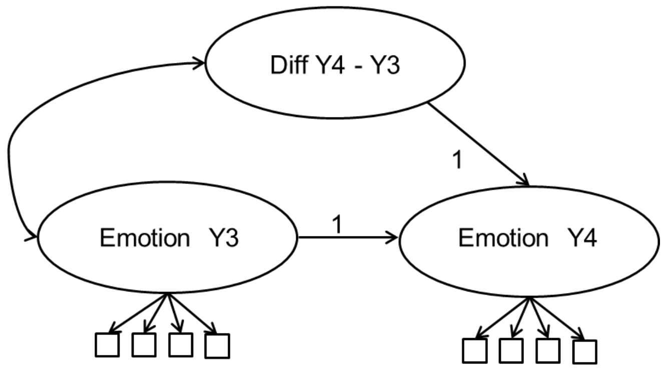
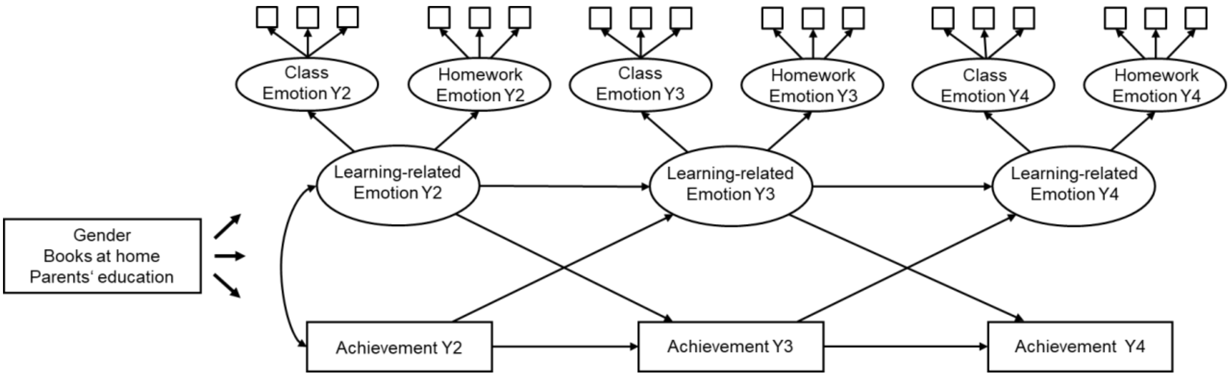


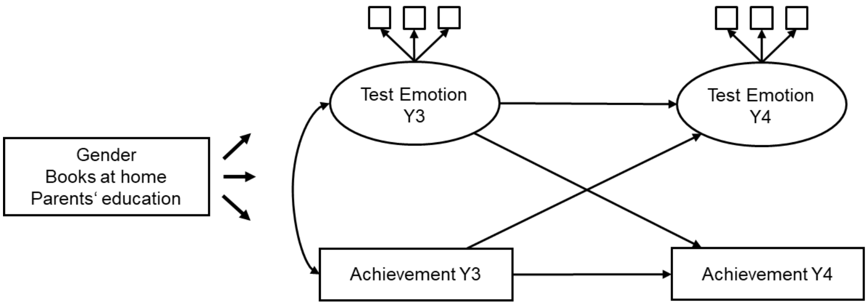
Figure 2

Basic structure of reciprocal effects models for learning-related emotions (panel A) and test emotions (panel B)

A



B



Note. The models include cross-lagged effects, autoregressive effects, and directional paths from the covariates to emotion and achievement at all waves. Correlations between the covariates and between residuals are not displayed.

Supporting Information

Table S1

Emotions Questionnaire–Elementary School (AEQ-ES): Sample Items

Learning emotions	
Enjoyment	I enjoy math class.
Boredom	Math homework bores me to death.
Anxiety	When I think about math class, I get nervous.
Test emotions	
Enjoyment	I look forward to math tests.
Anxiety	I get very nervous during math tests.

Table S2

CFAs for the Discrete Emotion Constructs with Separate Factors for Learning-Related and Test Emotions

Model	χ^2	*df*	CFI	TLI	RMSEA	SRMR
CFA Model Year 2	536.033***	246	.933	.918	.042	.045
CFA Model Year 3	977.720***	514	.935	.924	.037	.048
CFA Model Year 4	1088.545***	515	.937	.927	.041	.046

*** $p < .001$.

Table S3

Measurement Equivalence of Emotion Constructs Across Waves

	Configural Invariance				Metric Invariance				Scalar Invariance				Residual Invariance			
Model	CFI	TLI	RMSEA	SRMR	CFI	TLI	RMSEA	SRMR	CFI	TLI	RMSEA	SRMR	CFI	TLI	RMSEA	SRMR
Learning enjoyment	.993	.990	.022	.027	.992	.990	.022	.031	.987	.984	.027	.033	.977	.975	.035	.042
Test enjoyment	.999	.996	.024	.014	1.000	.999	.012	.019	1.000	1.000	.000	.021	1.000	1.000	.000	.026
Learning boredom	.974	.968	.030	.040	.974	.969	.029	.044	.971	.967	.030	.044	.959	.956	.035	.052
Learning anxiety	.960	.947	.031	.043	.958	.948	.030	.049	.953	.944	.031	.050	.905	.896	.043	.077
Test anxiety	.975	.966	.040	.033	.973	.967	.039	.042	.968	.964	.041	.041	.966	.965	.041	.048

Table S4

Fit Indexes for Reciprocal Effects Models with Test Scores

	χ^2	*df*	CFI	TLI	RMSEA	SRMR
	Cross-paths and autoregressive effects freely estimated					
Learning enjoyment	347.256***	221	.977	.971	.032	.037
Test enjoyment	54.733**	29	.987	.976	.036	.027
Learning boredom	467.559***	294	.970	.964	.030	.040
Learning anxiety	470.106***	294	.950	.941	.030	.045
Test anxiety	265.387***	141	.963	.956	.036	.039
	Cross-paths and autoregressive effects invariant across waves					
Learning enjoyment	373.173***	225	.972	.967	.034	.039
Learning boredom	492.453***	298	.966	.961	.031	.045
Learning anxiety	500.504***	298	.943	.933	.032	.054

** $p < .01$. *** $p < .001$.

Table S5

Reciprocal Effects Models for Emotions and Test Scores: Standardized Factor Loadings, Path Coefficients, and Residual Variances

	Enjoyment				Boredom		Anxiety			
	Learning enjoyment	Test scores	Test enjoyment	Test scores	Learning boredom	Test scores	Learning anxiety	Test scores	Test anxiety	Test scores
Factor loadings	.71-.92 [a]		.66-.91 [a]		.72-.88 [a]		.59-.80 [a]		.62-.82 [a]	
Autoregressive effects										
Y2 → Y3	**.61** (.04)	**.55** (.04)			**.51** (.05)	**.54** (.04)	**.57** (.06)	**.50** (.04)		
Y3 → Y4	**.65** (.04)	**.42** (.05)	**.58** (.05)	**.51** (.05)	**.48** (.07)	**.40** (.04)	**.60** (.07)	**.38** (.04)	**.62** (.04)	**.50** (.05)
Cross-lagged effects	*Test scores → Learning enjoyment*	*Learning enjoyment → Test scores*	*Test scores → Test enjoyment*	*Test enjoyment → Test scores*	*Test scores → Learning boredom*	*Learning boredom → Test scores*	*Test scores → Learning Anxiety*	*Learning anxiety → Test scores*	*Test scores → Test anxiety*	*Test anxiety → Test scores*
Y2 → Y3	.07 (.04)	**.07** (.03)			-.06 (.06)	**-.07** (.03)	-.07 (.04)	**-.13** (.02)		
Y3 → Y4	.05 (.03)	**.06** (.03)	**.08** (.04)	.02 (.05)	-.05 (.05)	**-.06** (.03)	-.06 (.03)	**-.12** (.03)	**-.11** (.03)	**-.09** (.04)
Effects of Covariates Y2 (for learning-related emotions) / Y3 (for test emotions)										
Gender	**.13** (.05)	**.17** (.04)	**.22** (.05)	**.15** (.03)	-.07 (.04)	**.17** (.03)	**-.15** (.06)	**.17** (.03)	**-.12** (.04)	**.15** (.03)
Cultural capital	.04 (.04)	.08 (.04)	.01 (.07)	**.10** (.05)	-.06 (.04)	**.09** (.04)	-.10 (.06)	**.09** (.04)	-.11 (.08)	**.11** (.05)
Education	**-.10** (.04)	**.26** (.05)	.02 (.06)	**.24** (.03)	.07 (.04)	**.26** (.04)	-.02 (.06)	**.26** (.04)	**-.11** (.05)	**.24** (.03)
Residual Variances										
Y2	.97	.89			.99	.89	.96	.89		
Y3	.61	.61	.95	.90	.74	.64	.60	.63	.96	.90
Y4	.57	.73	.62	.66	.74	.74	.59	.74	.54	.66

Note. [a] Range of factor loadings. Coefficients in parentheses are standard errors. Effects of covariates are reported for first measurement wave. **Bold** coefficients $p < .05$. All models were estimated with strong invariance of parameters across waves.