Achievement emotions and academic performance : Longitudinal models of reciprocal effects

Pekrun, Reinhard, Lichtenfeld, Stephanie, Marsh, Herbert W., Murayama, Kou and Goetz, Thomas

Paper accepted for publication in: Child Development

Achievement Emotions and Academic Performance: Longitudinal Models of Reciprocal Effects

Reinhard Pekrun

Stephanie Lichtenfeld

University of Munich

Herbert W. Marsh

Australian Catholic University and University of Oxford

Kou Murayama

University of Reading

Thomas Goetz

University of Konstanz and Thurgau University of Teacher Education

Author Note

Reinhard Pekrun, Department of Psychology, University of Munich, Munich, Germany; Stephanie Lichtenfeld, Department of Psychology, University of Munich, Munich, Germany; Herbert W. Marsh, Institute for Positive Psychology and Education, Australian Catholic University, Sydney, Australia, and Department of Education, University of Oxford, Oxford, UK; Kou Murayama, Department of Psychology, University of Reading, Reading, UK; Thomas Goetz, Department of Empirical Educational Research, University of Konstanz, Konstanz, Germany, and Thurgau University of Teacher Education, Thurgau, Switzerland.

This research was supported by a LMU Research Chair grant awarded to R. Pekrun by the University of Munich and four grants from the German Research Foundation (DFG) to R. Pekrun (PE 320/11-1, PE 320/11-2, PE 320/11-3, PE 320/11-4). Parts of this paper were presented at the annual meeting of the American Educational Research Association, Philadelphia, PA, April 2014, and at the International Congress of Applied Psychology, France, Paris, July 2014.

Correspondence concerning this article should be addressed to Reinhard Pekrun, Department of Psychology, University of Munich, Leopoldstrasse 13, 80802 Munich, Germany. E-mail: pekrun@lmu.de

Abstract

A reciprocal effects model linking emotion and achievement over time is proposed. The model was tested using five annual waves of the PALMA longitudinal study, which investigated adolescents' development in mathematics (grades 5-9; N=3,425 German students; mean starting age=11.7 years; representative sample). Structural equation modeling showed that positive emotions (enjoyment, pride) positively predicted subsequent achievement (math end-of-the-year grades and test scores), and that achievement positively predicted these emotions, controlling for students' gender, intelligence, and family socio-economic status. Negative emotions (anger, anxiety, shame, boredom, hopelessness) negatively predicted achievement, and achievement negatively predicted these emotions. The findings were robust across waves, achievement indicators, and school tracks, highlighting the importance of emotions for students' achievement and of achievement for the development of emotions.

Keywords: achievement emotion, anxiety, academic achievement, mathematics achievement, control-value theory

Research has shown that children's and adolescents' emotions are linked to their academic achievement. Typically, positive emotions such as enjoyment of learning show positive links with achievement, and negative emotions such as test anxiety show negative links (for overviews, see Goetz & Hall, 2013; Pekrun & Linnenbrink-Garcia, 2014; Zeidner, 1998). However, most of the available studies were correlational and do now allow any inferences about the causal ordering of emotion and achievement over time. As such, it remains unclear how the observed links should be interpreted. It is open to question if students' emotions impact their learning, if success and failure at learning influence the development of their emotions, if other variables cause the association, or if several of these possibilities are at work. Given the need to acquire knowledge about the antecedents of both students' achievement and their emotions, this is an issue of considerable theoretical and practical importance. To address this issue, the present investigation went beyond merely observing correlations at a single point in time and attempted to disentangle the temporal ordering of these constructs across multiple waves of data collection and a developmental time span of several school years.

The investigation is based on a reciprocal effects model of emotion and achievement which posits that the two variables reciprocally influence each other over time. This stands in contrast to traditional unidirectional perspectives, which suggest that the link between emotion and achievement is simply due to effects of emotions on students' learning and performance. For example, correlations between test anxiety and students' achievement were interpreted as indicating that anxiety impacts achievement, and test anxiety theories put forward various suggestions about mediating mechanisms (e.g., cognitive interference, motivation; Zeidner, 1998, 2014). In a similar vein, in studies on affect and performance more generally, researchers have been interested in the impact of moods and emotions on cognitive performance and created various theories targeting this influence (Clore & Huntsinger, 2009).

Certainly an analysis of the effects of emotions is important as it can document the functional relevance of emotions. However, what about the reverse causal direction, that is, what about the impact of achievement on the development of emotions? In other words, what about emotions as outcomes rather than causes of achievement? Herein we argue that this alternative causal direction is no less important. Beyond their functions, emotions are developmental outcomes that are in and of themselves important, because they are core components of identity, well-being, and health. By implication, researchers and practitioners alike should attend to the antecedents of students' emotions, and academic achievement is certainly one promising candidate---academic successes and failures possibly shape the development of emotions. As such, we concur with traditional perspectives in assuming that emotions impact achievement, but we also extend this notion and expect that achievement reciprocally influences emotion.

Empirical evidence on the causal ordering of students' emotions and their achievement is largely lacking, with a few exceptions pertaining to achievement-related anxiety. Specifically, longitudinal investigations suggested that K-12 students' test anxiety and academic achievement reciprocally influence each other (Meece, Wigfield, & Eccles, 1990; Pekrun, 1992). Furthermore, in a study of mathematics anxiety by Ma and Xu (2004), adolescents' achievement in mathematics had negative effects on their subsequent math anxiety, and anxiety had negative effects on subsequent achievement for two of the five time intervals included. The failure to find effects of anxiety on achievement for the other time intervals was likely due to the high stability of the achievement variable across waves (autogressive ßs > .95). For children's and adolescents' achievement emotions other than anxiety, evidence on reciprocal links with academic achievement is lacking.

In the following sections, we use Pekrun's (2006; Pekrun & Perry, 2014) control-value theory of achievement emotions to derive a theoretical framework for the reciprocal causation of emotion and achievement. This model expands upon previous models on the linkages of anxiety and boredom with achievement (Meece, Wigfield, & Eccles, 1990; Pekrun, 1992; Pekrun, Hall, Goetz, & Perry, 2014; Zeidner, 1998) by addressing not only negative emotions but positive emotions as well. We tested this model using a longitudinal dataset that examined adolescents' emotions and achievement in mathematics over a period of five school years.

A Reciprocal Effects Model of Emotion and Achievement

The control-value theory (Pekrun, 2006; Pekrun & Perry, 2014) integrates propositions from expectancy-value, attributional, and control approaches to achievement emotions (Folkman & Lazarus, 1985; Pekrun, 1992; Turner & Schallert, 2001; Weiner, 1985). Achievement emotions are defined as emotions related to achievement activities and their success and failure outcomes. The theory posits that these emotions are aroused by cognitive appraisals of control over, and the subjective value of, achievement activities and their outcomes. Control appraisals consist of perceptions of one's ability to successfully perform actions (i.e., academic self-concepts and self-efficacy expectations) and to attain outcomes (outcome expectations). Value appraisals pertain to the perceived importance of these activities and outcomes. Furthermore, the theory posits that these emotions, in turn, influence achievement behavior and performance. Since performance outcomes shape succeeding perceptions of control over performance, one important implication is that emotions, their appraisal antecedents, and their performance outcomes are linked by reciprocal causation. In terms of reciprocal causation, the theory is consistent with reciprocal effects models for variables such as students' self-concepts (Marsh & Craven, 2006; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005), achievement goals

(Linnenbrink & Pintrich, 2002), and anxiety (Pekrun, 1992).

Effects of Emotion on Achievement

In the control-value theory, two dimensions describing human affect are used to distinguish types of emotions, namely valence (positive vs. negative or pleasant vs. unpleasant) and activation (activating vs. deactivating). Using these dimensions renders four groups of emotions: positive activating (e.g., enjoyment, hope, pride), positive deactivating (e.g., relaxation, relief), negative activating (e.g., anger, anxiety, shame), and negative deactivating (e.g., boredom, hopelessness). The theory proposes that these emotions influence students' cognitive resources, motivation to learn, and use of learning strategies, thus impacting their achievement (for an in-depth discussion, see Pekrun & Linnenbrink-Garcia, 2012).

Positive activating emotions (e.g., enjoyment of learning) are thought to preserve cognitive resources and focus attention on the learning task, support interest and intrinsic motivation, and facilitate deep learning. Accordingly, these emotions are expected to positively influence students' academic achievement under most task conditions. The opposite pattern of effects is proposed for negative deactivating emotions (boredom, hopelessness). These emotions are thought to reduce cognitive resources and task-related attention, to undermine both intrinsic and extrinsic motivation, and to promote shallow information processing. Accordingly, negative deactivating emotions are expected to negatively influence students' achievement.

Achievement effects are posited to be more variable for the remaining two categories of emotion. Deactivating positive emotions (relaxation, relief) are thought to reduce attention, strategy use, and any immediate motivation to engage with learning tasks, but they can strengthen long-term motivation to reengage with learning. Activating negative emotions (anger, anxiety, shame) are thought to reduce cognitive resources by inducing irrelevant thinking, such as worries about failure in test anxiety, and to undermine intrinsic motivation. On the other hand, these emotions can trigger extrinsic motivation to invest effort to avoid failure. Moreover, they can facilitate the use of more rigid learning strategies, such as rote memorization. However, notwithstanding individual differences regarding effects, we expect that the average overall influence of positive deactivating emotions on achievement is positive, and that the average overall influence of negative activating emotions is negative. For negative activating emotions such as anxiety, this hypothesis is consistent with the available evidence, which indicates that the correlations between these emotions and academic achievement are typically negative (Hembree, 1988; Zeidner, 1998, 2014).

Reverse Effects of Achievement on the Development of Emotion

Achievement reciprocally influences the appraisals that are considered to be proximal antecedents of emotion. As implied by the control-value theory as well as other models of achievement emotion (e.g., Folkman & Lazarus, 1985), positive emotions are thought to be promoted when perceived competence and control over achievement activities are high. For example, students should enjoy learning when they judge themselves competent to master the learning task, provided they are interested in the material. Negative emotions should result when perceived competence and control are low. For example, anxiety about an upcoming important exam should be high if students judge themselves incompetent to pass it. One possible exception is boredom, which could be promoted by high perceived competence if coupled with low task demands (i.e., under-challenge); however, in an academic context, boredom also has been found to be linked to students' lack of perceived competence and control (e.g., Pekrun et al., 2010). Competence and control are thought to influence both students' momentary emotions within a specific situation and their habitual, re-occurring emotions, which are based on re-occurring appraisals and related control-value beliefs (for summaries of empirical evidence, see Daniels & Stupnisky, 2012; Pekrun & Perry, 2014).

Perceived competence and control depend on students' individual achievement history, with success strengthening control and failure undermining it. Hence, achievement is expected to have positive effects on perceived control. Since achievement has positive effects on control, and control has positive effects on positive emotions, it follows that students' achievement should have positive effects on the development of positive emotions. Similarly, since achievement has positive effects on control, and control has negative effects on negative emotions, it follows that achievement should have negative effects on the development of negative emotions.

Feedback Loops of Emotion and Achievement over Time

Because emotions are posited to influence achievement and achievement, in turn, to influence emotion, the two constructs are thought to be linked by reciprocal causation over time. Both effects are expected to be positive for positive emotions, amounting to positive feedback loops, and both effects are expected to be negative for negative emotions, which also amounts to positive feedback loops. We acknowledge that there may be negative feedback loops for negative activating emotions in some students and under some conditions (e.g., failure on an exam instigating a student's anxiety, and anxiety eliciting effort to avoid failing the next exam; Pekrun, 1992). However, the existing evidence summarized above implies that negative activating emotions typically are aroused by failure and contribute to subsequent failure, suggesting that feedback loops should be positive for these emotions as well in the average student.

Overview of the Present Research

We tested the proposed reciprocal effects model using a longitudinal investigation of adolescents' development in mathematics (*Project for the Analysis of Learning and Achievement* *in Mathematics*, PALMA; see Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009; Frenzel, Pekrun, Dicke, & Goetz, 2012; Murayama, Pekrun, Lichtenfeld, & vom Hofe, 2013; Murayama, Pekrun, Suzuki, Marsh, & Lichtenfeld, in press; Pekrun et al., 2007). To test models of reciprocal causal linkages, designs are needed that assess both variables at multiple points in time (Little, Preacher, Selig, & Card, 2007; McArdle, 2009; Rosel & Plewis, 2008). Although such designs cannot fully rule out alternative causal explanations, they are better suited to test causal propositions than cross-sectional designs or longitudinal designs that do not control for prior levels of outcome variables. The PALMA study involved annual assessments of both emotions and achievement, thus making it possible to conduct cross-lagged analyses examining reciprocal causation. This study design made it possible to conduct multiple tests for the effects of emotion on subsequent achievement, and of achievement on subsequent emotion, while controlling for prior emotion and achievement levels.

For the present analysis, we used the grade 5 to 9 data from the PALMA study. As such, the analysis involved five assessments for emotions and five assessments of achievement. These assessments span the time from the beginning of secondary school (grade 5) to the end of compulsory schooling in Germany (grade 9). At the start of secondary school, students are selected into one of three tracks, including lower-track schools (Hauptschule), medium-track schools (Realschule), and higher-track schools (Gymnasium), based on their elementary school achievement. There is no additional school transition until the end of secondary school and students usually remain in the same school. Whereas math teachers and the specific classroom context can change, the broad academic context for students' affective development remains relatively stable across this time period. Specifically, contextual factors defining the emotional salience of achievement, such as the visibility and frequency of feedback on achievement, remain stable during this period. The stability of context does not preclude changes in individual levels of emotion (e.g., due to repeated success or failure and the influence of teachers and peers). However, given the stability of context, we expected relations between students' trait-like emotions considered in this study and their achievement to be stable as well, with effects of these emotions on achievement, and effects of achievement on emotions, showing equivalence (i.e., developmental equilibrium) across each of the one-year intervals included.

Seven distinct mathematics emotions were measured, including math-related enjoyment, pride, anger, anxiety, shame, boredom, and hopelessness. These emotions were selected based on their frequency and theoretical relevance (Pekrun et al., 2007). They were measured as trait-like variables, that is, students' habitual, re-occurring emotions in mathematics. Habitual emotions can influence learning and achievement over a longer time span, in contrast to momentary emotional episodes. In addition, we considered summary constructs of positive and negative affect derived from integrating scores for positive and negative emotions, respectively. As compared with multiple discrete emotions, these constructs render a more parsimonious description of students' affective development (Linnenbrink, 2007).

Achievement was assessed by students' end-of-the-year grades in mathematics, which are derived from multiple evaluations across the school year and represent students' cumulative performance. As such, these grades are suited to examining the impact of emotions on the long-term development of achievement. In addition, test scores from the PALMA mathematical achievement test (see Pekrun et al., 2007) were included to examine the generalizability of the findings across different achievement outcomes. These scores reflect generic mathematical competencies whereas grades represent students' curriculum-related achievement in the classroom, which should be more closely related to their emotions. Accordingly, we expected effects to be stronger for grades than for the test scores.

Structural equation modeling was used to test the reciprocal effects model. To ensure that any observed relations were not mere artifacts of other plausible variables, we controlled for students' gender, intelligence, and family socio-economic status (SES) in the analysis. In addition, we examined the equivalence of relations across school tracks. We expected the effects linking emotion and achievement to be consistent over time and school tracks but modest in size due to controlling for autoregressive effects, intelligence, and demographic variables.

Method

Participants and Design

The sample consisted of German adolescents who participated in the PALMA longitudinal study (Pekrun et al., 2007). The study included annual assessments from grades 5 to 9 (2002-2006). Sampling and the assessments were conducted by the Data Processing and Research Center (DPC) of the International Association for the Evaluation of Educational Achievement (IEA). Samples were drawn from schools within the state of Bavaria and were representative of the student population of this state in terms of student characteristics such as gender, urban versus rural location, and family background (SES; for details, see Pekrun et al., 2007). At each grade level, the students answered the questionnaire towards the end of the school year. All instruments were administered in the students' classrooms by trained external test administrators.

At the first assessment (grade 5), the sample included 2,070 students from 42 schools (49.6% female, mean age = 11.7 years). The sample comprised students from all three school types within the Bavarian public secondary school system as described earlier, including lower-track schools (Hauptschule, 37.2%), intermediate-track schools (Realschule, 27.1%), and higher-track schools (Gymnasium, 35.7%). These three school types differ in average student achievement due to the selection of students by entry-level achievement (see Murayama et al., 2013). The distribution of students across tracks represents the distribution in the population. In each subsequent year, the study not only tracked the students who had participated in the previous assessment(s), but also incorporated those students who had not yet participated in the study but had become members of PALMA classrooms at the time of the assessment (for more details on sampling procedures, see Pekrun et al., 2007). This strategy resulted in the following sample sizes for the subsequent years: 2,059 students in grade 6 (50.0% female, mean age = 12.7 years); 2,397 students at grade 7 (50.1% female, mean age = 13.7 years); 2,410 students at grade 8 (50.5% female, mean age = 14.8 years); 2,528 students at grade 9 (51.1% female, mean age = 15.6 years). Across all five assessments (i.e., grades 5 to 9), a total of 3,425 students (49.7% female) took part in the study. 60.4% of the total sample participated in all five assessments, and 21.7%, 11.7%, 5.1%, and 1.1% completed four, three, two, or one assessment(s), respectively.

Measures

Emotions. Students' emotions in mathematics were measured using the Achievement Emotions Questionnaire-Mathematics (AEQ-M; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). The instructions for the instrument ask respondents to describe how they typically feel when attending class, doing homework, and taking tests and exams in mathematics; in this way, the AEQ-M assesses students' habitual, trait-like math-related emotions. The instrument comprises seven scales measuring mathematics enjoyment (9 items, e.g., "I enjoy my math class"), pride (8 items; e.g., "After a math test, I am proud of myself"), anger (8 items; e.g., "I am annoyed during my math class"), anxiety (15 items; e.g., "I worry if the material is much too difficult for me"), shame (8 items; e.g., "I am ashamed that I cannot answer my math teacher's questions well"), hopelessness (6 items; e.g., "During the math test, I feel hopeless"), and boredom (6 items; e.g., “My math homework bores me to death”). Participants responded on a 1 (*strongly disagree*) to 5 (*strongly agree*) scale, and the scores were summed to form the emotion indexes (Alpha range .86 to .92 across all scales and measurement occasions; see Table 1). The scores were also used to derive indexes for positive and negative affect factors combining positive and negative emotions, respectively (see Data Analysis section).

Achievement. Students’ achievement was assessed by their end-of-the-year grades in mathematics as retrieved from school documents and by standardized test scores.

End-of-the-year grades. These grades are summative scores based on multiple exams within each school year; they represent students’ achievement in the math curriculum for the respective year. Grades range from 1 (*excellent*) to 6 (*poor*). Grade scores were reversed prior to the analysis to ease interpretation.

Test scores. The test scores were derived from the PALMA Mathematics Achievement Test (Pekrun et al., 2007) which measures students’ competencies in arithmetics, algebra, and geometry. The test includes different test forms for different grade levels and includes anchor items to allow for the linkage of test forms across assessments. The obtained scores were scaled using one-parameter logistic item-response theory (Rasch scaling; see Murayama et al., 2013).

Background variables. Demographic variables (gender and SES) and intelligence were included as covariates in the analysis. Gender was coded 1=female, 2=male.

Socio-economic status. **SES** was assessed by parent report using the EGP classification (Erikson, Goldthorpe, & Portocarero, 1979), which consists of six ordered categories of parental occupational status. Higher values represent higher SES.

Intelligence. Intelligence was measured at Time 1 (grade 5) using the 25-item nonverbal reasoning subtest of the German adaptation of Thorndike's Cognitive Abilities Test (Kognitiver Fähigkeitstest [KFT 4–12 + R]; Heller & Perleth, 2000).

Strategy of Data Analysis

Structural equation modeling (SEM; *Mplus*, Version 7; Muthén & Muthén, 2012) was used to evaluate the reciprocal effects model. We estimated two sets of models. The first set used grades, and the second set used test scores as the achievement measure. In both sets, eight different models were estimated, including seven separate models for the discrete emotions and one integrative model combining all emotions into two second-order positive and negative affect factors. There was substantial multicollinearity between the emotion variables in the dataset (Table 1). As such, the present analysis combines two strategies to deal with multicollinearity, namely, using single variables (separate discrete emotion models) and combining them by constructing summary variables (integrative affect models). The separate discrete emotion models also served to examine if the links between emotion and achievement were sufficiently similar to combine emotions into the summary positive and negative affect constructs.

All of the models represent a cross-lagged format, with emotion at each assessment influencing subsequent achievement one year later, and achievement at each assessment influencing subsequent emotion one year later (Figure 1). As such, the discrete emotion models include four paths from emotion to achievement and four paths from achievement to emotion. In the affect models, there were eight paths from positive and negative affect to achievement, eight paths from achievement to positive and negative affect, as well as four paths from positive to negative affect and four paths from negative to positive affect (Figure 1). The emotion variables were modeled as latent constructs. The achievement measure and the three background measures (gender, SES, and intelligence) were evaluated as manifest variables. The background variables were included as covariates; for each of these variables, directional paths to all of the emotion variables and to all of the achievement variables were included.

We estimated two versions for all of the 16 models. In the first version, autoregressive coefficients, cross-paths, and factor residual variances were freely estimated. In the second version, all three parameters were constrained to be invariant across time intervals (developmental equilibrium; e.g., the effects of Time n emotion on Time n+1 achievement were constrained to be the same from each wave to the next).

Measurement models for latent variables. The emotion scale items were used as indicators for each of the latent emotion variables. Following recommendations by Pekrun et al. (2011), a correlated uniqueness approach was used by including correlations between residuals for items representing the same setting (attending class, doing homework, and taking tests and exams in mathematics). In addition, correlations between residuals for identical emotion items across measurement occasions were included to control for systematic measurement error.

The latent affect factors were constructed in a two-step procedure. We first conducted separate confirmatory factor analyses for each of the seven emotions across the five assessments and derived emotion factor scores from these analyses (it was not possible to conduct a confirmatory factor analysis with all emotion items across all assessments, i.e., 60 x 5 = 300 items, due to computational limitations). We then used these factor scores to construct one integrative affect measurement model. For this model, factor scores for the positive emotions served as indicators for positive affect, and factor scores for the negative emotions served as indicators for negative affect. As such, the two affect constructs represent second-order factors.

Measurement equivalence across waves and school tracks. Prior to the main SEM analyses, we sought to establish measurement equivalence of the latent emotion and affect constructs over time and schools tracks. For each of the emotion and affect variables, we sequentially evaluated models of configural, metric, scalar, and residual invariance (Meredith, 1993). Configural invariance is defined by equal patterns of factor loadings. Metric invariance additionally requires equal factor loadings, scalar invariance requires equal factor loading and intercepts, and residual invariance requires equal factor loadings, intercepts, and residual variances. To establish equivalence of constructs for analyzing correlations and path coefficients, metric invariance is the minimum needed (Chen, 2007; Steenkamp & Baumgartner, 1998). To compare model fit, we followed recommendations by Chen (2007). Provided adequate sample size, for testing metric invariance, a change of $\geq$ -.010 in CFI, supplemented by a change of $\geq$ .015 in RMSEA or a change of $\geq$ .030 in SRMR would indicate noninvariance; for testing scalar or residual invariance, a change of $\geq$ -.010 in CFI, supplemented by a change of $\geq$ .015 in RMSEA or a change of $\geq$ .010 in SRMR would indicate noninvariance. As recommended, we did not use the χ^2 difference test because it is overly sensitive to sample size (Marsh, Balla, & McDonald, 1988).

Hierarchical data structure, estimator used, and missing values. As students were nested in schools, we corrected for the clustering of the data using the “type=complex” option implemented in Mplus (Muthén & Muthén, 2012). As noted, schools in the German public secondary school system differ in average student achievement due to the between-schools tracking based on achievement, indicating that nestedness within schools needs to be considered. The <type=complex> corrects standard errors for nestedness while preserving use of the covariance matrix from the full sample to calculate parameters.

To estimate the model parameters, the robust maximum likelihood estimator (MLR) was employed which is robust to nonnormality of the observed variables. To make full use of the data from students with missing data, we applied the full information likelihood method (FIML; Enders, 2010). FIML has been found to result in trustworthy, unbiased estimates for missing values even in the case of large numbers of missing values (Enders, 2010) and to be an adequate method to manage missing data in longitudinal studies (Jeličič, Phelps, & Lerner, 2009). To examine the robustness of the analysis, we replicated the cross-lagged analyses for emotion and achievement with the subsample of students who participated in the study from the beginning (N = 2,070). As compared to the models using the full sample, there were no substantial differences in model fit (Δ CFI ≤ .007, Δ RMSEA ≤ .006, and Δ SRMR ≤ .007 for all of the models), and the substantive results were essentially the same (see Supplemental Material, Tables S6 and S7).

Goodness-of-fit indexes to evaluate model fit. We applied both absolute and incremental fit indices to evaluate the fit of the models, including the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root-mean-square-error of approximation (RMSEA), and the standardized-root-mean 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 were interpreted as indicating good fit (Browne & Cudeck, 1993; Hu & Bentler, 1999). We report these fit indexes to make the present analysis comparable with previous research. However, it should be noted that the recommended cutoff values are often not met with datasets derived from more complex studies, suggesting that they should be used with caution (Heene, Hilbert, Draxler, Ziegler, & Bühner, 2011; Marsh, Hau, & Wen, 2004).

Results

Preliminary Analysis

Alpha coefficients for the emotion scales and manifest correlations for the emotions and achievement are outlined in Table 1 (for information about distributions, see Table S1). Correlations between the emotion measures indicated that enjoyment and pride were positively related, as were anger, anxiety, shame, hopelessness, and boredom. The correlations between positive and negative emotions were negative. Overall, this pattern of relations is consistent with previous evidence on the structures of students' academic emotions (e.g., Pekrun et al., 2011). Enjoyment and pride correlated positively with mathematics achievement in each year, whereas anger, anxiety, shame, hopelessness, and boredom correlated negatively with achievement.

Confirmatory Factor Analysis (CFA) for the Emotion Constructs

To further examine the relations between emotions, item-based CFA models including the seven emotions were estimated. This was done separately for the five measurement occasions. The models showed a good fit to the data (Supplemental Material, Table S2), supporting the measurement quality of the emotion variables. The latent correlations between the emotion variables showed the same pattern as the manifest correlations (Table 1). These correlations are corrected for measurement error and indicate that the latent emotion variables are closely related but nevertheless distinct (for similar findings with university students, see Pekrun et al., 2011). This is also true for emotions that might be presumed to constitute opposite ends of a bipolar continuum, such as enjoyment and boredom, which showed moderately negative relationships. The strongest correlations were found for neighboring, like-valenced emotions such as enjoyment and pride, and anxiety, shame, and hopelessness. In interpreting these correlations, it is important to note that the present study used the AEQ-M to assess students' trait-like emotions. As noted by Pekrun et al. (2011), like-valenced trait emotions are known to be strongly correlated, in contrast to state emotions which show more divergence.

For positive and negative affect based on the emotion factor scores, we conducted an integrative CFA including both constructs across all five measurement occasions. The fit for this CFA model was good (Supplemental Material, Table S3, configural invariance model). Latent correlations between the positive and negative affect factors were r = -.19, -.23, -.25, -.23, and -.21 (all ps < .01) for Time 1, 2, 3, 4, and 5, respectively, showing that the two affect constructs were sufficiently distinct.

Measurement Invariance of the Emotion Constructs over Time and School Tracks

Measurement invariance across waves was tested separately for the seven emotions and for positive and negative affect. The configural invariance models showed a good fit to the data, with CFI > .93, RMSEA < .03, and SRMR < .05 for all seven discrete emotion constructs (Supplemental Material, Table S3). As compared with these models, the loss of fit for the metric invariance models was Δ CFI ≤ -.004, Δ RMSEA ≤ .001, and Δ SRMR ≤ .006 for all models, indicating clear support for metric invariance for all of the emotions. The loss of fit for the scalar invariance models was Δ CFI ≤ -.007, Δ RMSEA ≤ .004, and Δ SRMR ≤ .007 for all of the emotions, documenting that scalar invariance was supported as well. The loss of fit for the residual invariance models was Δ CFI < -.010 for all emotions except shame, Δ CFI = -.010, as well as Δ RMSEA ≤ .003 and Δ SRMR ≤ .008 for all emotions, indicating support for residual invariance. For positive and negative affect, the loss of fit was Δ CFI ≤ .008, Δ RMSEA ≤ .004, and Δ SRMR ≤ .005 for the metric, intercept, and residual invariance models, demonstrating support for invariance for these second-order constructs as well. In sum, the findings show that the latent emotion and affect variables showed strong measurement equivalence over time, thus meeting the requirements to be included in longitudinal analysis. Furthermore, in supplemental analyses using multi-group analysis, the emotion constructs also showed strong measurement equivalence across the three school tracks (see Supporting Information, Table S8).

Reciprocal Effects Models of Emotions and Achievement

The fit indexes provided support for the cross-lagged structural equation models for all seven emotions as well as positive and negative affect and across both measures of achievement. For all of the models freely estimating autoregressive effects, cross-lagged effects, and factor residual invariances, CFI was > .92, TLI > .90, RSMEA < .06, and SRMR < .08 (Table 2 and Supplemental Material, Table S4). When constraining autoregressive effects, cross-lagged effects, and factor residual variances to be equal across time intervals, the loss of fit was Δ CFI < .003, Δ RMSEA < .001, and Δ SRMR ≤ .003 for all of the models. These findings support the invariance of these parameters, suggesting developmental equilibrium in autoregressive stability and in the links of emotion and achievement across time. Accordingly, we adopted the constrained models for further interpretation, which have the additional advantage of providing more robust and precise parameter estimates (note that these constraints equalize unstandardized coefficients; to ease interpretation, we report standardized coefficients which can still differ due to the standardization procedure).

Emotions and grades. Factor loadings, path coefficients, and residual variances for the reciprocal effects models including grades are displayed in Table 3. In the enjoyment and pride models, both the emotion variables and students' achievement showed considerable stability over time, as indicated by the autoregressive effects for these variables. Furthermore, there were significant relations between the positive emotions and achievement at grade 5 in these models, latent *r*s = .26 and .26, *p*s < .001, for enjoyment and pride, respectively. Over and above these pre-existing relations, and despite autoregressive stability, results showed enjoyment and pride to positively predict each subsequent achievement outcome (ß range .11 to .13, *p*s < .001) while controlling for gender, SES, and intelligence. In addition, positive paths emerged from each achievement outcome to the subsequent enjoyment and pride variables (all βs = .11, *p*s < .001).

In the negative emotion models, there were substantial initial links between anger, anxiety, shame, boredom, and hopelessness at grade 5, latent *r*s = -.31, -.39, -.32, -.16, and -.37, respectively, *p*s < .001. Despite these links and the considerable stability of the emotion and achievement variables over time, anger, anxiety, shame, boredom, and hopelessness negatively predicted each subsequent achievement outcome (ß range -.08 to -.14, all *p*s < .001) while controlling for gender, SES, and intelligence. The effects were especially pronounced for anxiety and hopelessness (all ßs > -.11). In addition, negative paths from each achievement outcome to subsequent anger, anxiety, shame, boredom, and hopelessness were observed (ß range -.06 to -.14; all *p*s < .001).

These effects were similar across the two positive emotions, and similar across the five negative emotions, thus justifying their combination into positive and negative affect constructs. In the reciprocal effects model for positive and negative affect, the initial links with achievement were *r*s = .26 and -.33 for positive and negative affect, respectively, *p*s < .001. Despite these links and strong autoregressive coefficients for both positive and negative affect as well as achievement, positive affect positively predicted achievement, and negative affect negatively predicted achievement. Because both types of affect were included in the analysis, these findings indicate that positive and negative affect had independent predictive effects on achievement. Achievement, in turn, had positive predictive effects on positive affect and negative predictive effects on negative affect. Regarding cross-paths between positive and negative affect, we had not expected any effects of this type and none of the paths were significant.

Emotions and test scores. The findings for emotions and test scores replicated the results for grades, demonstrating generalizability across different achievement measures (Supplemental Material, Table S5). As expected, however, the effects were weaker than for grades. Positive emotions were positive predictors of test scores, ß range = .04 to .05, and negative emotions were negative predictors, ß range = -.03 to -.08, all *p*s < .001. Test scores were a positive predictor of positive emotions, ß range = .05 to .07, and a negative predictor of negative emotions, ß range = -.04 to -.11, all *p*s < .001. In the positive and negative affect model, positive affect was not a significant predictor of test scores (all ßs = .01, *ns*), whereas negative affect predicted test scores, ß range = -.06 to -.07, *p*s < .001. Test scores, in turn, were a positive predictor of positive affect, ßs = .03, *p*s < .01, and a negative predictor of negative affect, ß range = -.04 to -.05, *p*s < .001.

Effects of the covariates. Intelligence had positive effects on grades and test scores as well as negative effects on students' anger, anxiety, shame, and hopelessness (Tables 3 and S5). SES also had positive, albeit weaker, effects on math achievement. Gender had significant effects on all of the emotions except anger, indicating that girls reported lower enjoyment, pride, and boredom, and higher anxiety, shame, and hopelessness in mathematics than boys.

Equivalence of effects across school tracks. In supplemental analyses, we used multi-group analysis to examine the equivalence of cross-paths, autoregressive effects, and effects of covariates across the three school tracks. Comparing models constraining versus not constraining these coefficients to be invariant (using Chen's, 2007, criteria outlined in the Data Analysis section), the findings provide robust support for invariance across tracks for all of the emotion and affect constructs included and both math grades and test scores (see Tables S9, S10).

Discussion

The findings of this study provide robust evidence for the proposed reciprocal effects model of emotion and achievement. As indicated by longitudinal SEM, adolescents' math-related positive emotions (enjoyment and pride) positively predicted their subsequent end-of-the-year math grades, and grades, in turn, positively predicted the development of positive emotions. Math-related negative emotions (anger, anxiety, shame, hopelessness, and boredom) were negative predictors of subsequent math grades, and grades, in turn, were a negative predictor for the development of negative emotions. Similar predictive effects were obtained for the integrative constructs of positive and negative affect, respectively, and for test scores as a measure of achievement. The findings were consistent across models for the seven discrete emotions, the combined positive and negative affect model, four time intervals, two different measures of achievement (grades, test scores), and the three school tracks while controlling for students' gender, SES, and intelligence. All of the effects were significant with the single exception of the effects of positive affect on test scores.

Because prior links between emotion and achievement as well as intelligence and demographic background variables were controlled, the path coefficients are likely to represent effects of emotion on achievement, and vice versa, rather than simply the influence of prior emotion, prior achievement, gender, intelligence, or socio-economic status. As expected, the size of these coefficients was modest. However, it is important to note that the coefficients represent incremental predictive effects due to prior emotion and achievement being controlled. Thus, the coefficients represent effects of each variable on change in the other from one assessment to the next, rather than effects on the absolute levels of these variables. Furthermore, both emotion and achievement showed considerable stability over time, leaving little variance to be explained and making it difficult to detect the effects of additional variables. From this perspective, the consistency of effects lends credibility to the notion that emotion and achievement are indeed linked by reciprocal causation over time.

Reciprocal Effects Linking Emotion and Achievement

The findings are congruent with previous evidence showing that emotions and academic achievement are correlated (Goetz & Hall, 2013; Pekrun & Linnenbrink-Garcia, 2014; Zeidner, 1998). However, they go beyond correlational evidence by disentangling the directional effects underlying the emotion-achievement link. Specifically, the findings suggest that emotions indeed have an influence on adolescents' achievement, over and above the effects of general cognitive ability and prior accomplishments. These effects are in line with Pekrun's (2006) control-value theory which posits that emotions influence learning and achievement outcomes.

Of specific importance is the finding that adolescents' positive emotions in mathematics had positive predictive effects on their math grades over time. Previous research has produced mixed findings on the relation between students' positive affect and their learning, with most studies reporting positive relations (see Linnenbrink, 2007) but some others null findings (e.g., Pekrun, Elliot, & Maier, 2009). The present analysis suggests that positive emotions can have positive effects, in line with theory and the views of educational practitioners. However, the effects were weaker for positive emotion than for the negative emotion constructs, and did not reach significance for the predictive effect of positive affect on test scores. Future research should examine possible reasons why negative emotion is a stronger predictor of students' academic achievement than positive emotion. This difference may relate to general asymmetries in the impact of negative versus positive states and events on human memory and action (see e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001).

The results also contribute to our understanding of the developmental origins of students' emotions. The findings suggest that achievement impacts the development of emotions. More specifically, it appears that doing well in school can strengthen students' positive emotions and reduce their negative emotions over time, whereas doing poorly in school undermines positive emotions and exacerbates negative emotions. These effects are likely mediated by students' perceptions of competence and control over achievement, with high control promoting enjoyment and pride and low control leading to negative emotions (e.g., Pekrun et al., 2010).

Taken together, these effects amount to positive developmental feedback loops linking emotions and achievement. As noted, a few longitudinal studies have found that students' test anxiety and their achievement were linked by positive feedback loops (Meece, Wigfield, & Eccles, 1990; Pekrun, 1992). The present research adds to this literature by showing that emotions other than anxiety share similar links with achievement. As such, it would appear that unidirectional models are unable to adequately capture the complex reality of students' emotions. Rather, systems-oriented perspectives are needed that take more complex patterns of causal links into account, including feedback loops between emotions, their antecedents, and their effects.

Discrete Emotions versus General Affect

It is noteworthy that the cross-paths were similar across different discrete emotions. For effects of achievement on emotion, this is to be expected, as success and failure are thought to impact the development of different positive and negative emotions in similar ways. As outlined in our reciprocal effects model, success is expected to generally increase perceived control, thus enhancing positive emotions, and failure is expected to decrease control, leading to negative emotions. However, regarding effects of emotion on achievement, emotion theories such as the control-value theory (Pekrun, 2006) imply that the effects of some emotions (e.g., deactivating negative emotions such as boredom) may be more consistent than the effects of other emotions (e.g., activating negative emotions such as anxiety). Instead, the findings clearly indicate that the predictive effects of emotions on students' long-term achievement were also similar across different emotions. Accordingly, whereas constructs of discrete emotions are needed to explain the impact of emotions on functional mechanisms and different types of cognitive performance, parsimonious summary constructs of positive and negative affect may be sufficient to explain their relations with overall academic achievement. This possibility is underscored by the robust findings for positive and negative affect documented in the present analysis.

Effects of Gender, Intelligence, and SES

The findings on gender differences are consistent with previous evidence showing that girls report less enjoyment and more anxiety and shame in mathematics even if they perform as well as boys. Lower competence beliefs and perceived values in mathematics may be possible explanations (Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013). However, girls reported less boredom than boys, in line with previous evidence (Pekrun et al., 2010). As such, the findings suggest that girls exhibit a more maladaptive profile of math emotions, except for boredom.

As expected, intelligence had substantial predictive effects on the achievement variables. Furthermore, intelligence had negative effects on math-related anger, anxiety, shame, and hopelessness. Given that students' mathematics achievement was included in the analysis, this finding suggests that higher general cognitive ability can help to reduce negative mathematics emotions, above and beyond any effects of students' academic success in mathematics. Finally, SES also had positive, albeit weaker, effects on math achievement, suggesting that the family exerts an influence on students' achievement, over and above any effects of cognitive ability.

Limitations, Suggestions for Future Research, and Implications for Practice

The present study represents a significant advancement over previous research, because it documents reciprocal effects of emotion and achievement over time while controlling for general cognitive ability and critical demographic background variables. Nevertheless, several limitations should be considered when interpreting the study findings and can be used to suggest directions for future research.

Methodological considerations. As compared with experimental studies, the power of non-experimental field studies to derive causal conclusions is limited. As such, although the present analysis used multi-wave longitudinal structural equation modeling and controlled for related variables and autoregressive effects, the possibility still exists that our findings are attributable to other variables that were not included in the study. On the other hand, field studies may be more ecologically valid than experimental emotion studies, which are limited in terms of situational representativeness and ethical concerns about experimentally manipulating emotions. Furthermore, statistical power is higher in field studies such as the present one due to large sample size. To balance the benefits and drawbacks of different methodologies and make headway in this avenue of research, future studies should further pursue the approach taken herein while complementing this approach with experimental studies.

Achievement was assessed by students' end-of-year grades and test scores. By using grades, we sought to employ an ecologically valid measure of student achievement (for a similar procedure, see Pekrun et al., 2014). As is typical for grades, more detailed information about reliability was not available; as such, it was not possible to disattenuate the link between emotions and grades for potential unreliability of this achievement measure. However, in German secondary schools, end-of-the-year grades are summative scores based on multiple exams within each school year, which may boost their reliability in comparison to grades on single exams. In the present study, the stability of grades across years (all ßs > .50) could be considered as a lower bound to reliability. Furthermore, from the perspective of grades as sources of students' emotional development, they could be seen as having almost perfect reliability---grades, rather than objective achievement, provide the feedback that shapes students' perceptions of success and failure and any development based on these perceptions. In addition, an advantage of grades is that they represent achievement in terms of the math curriculum taught in students' classes. They represent the specific contents learned by students and may be superior to alternative measures in terms of curricular validity. Finally, the findings based on grades proved to be generalizable, as the results were essentially the same for test scores.

Substantive issues. The present research examined academic emotions as experienced by adolescents in the domain of mathematics. It is open to question whether the present findings would generalize to other age groups, such as elementary school children or post-secondary students. Furthermore, it is possible that there is individual variation in the link between emotions and achievement. To examine such variation, within-person analyses of the relations between emotion and achievement over time are needed (e.g., by using experience sampling methodology; Goetz, Sticca, Pekrun, Murayama, & Elliot, 2016). Because the present research involved samples of German adolescents, it also remains an open question as to whether the findings would generalize to students in other cultures. Additionally, future research should explore if these findings generalize to emotions in achievement domains other than mathematics,

The study considered a broad range of important mathematics emotions but did not include an exhaustive list of emotions. It is open to question whether the observed reciprocal effects would also occur for emotions not assessed herein. Specifically, the study did not include students' deactivating positive emotions, such as relief and relaxation. Future studies could explore how these emotions are linked to students' academic achievement. Furthermore, the present study examined students' trait-like emotions which are known to be highly correlated (Pekrun et al., 2011), which makes it difficult to determine unique variance in achievement attributable to different emotions. Future research should examine the unique impact of multiple state emotions, which are less correlated (Goetz et al., 2016), on students' learning.

Finally, the study addressed the overall developmental relations between emotion and achievement but did not examine the mechanisms that mediate the observed links. In the proposed model of reciprocal effects, it is posited that effects of emotion on achievement are due to the influence of emotions on cognitive resources, motivation, and strategy use. The effects of achievement outcomes on the development of emotion are thought to be mediated by perceptions of competence and control over performance, and could additionally be mediated by value appraisals. More research on the link between emotion and achievement as mediated by these cognitive and motivational mechanisms is needed to better understand students' emotions and their relations with important school outcomes.

Implications for educational practice. Two important messages follow from the present research. First, the results suggest that emotions have effects on adolescent students' academic achievement, and that these effects are not merely an epiphenomenon of prior performance---more likely, they represent a true causal influence of students' emotion experiences. By implication, the findings suggest that educators, administrators, and parents alike should consider intensifying efforts that strengthen adolescents' positive emotions and minimize their negative emotions. Second, the results imply that achievement outcomes reciprocally influence students' emotions, suggesting that successful performance attainment and positive achievement feedback can facilitate the development of positive emotions, and failure experiences can contribute to the development of negative emotions. Accordingly, providing students with opportunities to experience success (e.g., using intrapersonal standards to evaluate achievement; emphasizing mastery over competition goals) may help to promote positive emotions and prevent negative emotions (also see Pekrun, Cusack, Murayama, Elliot, & Thomas, 2014). By documenting the influence of achievement outcomes on students' emotions, the present findings elucidate one important factor that can be targeted by educators to reduce students' negative affect and facilitate the development of emotional well-being.

References

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323-370. Doi: 10.1037//1089-2680.5.4.323

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. 136-161). Thousand Oaks, CA: Sage.

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

Clore, G. L., & Huntsinger, J. R. (2009). How the object of affect guides its impact. *Emotion Review, 1,* 39-54. doi: 10.1177/1754073908097185

Daniels, L. M., & Stupnisky, R. H. (2012). Not that different in theory: Discussing the control-value theory of emotions in online learning environments. *Internet and Higher Education, 15,* 222-226. doi:10.1016/j.iheduc.2012.04.002

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

Erikson, R., Goldthorpe, J. H., & Portocarero, L. (1979). Intergenerational class mobility in three Western European Societies: England, France, and Sweden. *British Journal of Sociology, 30,* 341-415.

Folkman, S., & Lazarus, R. S. (1985). If it changes it must be a process: Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology, 48,* 150-170.

Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal* *of Educational Psychology, 101,* 705-716. doi: 10.1037/a0014695

Frenzel, A. C., Pekrun, R., Dicke, A. L., & Goetz, T. (2012). Beyond quantitative decline: Conceptual shifts in adolescents' development of interest in mathematics. *Developmental Psychology, 48,* 1069-1082. doi: 10.1037/a0026895

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science, 24,* 2079-2087. doi: 10.1177/0956797613486989

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

Goetz, T., Sticca, F., Pekrun, R., Murayama, K., & Elliot, A. J. (2016). Intraindividual relations between achievement goals and discrete achievement emotions: An experience sampling approach. *Learning and Instruction. 41,* 115-125. doi: 10.1177/095679761348698

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

Heller, K., & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4–12 + R) [Cognitive Abilities Test for Grades 4 to 12, revision (KFT 4–12 + R)]. Göttingen, Germany: Hogrefe.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58,* 47-77. doi: 10.3102/00346543058001047

Hu, L. -T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424-453. doi: 10.1037/1082-989X.3.4.424

Jeličič, 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*, 1195-1199.

Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development, 31,* 357-365. doi: 10.1177/016502540707757

Linnenbrink, E. A. (2007). The role of affect in student learning: A multi-dimensional approach to considering the interaction of affect, motivation, and engagement. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 107-124).

Linnenbrink, E. A., & Pintrich, P. R. (2002). Achievement goal theory and affect: An asymmetrical bidirectional model. *Educational Psychologist, 37,* 69-78. doi: 10.1207/S15326985EP3702_2

Ma, X., & Xu, J. 2004. The causal ordering of mathematics anxiety and mathematics achievement: a longitudinal panel analysis. *Journal of Adolescence, 27,* 165–179. doi: 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,* 391– 410. doi:10.1037/0033-2909.103.3.391

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fix indexes and dangers in overgeneralizing

Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11,* 320-341. doi: 10.1207/s15328007sem1103

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective. *Perspectives on Psychological Science, 1,* 133-163. doi: 10.1111/j.1745-6916.2006.00010.x

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,* 397-416. doi: 10.1111/j.1467-8624.2005.00853.x

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60,* 577-605. doi: 10.1146/annurev.psych.60.110707.153612

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,* 60-70. doi: 10.1037/0022-0663.82.1.60

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525-543

Murayama, K., Pekrun, R., Lichtenfeld, S., & vom Hofe, R. (2013). Predicting long-term growth in students' mathematics achievement: The unique contributions of motivation and cognitive strategies. *Child Development, 84,* 1475-1490. doi: 10.1111/cdev.12036

Murayama, K., Pekrun, R., Suzuki, M., Marsh, H. W., & Lichtenfeld, S. (in press). Don't aim too high for your kids: Parental over-aspiration undermines students' learning in mathematics. *Journal of Personality and Social Psychology.*

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide.* Los Angeles, CA: Author.

Pekrun, R. (1992). The expectancy-value theory of anxiety: Overview and implications. In D.G. Forgays, T. Sosnowski, & K. Wrzesniewski (Eds.), *Anxiety: Recent developments in self-appraisal, psychophysiological and health research* (pp. 23-41). Washington, DC: Hemisphere.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18,* 315-341. doi: 10.1007/s10648-006-9029-9

Pekrun, R., Cusack, A., Murayama, K., Elliot, A. J., & Thomas, K. (2014). The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction, 29,* 115-124. doi: 10.1016/j.learninstruc.2013.09.002

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology, 101,* 115-135. doi: 10.1037/a0013383

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,* 531-549. doi: 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,* 36-48. doi: 10.1016/j.cedpsych.2010.10.002

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*, 696-710.

Pekrun, R., vom Hofe, R., Blum, W., Frenzel, A. C., Goetz, T. & Wartha, S. (2007). Development of mathematical competencies in adolescence: The PALMA longitudinal study. In M. Prenzel (Ed.), *Studies on the educational quality of schools* (pp. 17-37). Münster, Germany: Waxmann.

Pekrun, R., & Linnenbink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 259-282). New York: Springer.

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

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *Handbook of emotions in education* (pp. 120-141). New York: Taylor & Francis.

Rosel, J., & Plewis, I., (2008). Longitudinal data analysis with structural equations. *Methodology, 4,* 37-50. doi: 10.1027/1614-2241.4.1.37

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

Turner, J. E., Schallert, D. L. (2001). Expectancy-value relationships of shame reactions and shame resiliency. *Journal of Educational Psychology, 93,* 320-329. doi: 10.1037//0022-0663.93.2.320

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92,* 548-573. doi: 10.1037/0033-295X.92.4.548

Zeidner, M. (1998). *Test anxiety. The state of the art.* New York: Plenum.

Zeidner, M. (2014). Anxiety in education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *Handbook of emotions in education* (pp. 265-288). New York: Taylor & Francis.

Table 1

Alpha Coefficients and Pearson Product-Moment Correlations for Emotions and Achievement

	Enjoyment	Pride	Anger	Anxiety	Shame	Boredom	Hopelessness
Enjoyment	(.87) [a]	.83	-.63	-.53	-.36	-.60	-.48
	(.87)	.84	-.65	-.51	-.33	-.63	-.51
	(.88)	.86	-.65	-.48	-.30	-.62	-.49
	(.85)	.86	-.61	-.46	-.30	-.57	-.49
	(.89)	.88	-.56	-.42	-.23	-.50	-.46
Pride	.73	(.87)	-.42	-.37	-.25	-.39	-.38
	.74	(.88)	-.51	-.42	-.27	-.50	-.44
	.75	(.88)	-.50	-.40	-.26	-.47	-.43
	.76	(.89)	-.48	-.37	-.25	-.47	-.43
	.78	(.89)	-.46	-.35	-.18	-.43	-.39
Anger	-.55	-.35	(.87)	.88	.76	.84	.93
	-.55	-.40	(.88)	.86	.73	.82	.82
	-.56	-.39	(.87)	.86	.69	.79	.83
	-.53	-.39	(.87)	.86	.68	.72	.85
	-.49	-.37	(.88)	.87	.68	.75	.84
Anxiety	-.41	-.29	.74	(.90)	.92	.67	.90
	-.39	-.31	.74	(.90)	.92	.60	.91
	-.35	-.29	.74	(.91)	.87	.53	.92
	-.33	-.26	.73	(.91)	.88	.51	.92
	-.32	-.26	.73	(.92)	.87	.55	.91
Shame	-.27	-.19	.65	.78	(.86)	.55	.82
	-.23	-.18	.62	.77	(.88)	.48	.79
	-.20	-.16	.58	.74	(.87)	.37	.78
	-.19	-.16	.57	.75	(.87)	.36	.78
	-.14	-.09	.58	.74	(.89)	.42	.78
Boredom	-.51	-.27	.70	.44	.37	(.86)	.63
	-.53	-.35	.70	.39	.31	(.89)	.60
	-.52	-.33	.66	.33	.25	(.90)	.54
	-.48	-.32	.61	.29	.23	(.90)	.56
	-.41	-.29	.64	.32	.28	(.90)	.57
Hopelessness	-.41	-.34	.72	.83	.74	.43	(.86)
	-.43	-.38	.74	.86	.73	.42	(.88)
	-.42	-.37	.74	.86	.71	.37	(.88)
	-.43	-.37	.75	.86	.70	.37	(.87)
	-.43	-.37	.76	.86	.68	.38	(.83)
Achievement	.20	.18	-.30	-.37	-.33	-.37	-.12
(end-of-year	.25	.22	-.30	-.38	-.34	-.37	-.09
grades)	.34	.29	-.34	-.37	-.29	-.39	-.15
	.41	.36	-.36	-.37	-.29	-.39	-.15
	.45	.38	-.42	-.40	-.29	-.45	-.22

Note. [a] 1st, 2nd, 3rd, 4th, 5th coefficient in each column: Grade 5, 6, 7, 8, and 9, respectively. Coefficients below main diagonal are manifest correlations. Coefficients above main diagonal are latent correlations based on confirmatory factor analyses for each wave. Coefficients in parentheses are Cronbach's Alphas. $p < .01$ for all coefficients.

Table 2

Reciprocal Effects Models for Emotion and Grades: Fit Indexes

	χ^2	*df*	CFI	TLI	RMSEA	SRMR
Model	*Cross-paths, autoregressive effects, and residual variances freely estimated*					
Enjoyment	4125.280**	1147	.940	.928	.027	.052
Pride	2729.201**	722	.940	.928	.028	.048
Anger	3238.875**	918	.941	.927	.027	.049
Anxiety	9091.434**	2992	.920	.909	.024	.050
Shame	2168.850**	907	.965	.957	.020	.044
Boredom	1384.409**	532	.974	.966	.021	.038
Hopelessness	2018.158**	562	.959	.949	.027	.055
Positive and negative affect	6837.618**	685	.947	.930	.051	.075
	Cross-paths, autoregressive effects, and residual variances invariant across waves					
Enjoyment	4210.435**	1165	.938	.927	.027	.053
Pride	2794.131**	740	.942	.930	.028	.049
Anger	3285.829**	936	.940	.928	.027	.050
Anxiety	9148.887**	3010	.920	.909	.024	.050
Shame	2244.200**	925	.964	.956	.020	.045
Boredom	1500.094**	550	.971	.963	.022	.041
Hopelessness	2058.064**	580	.959	.950	.027	.055
Positive and negative affect	6976.520**	721	.946	.933	.050	.078

** $p < .01$.

Table 3

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

	Enjoyment model		Pride model		Anger model		Anxiety model		Shame model	
	Enjoyment	Grades	Pride	Grades	Anger	Grades	Anxiety	Grades	Shame	Grades
Factor loadings	.37-.81[a]		.55-.77[a]		.58-.77[a]		.44-.77[a]		.48-.78[a]	
Autoregressive effects										
T1 → T2	.67***	.57***	.62***	.57***	.58***	.57***	.60***	.56***	.62***	.58***
T2 → T3	.66***	.59***	.64***	.59***	.61***	.59***	.64***	.58***	.61***	.60***
T3 → T4	.66***	.61***	.65***	.61***	.62***	.60***	.66***	.60***	.60***	.62***
T4 → T5	.65***	.59***	.65***	.59***	.62***	.58***	.68***	.58***	.60***	.60***
Cross-lagged effects	*Grades → Enjoyment*	*Enjoyment → Grades*	*Grades → Pride*	*Pride → Grades*	*Anger → Grades*	*Grades → Anger*	*Grades → Anxiety*	*Anxiety → Grades*	*Grades → Shame*	*Shame → Grades*
T1 → T2	.11***	.13***	.11***	.11***	-.12***	-.10***	-.08***	-.11***	-.06***	-.09***
T2 → T3	.11***	.13***	.11***	.12***	-.13***	-.10***	-.08***	-.13***	-.06***	-.09***
T3 → T4	.11***	.13***	.11***	.12***	-.14***	-.10***	-.07***	-.14***	-.06***	-.09***
T4 → T5	.11***	.12***	.11***	.12***	-.13***	-.10***	-.07***	-.14***	-.06***	-.08***
Effects of Covariates at T1										
Gender	.14***	.02	.17***	.02	-.03	.02	-.16***	.02	-.09**	.02
SES	-.05**	.09***	.05*	.09***	.03	.09***	-.04	.09***	-.03	-.09***
Intelligence	-.02	.40***	-.00	.40***	-.12***	.40***	-.18***	.40***	-.17***	.40***
Residual Variances										
T1	.98	.82	.97	.82	.98	.82	.94	.82	.96	.82
T2	.50	.57	.57	.58	.62	.57	.59	.57	.55	.58
T3	.51	.56	.54	.56	.59	.56	.53	.56	.58	.56
T4	.52	.58	.53	.58	.57	.57	.50	.58	.60	.58
T5	.52	.56	.52	.56	.57	.55	.50	.56	.61	.56

Table 3 (continued)

	Boredom model		Hopelessness model		Positive and negative affect model			
	Boredom	Grades	Hopelessn.	Grades	Pos. affect [b]	Neg. affect [b]	Grades	
Factor loadings	.56-.77 [a]		.63-.85 [a]		.77-.96 [a]	.41-.93 [a]		
Autoregressive effects								
T1 → T2	.63***	.59***	.53***	.56***	.80***	.74***	.54***	
T2 → T3	.65***	.61***	.57***	.59***	.81***	.76***	.56***	
T3 → T4	.66***	.63***	.58***	.60***	.82***	.78***	.57***	
T4 → T5	.66***	.61***	.59***	.58***	.82***	.79***	.56***	
Cross-lagged effects	*Grades → Boredom*	*Boredom → Grades*	*Grades → Hopelessn.*	*Hopelessn. → Grades*	*Grades → Pos. affect*	*Grades → Neg. affect*	*Pos. affect → Grades*	*Neg. affect → Grades*
T1 → T2	-.06***	-.08***	-.11***	-.11***	.05***	-.04***	.10***	-.08***
T2 → T3	-.06***	-.08***	-.12***	-.12***	.05***	-.04***	.10***	-.08***
T3 → T4	-.06***	-.09***	-.12***	-.13***	.05***	-.04***	.10***	-.09***
T4 → T5	-.06***	-.09***	-.11***	-.13***	.05***	-.04***	.10***	-.09***
Effects of Covariates at T1								
Gender	.09**	.02	-.16***	.02	.15***	-.13***	.02	
SES	-.03	.09***	-.04	.09***	-.05**	-.03	.09***	
Intelligence	.00	.40***	-.13***	.40***	-.02	-.15***	.40***	
Residual Variances								
T1	.99	.82	.95	.82	.97	.96	.82	
T2	.59	.58	.66	.58	.34	.41	.58	
T3	.56	.56	.61	.58	.33	.36	.57	
T4	.54	.57	.60	.56	.32	.35	.59	
T5	.53	.55	.59	.56	.32	.33	.57	

Note. [a] Range of factor loadings. $p < .001$ for all loadings. [b] Cross-paths between positive and negative affect were not significant (all $ps > .05$).

* $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 1. Basic structure of cross-lagged reciprocal effects models. Upper part: emotion and achievement. Lower part: positive affect, negative affect, and achievement. The models include cross-lagged effects, autoregressive effects, and directional paths from the covariates to emotion or affect and achievement at all waves. Correlations between the covariates and between residuals are not displayed.