University of Reading

Mathematics motivation in students with low cognitive ability: a longitudinal study of motivation and relations with effort, self-regulation, and grades

Article

Accepted Version

Tracey, D., Morin, A. J. S., Pekrun, R., Arens, A. K., Murayama, K., Lichtenfeld, S., Frenzel, A. C., Goetz, T. and Maïano, C. (2020) Mathematics motivation in students with low cognitive ability: a longitudinal study of motivation and relations with effort, self-regulation, and grades. American Journal on Intellectual and Developmental Disabilities, 125 (2). pp. 125-147. ISSN 1944-7558 doi: 10.1352/1944-7558-125.2.125 Available at https://centaur.reading.ac.uk/85988/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1352/1944-7558-125.2.125

Publisher: American Association on Intellectual and Developmental Disabilities

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in

the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Running head: MOTIVATION AND LOW COGNITIVE ABILITY

Mathematics Motivation in Students with Low Cognitive Ability: A Longitudinal Study of Motivation and Relations with Effort, Self-Regulation, and Grades

Danielle Tracey [1]†, Alexandre J.S. Morin[2]†, Reinhard Pekrun [3], A. Katrin Arens [4] , Kou Murayama [5], Stephanie Lichtenfeld [6], Anne C. Frenzel [7], Thomas Goetz [8], Christophe Maïano [9]

[1] School of Education, Western Sydney University, Sydney, Australia

[2] Department of Psychology, Concordia University, Canada

[3] Department of Psychology, University of Munich, Germany and Institute for Positive Psychology and Education, Australian Catholic University, Sydney, Australia

[4] German Institute for International Educational Research, Frankfurt, Germany

[5] School of Psychology and Clinical Language Sciences, University of Reading, UK

[6] School of Education, University of Durham, Durham, UK and Department of Psychology, University of Munich, Munich, Germany

[7] Department of Psychology, University of Munich, Munich, Germany

[8] Department of Empirical Educational Research, University of Konstanz, Germany and Thurgau University of Teacher Education, Switzerland

[9] Département de psychoéducation et de psychologie, Université du Québec en Outaouais, Saint-Jérome, Canada

† Because the first two authors (D.T., & A.J.S.M.) contributed equally to the preparation of this paper, their order was determined at random: both should be considered first authors.

***Corresponding author:**
Danielle Tracey
School of Education, Western Sydney University
Locked Bag 1797 Penrith New South Wales 2751 Australia
Email: d.tracey@westernsydney.edu.au

Funding: Data collection for this study was funded by four grants from the German Research Foundation (DFG) awarded to the third author (PE 320/11-1, PE 320/11-2, PE 320/11-3, PE 320/11-4). Preparation of this article was also supported by research grants from the Australian Research Council (DP140101559) awarded to the first, second, and last authors, from the Social Sciences and Humanities Research Council of Canada (430-2012-0091, 435-2014-0909) awarded to the first, second, and last authors, and from a LMU Research Chair grant from the University of Munich awarded to the third author.

This is the prepublication version of the following manuscript:

Tracey, D., Morin, A.J.S., Pekrun, R., Arens, A.K., Murayama, K., Lichtenfeld, S., Frenzel, A.C., Goetz, T., & Maïano, C. (In press). Mathematics Motivation in Students with Low Cognitive Ability: A Longitudinal Study of Motivation and Relations with Effort, Self-Regulation, and Grades**.** *American Journal on Intellectual and Developmental Disabilities.*

© 2019. This paper is not the copy of record and may not exactly replicate the authoritative document published in *American Journal on Intellectual and Developmental Disabilities*

Abstract

Expectancy-value theory (EVT) is a popular framework to understand and improve students' motivation. Unfortunately, limited research has verified whether EVT predictions generalize to students with low levels of cognitive ability. This study relies on Grade 5 and 8 data from 177 students with low levels of cognitive ability and a matched sample of 177 students with average to high cognitive ability from the German 'Project for the Analysis of Learning and Achievement in Mathematics'. Results showed that students with low levels of cognitive ability were able to differentiate EVT components. Both groups demonstrated a similar downward developmental trend in motivation from early to middle adolescence, and similar relations between EVT components and levels of efforts, self-regulation, and mathematics class grades.

Key words: motivation, class grades, low cognitive ability, mathematics education.

Understanding the motivational factors that drive students to achieve, apply effort, and self-regulate their learning has occupied educational psychology for more than half a century. Among the theoretical approaches that have guided this research, expectancy-value theory (EVT; e.g., Atkinson, 1957; Eccles, 2009, 2011; Eccles et al. 1983; Heckhausen, 1991; Lauerman, Eccles, & Pekrun, 2017; Pekrun, 1988, 1992, 1993) is one of the longest lasting dominant theoretical frameworks due to its ability to describe both the nature of students' motivation and its impact on student achievement. Regrettably, little EVT research has focused on identifying the nature and impact of motivational components for students with low levels of cognitive ability (Gilmore & Cuskelly, 2009), possibly due to the implicit belief that finer grained distinctions between motivational constructs are harder to achieve with this population, especially at younger ages. With academic success presenting as a key challenge for students with low levels of cognitive ability, and motivation positioned as a core driver of achievement, the application of EVT may provide an opportunity to optimize the achievement of this specific group of students. Unfortunately, the limited research conducted in this area typically relies on small cross-sectional samples with a diversity of disabilities.

In the present study, EVT is applied to test whether the core EVT components of expectancy (operationalized as academic self-concept in educational psychology) and value (based on a differentiation between intrinsic and extrinsic components) are distinguished by students with low levels of cognitive ability during two developmental periods (early and middle adolescence) and how this compares to students with average to high levels of cognitive ability. Furthermore, the present study examines whether these core EVT constructs are able to predict effort, self-regulation of learning, and class grades in mathematics, and the extent to which these relations differ or generalize across students with low, versus average to high, levels of cognitive ability.

Expectancy-Value Theory

In educational psychology, modern EVT proponents (e.g., Eccles & Wigfield, 2002; Lauerman et al., 2017; Wigfield & Eccles, 2000; 2002) suggest that students' academic motivation and subsequent achievement-related behaviors should increase when students (a) believe in their ability to succeed in an academic area of study (expectancy); and (b) place value on that academic area of study (value). Theoretically, the expectancy component of EVT is closely associated with both self-efficacy (e.g. Bandura, 1997) and academic self-concept (e.g. Marsh, 2007). Although it is argued that these two components are theoretically distinct (e.g. Pajares & Miller, 1994), empirically they have shown to be highly correlated, difficult to differentiate, and thus collapsed into a single construct where measures of either self-efficacy or academic self-concept can be used to interchangeably to reflect expectancy of success (Eccles & Wigfield, 2002; Guo, Parker, Marsh, & Morin, 2015; Marsh et al., 2019; Nagengast et al., 2011).

The value component of EVT is known to encompass students' intrinsic valuing of a specific academic area (referring to the extent to which they enjoy this field of study), as well as the extrinsic, or utility, value ascribed to the specific academic area (referring to the extent to which they believe that this field of study will help them to achieve long term goals; e.g., Eccles & Wigfield, 2002). According to EVT (e.g., Eccles & Wigfield, 2002; Eccles et al., 1983; Pekrun, 1993) both the expectancy and value components are expected to have direct causal relations with achievement-related behaviors such as effort, self-regulation of learning, and resulting class grades. Within the value component, intrinsic value is purported to be more beneficial than extrinsic value (Crumpton & Gregory, 2011; Taylor et al., 2014). Of all EVT components, current empirical studies consistently position expectancy (operationalized as academic self-concept) as one of the most robust predictor of academic success (e.g., Arens et al., 2017; Guo, Marsh, Morin, Parker, & Kaur, 2015; Guo et al., 2016; Pajares & Kranzler, 1995).

Although not typically accentuated in modern EVT (for exceptions, see Lauerman et al., 2017; Pekrun, 1993), traditional EVT approaches also proposed that high motivation can only be achieved when both expectancy and value components are high, suggesting a synergistic interaction effect between these two components (Atkinson, 1957; Feather, 1982; Heckhausen, 1991). Recent methodological advances capitalizing on the availability of large scale studies have indeed provided tentative empirical evidence of such a multiplicative relation between expectancy and value components in various domains (Guo, Marsh, Parker, Morin, & Dicke, 2017; Nagengast et al., 2011) and specifically in relation to mathematics (e.g., Guo, Marsh et al., 2015; Guo, Parker et al., 2015; Trautwein et al, 2012). However, even with large samples, these interactions remain small and inconsistent across studies, suggesting that the main effect of each component remains key to achieving a clear understanding of academic success.

Relations among Motivational Constructs, Effort, Self-Regulation of Learning, and Class Grades

Proponents of Bandura's (1997) social-cognitive theory instruct that human functioning is determined by reciprocal relations between cognitive, behavioral, personal, and environmental factors (e.g., Crothers, Hughes, & Morine, 2008). Viewing learning through the lens of social-cognitive theory, educational psychology has interrogated the role that student motivation plays in bolstering student educational success for more than five decades. This area of research has been driven, in large part, by the proposition that student motivation directly influences class grades and academic achievement (Korpershoek, 2016) and that its pliability casts it as a key target in the quest to optimize academic success. With a specific focus on mathematics motivation and achievement, Watts et al.'s (2015) longitudinal study confirmed that even when controlling for prior mathematics achievement, expectancy of success (operationalized as mathematics self-concept) was a strong predictor of later mathematics achievement.

Despite the appeal of motivation as a causal contributor to academic achievement, this relation may not, however, be as clear as initially thought (Korpershoek, 2016; Rotgans & Schmidt, 2012). A review of early motivational research has championed the link between motivation and achievement, suggesting that this link accounts for 16-20% of the variance in student class grades and achievement (Walberg, 1984). However, focusing on achievement goals as an indicator for motivation, a more recent meta-analysis (Huang, 2012) suggests that this relation might be smaller in magnitude than what these early studies suggested. Similarly, empirical studies have reported only small effects of motivation on academic achievement (Korpershoek, 2016; Rotgans & Schmidt, 2012), leading researchers to realize that these relations might be more complex than initially believed.

Current research purports that the relatively low correlation between motivation components and achievement may indicate that motivation constructs are not directly related to grades or achievement, but rather may drive specific learning behaviors (such as self-regulation and effort) which in turn may facilitate student achievement (e.g., Rotgans & Schmidt, 2012). Specific to middle school mathematics achievement, Cleary and Kitsantas (2017) identified that both cognitive and behavioral latent factors exhibited unique effects on mathematics achievement after controlling for prior achievement. Pintrich (2000, 2004) proposed an adaptation of social-cognitive theory encompassing both self-regulation and motivation components, and hypothesized that motivational beliefs influence behavioral strategies (such as effort and self-regulation) which ultimately determine subsequent class grades and academic achievement. The role of these behavioral strategies is substantiated by Liao, Ferdenzi and Edlin (2012) and Nagengast et al.'s (2011) respective findings that motivation did not directly affect class grades or academic achievement, whereas self-regulation of learning did. Still, via stronger longitudinal analyses, Guo, Marsh et al. (2015) demonstrated a reciprocal relation between expectancy and achievement, and between intrinsic value and achievement. Tackling the issue of temporal associations, findings revealed that both expectancy and intrinsic value predicted students' levels of post-secondary education, with expectancy cast as the strongest predictor.

Although statistical advances have assisted researchers in the investigation of relations amongst motivation constructs, effort, self-regulation of learning and academic performance (i.e., class grades or achievement), methodological flaws remain such as the dominant reliance on cross-sectional designs that limit the ability of researchers to make causal claims (e.g., Gilmore & Cuskelly, 2009; Taylor et al., 2014), and a typical emphasis on general academic functioning rather than on domain-specific constructs predicated on social cognitive theory which favors the importance of context (e.g., Crothers et al., 2008; Schunk & Zimmerman, 2003). Another ominous flaw is the omission of studies adopting an EVT framework to investigate relations between motivational beliefs (expectancy, value), behaviors (e.g., effort, self-regulation of learning) and academic performance (i.e., class grades or achievement) for students with low levels of cognitive ability (Huang, 2012) at different developmental stages. Research attests that cognitive ability is the strongest predictor of class grades and academic achievement (Korpershoek, 2016) and that motivational beliefs themselves are affected by prior achievement levels or class grades (e.g., Guo, Parker et al., 2015). With achievement and motivation deficits cast as significant impediments for students with low levels of cognitive ability (Gilmore & Cuskelly, 2009), there is a pressing need to understand the mechanisms at play in relations between motivational constructs, behaviors, and class grades via robust studies that may help to optimize success for this vulnerable group of students.

Students with Low Levels of Cognitive Ability

Following a long history of international shifts in educational philosophy, the inclusion movement now dominates as the most frequently adopted perspective for educating students with intellectual disabilities and cognitive deficits (Craven et al., 2015; Inclusion International, 2009). Despite this

philosophy, Göransson, Hellblom-Thibblin, and Axdorph (2016) lament that the empirical literature has neglected one of the paramount issues that guides major trends within general education: How best to raise academic skills among students with low levels of cognitive ability, such as mathematics skills, which are the focus of the present study. This dearth of research is highlighted in Shurr and Bouck's (2013) systematic review, which found that only 2% of the research on students with intellectual disabilities published during 1996 to 2010 focused on this issue. To realize inclusive education, there must be a move away from a deficit model that merely seeks to identify problems, and a commitment to optimizing their likelihood of academic success. With motivational interventions demonstrating promise in enhancing academic achievement (effect size d = 0.52, Lazowski & Hulleman, 2016) and beliefs such as interest and achievement goals (d = 0.54, Lazowski & Hulleman, 2016) for typically developing students, motivational research and intervention may present as an underexplored remedy for students with low cognitive abilities.

Despite high demand, strong empirical motivational research with students with low levels of cognitive ability is very rare (Frielink, Schuengel & Embregts, 2017) and continues to be plagued by methodological shortcomings. Methodological weaknesses include small sample sizes (Crumpton & Gregory, 2011; Gilmore & Cuskelly, 2009), cross-sectional designs (Zisimopoulos & Galanaki, 2009), and a lack of focus on individuals with low levels of cognitive ability as opposed to other difficulties (Crumpton & Gregory, 2011; Dryer et al., 2016; Zisimopoulos & Galanaki, 2009). Although emerging research shows that adults with low levels of cognitive ability can distinguish intrinsic from extrinsic motivation (Frielink, Schuengel & Embregts, 2017), theoretical models such as EVT have either never been tested among students with low levels of cognitive ability, or at least never been used to guide empirical investigations (Dryer et al., 2016), which has hindered the advancement of theoretical knowledge and educational practice for these students. As a result, it is unknown whether students with low levels of cognitive ability are able to differentiate among the expectancy and value components of motivation, and there is an established view that youth with low levels of cognitive ability may be unable to do so (Katz & Cohen, 2014; also Tracey, Craven, & Marsh, 2015).

A systematic search of the literature only identified three longitudinal studies of academic motivation among individuals with low levels of cognitive ability. Gilmore and Cuskelly (2009) longitudinally followed 25 children with Down syndrome over two time points (aged 4 to 6 years at time 1 and 11 to 15 years at time 2). Their results showed that mastery motivation (operationalized as persistence and preference for challenge) was important above and beyond cognitive ability in the prediction of academic ability for children with Down syndrome. Participants who were more motivated at time 1 performed more competently in reading and mathematics at time 2, even when cognitive ability at time 1 was controlled for. Blair, Greenberg, and Crnic (2001) followed 41 children with low levels of cognitive ability (initially aged 1 to 5 years) over a 12- month interval. They concluded that changes in goal-directed behaviors for children with low levels of cognitive ability were similar to those found among typically developing children. However, they did not consider the impact of these developmental trends on achievement-related behaviors. Finally, Crumpton and Gregory (2011) conducted a one year longitudinal study among a sample of 44 high school low achievers and found that students' perceived intrinsic motivation and utility value were related to behavioral engagement in classroom tasks, but not to class grades. This finding appears to support Pintrch's (2000, 2004) view that behavioral strategies (such as effort and self-regulation) are driven by motivation more so than class grades or achievement. Still, despite their interest, these studies all remained limited by the reliance on small non-representative samples of students and their inconsistent and limited operationalizations of academic motivation and outcomes.

Given that students with low levels of cognitive ability face a stark educational disadvantage, the void of empirical and theoretical evidence to help reduce this disadvantage justifies recent calls (Gilmore & Cuskelly, 2009; Korpershoek, 2016) for research evidence aiming to improve our understanding of academic motivation for students with low levels of cognitive ability. Indeed, motivational interventions have been found to impact academic achievement (as noted earlier: d = 0.52, Lazowksi & Hulleman, 2016) more than comprehensive school reform models (d = 0.11, Borman et al., 2003). Furthermore, sophisticated designs are required to determine the predictive capacity of motivation components for measures of academic performance (i.e., class grades or achievement) and possible mediators, with the objective of guiding programs to alleviate disadvantage for this vulnerable group of students.

The Present Study

The present study is premised on EVT (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000; 2002; Pekrun, 1993) and aims to elucidate the nature of motivation for students with low levels of cognitive ability

during and across two distinct developmental periods, as well as the relations between motivational constructs and the key academic outcomes of effort, self-regulation of learning, and class grades in mathematics. Importantly, this study relies on a longitudinal design and on a larger sample than what is typically seen in this field of research, as well as on a matched sample of students with average to high levels of cognitive ability used for comparison purposes.

Objective 1. The first aim of the study was to determine whether the expectancy and value components proposed by EVT can be differentiated equally well by students with low levels of cognitive ability as by their peers with average to high levels of cognitive ability, and whether motivational levels differ as participants move from early adolescence (Grade 5, which occurs right after the transition to secondary school in Germany) to middle adolescence (Grade 8).

Objective 2. Second, this study aims to investigate whether the motivational constructs proposed by EVT predicts students' levels of effort, self-regulation of learning, and class grades in mathematics, and whether this prediction is similar for students with low levels of cognitive ability and for those with average to high levels of cognitive ability. In doing so, we also verify whether these predictions and group similarities or differences generalize to early and middle adolescence. Motivation beliefs are malleable to intervention (Lazowski & Hulleman, 2016; Wigfield et al., 2015), therefore, distinguishing how they influence critical educational outcomes such as effort, self-regulation of learning, and achievement in mathematics is likely to have important implications for instructional and intervention planning for students with low levels of cognitive ability.

Objective 3. Third, using a more robust longitudinal mediation analysis, we assess the relations between the EVT components measured in early adolescence and changes in academic outcomes occurring across the early and middle adolescence period, and verify whether changes in behavioral strategies (effort and self-regulation) can mediate the relations between the EVT components and class grades in mathematics. In doing so, we also test the extent to which these relations are similar for students with low levels of cognitive ability and those with average to high levels of cognitive ability.

Our decision to rely on class grades, rather than on standardized achievement tests, is anchored in the fact that, despite their limitations (e.g., reliability: Graham, 2015; grading-on-a-curve leading to limited between-class differences: Arens & Morin, 2016; Arens, Morin, & Waterman, 2015; Marsh et al., 2005, 2015) class grades are a key outcome for students in educational systems throughout the planet. Class grades are the main outcome taken into account in determining long term educational outcomes (attainment, employment opportunity, postsecondary admission, lifetime income, etc.). This is especially true in Germany where standardized acheivement tests are rarely used and class grades define entry into post secondary education (e.g., University). As such, class grades represent a highly ecologically valid educational outcome to consider.

Method

Sample and Procedures

The present study uses a longitudinal dataset from the Project for the Analysis of Learning and Achievement in Mathematics (PALMA; see Arens et al., 2017; Marsh, Pekrun, Lichtenfeld, Guo, Arens, & Murayama, 2016; Murayama, Pekrun, Suzuki, Marsh, & Lichtenfeld, 2016; Pekrun et al., 2007; Pekrun, Lichtenfeld, Marsh, Murayama, & Goetz, 2017). PALMA is a large-scale longitudinal study investigating the development of mathematics achievement and its determinants (e.g., motivation, classroom instruction, family variables) in German secondary schools located in the German state of Bavaria and starting in Grade five. Sampling and yearly assessments were conducted by the German Data Processing and Research Center (DPC) of the International Association for the Evaluation of Educational Achievement (IEA). Ethical approval for the PALMA study was given by the ethics committee of the State Ministry of Education and Cultural Affairs of Bavaria. The sample represented the typical student population in the German federal state of Bavaria in terms of age, gender, home language, urban-rural location, and SES. Participation rate at the school level was 100%, and two classes were randomly selected from each school. In PALMA, students answered a questionnaire towards the end of each successive school year. All instruments were administered in the classrooms by trained test administrators. Participation in the study was voluntary and confidential, and parental consent was obtained for everyone.

At the first measurement wave in grade five (grade five is the time of entry into secondary schools in Germany), the sample included 2,070 students (49.6% female). The students then had a mean age of 11.75 (SD = 0.68) which is the typical age for fifth grade students in Germany. For subsequent data collections, the study did not only follow students who had already participated in earlier assessments, but also included

students who more recently entered classrooms participating in the PALMA study and thus had not yet participated in the study (for more details on the sampling procedure, see Pekrun et al., 2007). By ninth grade, the time at which students with low levels of cognitive ability were likely to leave school, a total sample of N = 3,425 (50% boys) had participated in at least one assessment.

The sample of students with low levels of cognitive ability was identified using the figural subtest of the German adaptation of Thorndike's Cognitive Abilities Test (Kognitiver Fähigkeitstest, KFT 4-12+R; Heller & Perleth, 2000) at all measurement waves (an example of item from this subtest is reproduced in Appendix A of the online supplements). Age-related sample-specific IQ norms (M = 100, SD = 15) for this test were created based on the total sample of students participating in the PALMA study. Based on these norms, a sample of 177 students (58.2% boys; M_{age} = 11.86), who consistently obtained an IQ less than or equal to 85, were retained (M_{IQ} = 74.92; SD_{IQ} = 5.43).

In order to obtain a comparable sample of youth presenting average to high levels of cognitive ability of the same size as this sample of youth with low levels of cognitive abilities, we relied on a one-to-one matching procedure to select participants comparable in terms of age, gender, and home language (e.g., Morin et al., 2017a, 2017b). This matching process was conducted using a genetic search algorithm automated as part of the R Matching package (Sekhon, 2011). This algorithm combines propensity score matching (based on a logistic regression of the grouping variable on the covariates) with multivariate matching in a way to maximise balance (equivalence) across groups on the covariates used in the matching procedure (gender, age, and household language). A more extensive presentation of this algorithm is provided in Sekhon (2011). The resulting matched sample of 177 students with average to high levels of cognitive ability (57.1% boys; M_{age} = 11.90; M_{IQ} = 100.62; SD_{IQ} = 14.88) can be considered to be equivalent to the sample of students with low levels of cognitive ability, while also preserving the representative nature of the PALMA data set in terms of gender, age, and household language.

Given the reduced sample size utilized here, it was not possible to conduct analyses incorporating all six measurement waves. As such, the decision was made to focus on the first (Time 1) and fourth (Time 2) measurement wave so as to ascertain the extent to which our results would generalize to early (Grade 5) and middle (Grade 8) adolescent students (data on class grades was not available in Grade 9).

Measures

Expectancy. Students' mathematics self-concept was used to assess the expectancy component. In PALMA, six items are used to measure mathematics self-concept (e.g., "I can solve mathematics tasks easily"; "It is easy for me to understand mathematics"). These items are rated on a 5-point Likert scale (1 = *not at all true* to 5 = *completely true*). Scale score reliability estimates for this scale were good for the total sample at both measurement waves (Time 1: α = .86; Time 2: α = .91) as well as for students with low (Time 1: α = .86; Time 2: α =.88) and average to high (Time 1: α = .87; Time 2: α =.92) levels of cognitive abilities. Previous studies using PALMA established the validity of this measure by demonstrating its association with mathematics achievement and effort (Arens et al., 2017; Marsh et al., 2016, 2019).

Value. The value components were measured using a subscale targeting intrinsic value (six items: e.g., "I am interested in mathematics"; "The things that we are doing in mathematics lessons are very exciting for me") and a second subscale targeting extrinsic value (two items: "I think that you cannot succeed in life without mathematics"; "I think that the things I learn in mathematics are useful for my future"). All items were rated on a 5-point Likert scale (1 = *not at all true* to 5 = *completely true*). The scale score reliability estimate for the intrinsic value scale was good at both time poins in the total sample (Time 1: α = .89; Time 2: α = .87), as well as for students with low (Time 1: α = .87; Time 2: α = .87) and average to high (Time 1: α = .91; Time 2: α = .88) levels of cognitive abilities. However, the scale score reliability of the extrinsic value subscale was not satisfactory (α = .44 to .55), due to the small number of items included in this scale (Streiner, 2003). When the Spearman-Brown formula was used to adjust these estimates based on eight equivalent items, all estimates reached acceptability levels among the total sample (Time 1: α = .80; Time 2: α = .79), as well for students with low (Time 1: α = .77; Time 2: α = .76) or average to high (Time 1: α = .83; Time 2: α = .83) levels of cognitive abilities. However, the latent variable models used here (see the analysis section) were explicitly corrected for measurement errors.

Class grades. Students' class grades in mathematics were obtained from students' end-of-school year reports. In Germany, school grades range from 1 to 6 with 1 depicting the best grade and 6 the worst grades. For ease of interpretation, school grades were reversed-coded prior to the analyses. In the present matched samples, the average one year test-retest correlation for class grades across all years of PALMA is r = .52 for the group of students with low levels of cognitive abilities and r = .67 for the group of students

with average to high levels of cognitive abilities, which is aligned with the one-year test-retest correlations (*r* =.57 to .62) reported for the full PALMA data set (Pekrun et al., 2017).

Effort. Effort in mathematics was assessed using seven items asking students for the amount of effort they were willing to invest in mathematics tasks and in learning mathematics in general (e.g., "In math, I try to do the best that I can"). These items were rated on a 5-point Likert scale (1= *not at all true* to 5 = *completely true).* The scale score reliability estimate of this scale was satisfactory at both time waves among the total sample (Time 1: α = .79; Time 2: α = .80), as well as for students with low (Time 1: α = .83; Time 2: α = .79) or average to high (Time 1: α = .71; Time 2: α = .80) levels of cognitive abilities.

Self-regulation of learning. Self-regulation of learning in mathematics was assessed using six items asking students about self-regulated behaviors that can be used during the mathematics learning process. These items address students' self-directed goal setting (e.g., "When learning mathematics, I set myself objectives that I aim to achieve during the learning process"), self-monitoring ("I notice when my progress in mathematics is too slow"), and self-evaluation (e.g., "I can evaluate quite correctly how good I am in mathematics at the moment"). The items were rated on a 5-point Likert scale (*1 = not at all true* to 5 = *completely true*). The scale score reliability estimate of this scale was satisfactory at both time waves among the total sample (Time 1: α = .67; Time 2: α = .72), as well as for students with low (Time 1: α = 73; Time 2: α = .65) or average to high (Time 1: α = .60; Time 2: α = 75) levels of cognitive abilities.

Analyses

Model Estimation

All analyses were conducted using Mplus 7.4's robust Maximum Likelihood (MLR) estimator, which is robust to the non-normality of the Likert response scales used in the present study (Muthén, & Muthén, 2015). Full Information Maximum Likelihood (FIML) was used to handle the missing data present in the current study (Enders, 2010). FIML allowed us to estimate the longitudinal models using the full available sample of 354 (177 matched participants from each group), rather than relying on a suboptimal listwise deletion strategy focusing only on participants having completed both assessments (Enders, 2010). This procedure has comparable efficacy to multiple imputation, while being more computationally efficient (Enders, 2010; Jeličič, Phelps, & Lerner, 2009; Larsen, 2011), and allows missing data to be conditional on all observed and latent variables included in the model. Additional information on missing data and on more technical details of model estimation, are provided in Appendix B of the online supplements. Correlations and descriptive statistics for all variables are reported in Appendix C of the online supplements. These statistics support reasonable associations among constructs, a reasonable amount of differentiation among constructs, and no apparent evidence of multicollinearity.

Objective 1. Measurement Models and Measurement Invariance

Model Specification. To make sure that participants from both groups were able to differentiate among the various EVT components assessed in the present study, we first estimated a series of confirmatory factor analyses (CFA) in which the three EVT components (expectancy, intrinsic value, and extrinsic value) were assessed based on their a priori indicators. In these models, each item was used to define its a priori factor, all factors were allowed to correlate, and no cross-loading or correlated uniqueness was allowed. These CFAs were used to tests the invariance of this measurement model across groups of participants and time points in the following sequence (Millsap, 2011): (1) configural invariance (the same measurement model is estimated in both groups or time-points); (2) weak invariance (factor loadings); (3) strong invariance (factor loadings and item intercepts); (4) strict invariance (factor loadings, item intercepts, and items uniquenesses); (5) variance/covariance invariance (factor loadings, item intercepts, items uniquenesses, and latent variances and covariances); and (6) latent means invariance (factor loadings, item intercepts, items uniquenesses, latent variances and covariances, and latent means).

The extrinsic value factor was estimated using only two indicators, creating a locally underidentified construct (although the overall model remains overidentified). This construct was thus locally-identified using essentially tau-equivalent constraints (ETEC; Little, Lindenberger & Nesselroade, 1999). This technique involves placing equality constraints on the loadings of both indicators to locate the construct at the true centroid of the indicators. These ETEC were incorporated as an additional step in the invariance sequence (2b) to allow for proper tests of weak invariance.

Unfortunately, possibly because of the reduced sample size available in the present study, it was not possible to simultaneously test the 2 (groups) by 2 (time points) invariance of the model (these tests systematically failed to converge on proper solutions). For this reason, separate measurement invariance tests were conducted across groups of participants first at Time 1, and then at Time 2. Then, the measurement

invariance of the model was also tested across time points using the total sample of participants. In these longitudinal models, a priori correlated uniquenesses between matching indicators of the factors utilized at the different time-points had to be included to avoid obtaining inflated stability estimates for the latent constructs (e.g., Marsh, Abduljabbar et al., 2013).

Model Comparison. The degree of correspondence between the measurement models and the data was assessed using goodness-of-fit indices (Hu, & Bentler, 1999; Marsh, Hau, & Wen, 2004) including the comparative fit index (CFI), the Tucker-Lewis index (TLI) and the root mean square error of approximation (RMSEA) and its 90% confidence interval, in addition to the chi-square test of exact fit. Given the known oversensitivity of the chi-square test of exact fit (χ^2) to sample size and minor model misspecifications (e.g., Marsh, Hau, & Grayson, 2005), we relied on values greater than or equal to .90 for the CFI and TLI to indicate adequate model fit, although values greater than or equal to.95 are preferable. Values smaller or equal to .10, .08 or .06 for the RMSEA were taken to support marginal, acceptable and excellent model fit, respectively. Like the chi-square, chi-square difference tests are sensitive to sample size and minor misspecifications. Nested models were thus compared using changes (Δ) in goodness-of-fit indices, with ΔCFI/TLI $\leq$ .010 and ΔRMSEA $\leq$ .015 taken to support the equivalence of the models (Chen, 2007; Cheung & Rensvold, 2002).

Objective 2. Predictive Analyses

We relied on three predictive models in which the EVT variables were allowed to predict scores on the latent (effort and self-regulation) and observed (class grades) outcomes. In the first model, analyses were conducted based on Time 1 measures. In the second model, these analyses were conducted based on the Time 2 measures to verify the extent to which results would be replicated across developmental periods. Although these analyses were cross-sectional and thus not perfectly suited to tests of mediation, they provided a way to directly test whether the relations between the EVT predictors and the academic outcomes would be replicated across developmental periods. In our third model, a longitudinal SEM model was used in which the EVT predictors measured at Time 1 were allowed to predict all outcomes (effort, self-regulation, class grades) at Time 2. Across all predictive models, in order to maintain a proper level of control for measurement errors, we relied on single-item latent Structural Equation Models (SEM; e.g., Bollen, 1989) – except for class grades, which was specified as an observed score. Additional details are provided in Appendix B of the online supplements.

Objective 3. Longitudinal Mediation.

In order to test our a priori mediational hypothesis, we relied on a fourth predictive model. In this model, Time 2 class grades were controlled for Time 1 class grades, and Time 1 and Time 2 measures of the mediators (effort and self-regulation) were used to estimate latent change factors (i.e., McArdle, 2009, also see Appendix B in the online supplements) representing changes in these constructs occurring between Time 1 and Time 2. As noted by Morin, Marsh, and Nagengast (2013), this model provided a way to test for mediation based on two repeated measurements while ensuring a complete temporal ordering of the variables involved in the mediation chain: Time 1 EVT constructs predict change in the mediators occurring between Time 1 and Time 2, which themselves predict Time 2 class grades controlled for class grades.

The predictive model used for these analyses is illustrated in Figure 1, from which the measurement part (single-indicator latent variables) and the correlations have been taken out for ease of illustration. In these analyses, the goal was to contrasts two alternative predictive models: (a) a fully mediated model (bold arrows in Figure 1), in which the EVT factors were allowed to predict the effort and self-regulation change factors, which in turn were allowed to predict observed class grades at Time 2; (b) a partially mediated model (dashed and bold arrows in Figure 1) in which direct paths were also included between the EVT factors at Time 1 and the observed class grades at Time 2. The statistical significance of the indirect effects forming the mediation mechanisms was tested using bias-corrected bootstrap 95% confidence intervals (CI; using 1000 bootstrap samples) as recommended by Cheung and Lau (2008; Lau & Cheung, 2012): CI excluding 0 indicated that the indirect effect could be considered to be statistically significant.

Model Comparison Strategy: Objectives 2 and 3

Because it was not possible to compare model fit in the estimation of our predictive models (see Appendix B of the online supplements for details), we relied on a model-building strategy to select the optimal final models. We started by the estimation model in which all paths were freely estimated across both groups of participants. In order to verify whether the predictive paths could be considered to be equivalent across groups, we relied on the multivariate delta method (Raykov & Marcoulides, 2004), implemented via the Mplus MODEL CONSTRAINT command, to test the equality of these paths across

samples. Based on this examination, we then estimated a model in which the paths that could be considered to be equivalent across samples were constrained to equality. In the fourth (mediation) model, the parameter estimates from these two solutions (invariant and non-invariant) were contrasted to see whether there was any evidence of statistical significance for the partial mediation paths (i.e., the direct effects of the EVT predictors on class grades). Non-significant partial mediation paths were then dropped from the model, and a final model was estimated.

Results

Objective 1. Measurement Invariance and Differentiation among EVT Components

The results from the CFA estimated to test the measurement invariance of the EVT ratings across groups and time points are reported in Table 1. These results first show that all configural models, in which the same basic measurement model was estimated across groups or time points with no added equality constraints, provided an acceptable level of fit to the data, with CFI and TLI values ≥ .909 and RMSEA values ≤ .077. In addition, at both time points, the results support the weak (factor loadings) and strong (item intercepts) invariance of these measurement models across groups of participants (ΔCFI and ΔTLI ≤ .010; ΔRMSEA ≤ .015), but not their strict (item uniqueness) measurement invariance (ΔCFI and ΔTLI > .010), suggesting that at least some item ratings incorporate levels of item-specific measurement error that differ across samples. We thus pursued a model of partial strict invariance to locate the items responsible for these differences, based on an inspection of parameter estimates and modification indices (e.g., Byrne, Shavelson, & Muthén, 1989). This examination revealed that the invariance constraints had to be relaxed across groups for four items at Time 1, and two items at Time 2, revealing mainly that responses to these items provided by the sample of students with low levels of cognitive ability incorporated a slightly larger amount of measurement error. The next models supported the partial strict invariance of EVT ratings, as well as the complete invariance of the latent variances, covariances, and means across samples at both time points (ΔCFI and ΔTLI ≤ .010; ΔRMSEA ≤ .015). This result shows that the levels of motivation components, their variances, and their interrelations in students with low levels of cognitive ability are indistinguishable from those of students with average to high levels of cognitive ability at both time points, when estimated using latent variable models controlled for measurement errors.

Tests of measurement invariance conducted across time points for the complete sample of participants (merging the two groups was supported by evidence of partial strict, latent variances-covariance, and latent mean invariance of the EVT ratings across samples) supported the weak, strong, and strict measurement invariance of the models, as well as the invariance of the latent variances and covariances across time points (ΔCFI and ΔTLI ≤ .010; ΔRMSEA ≤ .015). However, the results revealed significant latent means differences across time points (ΔCFI and ΔTLI > .010), showing that levels of expectancy, intrinsic value, and extrinsic value tended to decrease respectively by -.437 *SD* units, -.918 *SD* units, and -.652 *SD* units over time for both groups of participants.

When the parameter estimates from the most invariant models are examined, they revealed factors that are generally well-defined by moderate to strong factor loadings ranging across models, from .615 to .866 (M = .748) for the expectancy factor, from .580 to .866 (M = .729) for the intrinsic value factor, and from .458 to .666 (M = .561) for the extrinsic value factor. As expected, the three EVT factors were moderately correlated, albeit remaining clearly distinguishable for both groups of participants (r = .493 to .736, M = .615). Finally, rank-order stability in EVT ratings was found to be moderate to low in magnitude across the Grade 5 to 8 time interval, with stability estimates of r = .567 for expectancy, r = .436 for intrinsic value, and r = .204 for utility value.

Objective 2: Simple Predictive Models

Time 1. Examination of the parameter estimates obtained from the Time 1 model supported the equivalence of the predictions across groups of participants. A final model in which all predictions were constrained to be equivalent across groups was thus estimated and achieved a satisfactory level of fit to the data (χ^2 = 23.160, df = 12; CFI = .975; TLI = .937; RMSEA = .080). Results from this model, reported in the top section of Table 2, reveal significant relations whereby higher scores on extrinsic value were associated with higher levels of effort and self-regulation, but with lower class grades. In contrast, higher scores on expectancy were associated with higher class grades. These explained 75.6% (average to high cognitive ability) to 80.8% (low levels of cognitive abilities) of the variance in effort scores, 69.9% (average to high levels of cognitive abilities) to 73.3% (low levels of cognitive abilities) of the variance in self-regulation scores, and 22.9% (average to high levels of cognitive abilities) to 31.2% (low levels of cognitive abilities) of the variance in class grades.

Time 2. Parameter estimates obtained from the Time 2 model supported the equivalence of the predictions across groups of participants. A final model in which all predictive relations were constrained to be equivalent across groups was thus estimated and generally achieved a satisfactory level of fit to the data (χ^2 = 27.073, *df* = 12; CFI = .963; TLI = .908), albeit the RMSEA value was marginal (.100). Results from this model, reported in the second section of Table 2, revealed significant positive relations whereby higher scores on extrinsic value were associated with higher effort and self-regulation. In addition, higher scores on expectancy were associated with higher levels of self-regulation and class grades. These relations explained 31.1% (low levels of cognitive abilities) to 31.8% (average to high levels of cognitive abilities) of the variance in effort scores, 43.6% (average to high levels of cognitive abilities) to 54.2% (low levels of cognitive abilities) of the variance in self-regulation scores, and 50.5% (average to high levels of cognitive abilities) to 54.5% (low levels of cognitive abilities) of the variance in class grades.

Longitudinal prediction model. Examination of the parameter estimates obtained from the longitudinal model in which Time 1 EVT ratings predicted Time 2 outcomes supported the equivalence of the predictions across groups of participants. A final model in which all predictions were constrained to be equivalent across groups was thus estimated and achieved an excellent level of fit to the data (χ^2 = 14.268, *df* = 12; CFI = .993; TLI = .981; RMSEA = .033). Results from this model, reported in the third section of Table 2, reveal significant positive relations whereby scores on the expectancy component at Time 1 were associated with higher levels of self-regulation and class grades at Time 2. None of the other coefficients were significant. These relations explained 21.8% (average to high levels of cognitive abilities) to 33.9% (low levels of cognitive abilities) of the variance in self-regulation scores, and 15.4% (average to high levels of cognitive abilities) to 25.5% (low levels of cognitive abilities) of the variance in class grades.

Objective 3. Longitudinal Mediation.

Examination of the parameter estimates obtained from the model of partial mediation supported: (a) the equivalence of the predictions across groups of participants; and (b) the need to maintain one direct predictive path between expectancy and class grades. A final, partially mediated model in which all predictions were constrained to be equivalent across groups of participants was thus estimated and achieved a satisfactory level of fit to the data (χ^2 = 49.731, *df* = 26; CFI = .965; TLI = .902; RMSEA = .073). Results from this model are reported in the last section of Table 2, and illustrated in Figure 2. This model provides an interesting complementary perspective to the cross-sectional results in showing that initial levels of intrinsic value at Time 1 predict more increase of effort over time, while initial levels of extrinsic value at Time 1 predict a reduction of effort over time. In addition, these results support those from the cross-sectional models in positioning expectancy at Time 1 as the sole positive predictor of increases over time in class grades. These results reveal no evidence of significant mediation effects. These relations explained 38.6% (average to high cognitive ability) to 43.7% (low cognitive ability) of the variance in changes in effort scores over time, as well as 28.6% (average to high cognitive ability) to 36.9% (low cognitive ability) of the variance in class grades at Time 2.

Discussion

Extensive theoretical and empirical work has been conducted to examine the structure of motivation within the EVT framework, as well as the impact of motivation components on educational outcomes such as effort, self-regulation of learning, and class grades (e.g., Nagengast et al., 2011; Trautwein et al., 2012). The educational disadvantage experienced by students with low levels of cognitive ability makes it clear that these students warrant equal, if not accrued, research attention as their typically developing peers in order to advance the development of effective educational practices. The volume and methodological rigor applied to research conducted with this specific population, however, is inadequate as yet, resulting in insufficient knowledge related to issues as basic as whether students with low levels of cognitive ability have the same ability to differentiate among motivational components, and the impact of these components on educational outcomes. This study sought to address some of the shortcomings of previous research and to contribute new knowledge in this area, based on the premise that all students have a right to curricula and pedagogy that seek to maximise academic success regardless of cognitive ability.

Differentiation of Among EVT Components and Developmental Differences

The first objective of this study was to ascertain how well students with low levels of cognitive ability are able to differentiate among the expectancy and value components of academic motivation, and whether this ability differs across two distinct developmental periods (early versus late adolescence). Our results revealed that whilst the three EVT components were found to be moderately correlated, they remained distinguishable for both groups of participants, during both developmental periods. The underlying

measurement models were essentially the same (i.e., measurement invariance) for students with low versus average to high levels of cognitive ability during early and middle adolescence. This finding challenges the assumption that these students are unable to provide reliable and well-differentiated self-reports (e.g., Katz & Cohen, 2014) by demonstrating that they are indeed able to do so, at least as well as their typically developing peers (Frielink, Schuengel, & Embregts, 2017). This study also examined whether motivational beliefs differed as participants moved from early to middle adolescence. An interesting finding was that there were no mean level differences in motivational levels across groups, which appears to contradict the general perception that students with low levels of cognitive ability are less motivated than their peers (Gilmore & Cuskelly, 2009). Furthermore, in accordance with previous research demonstrating a decrease in motivation levels during the adolescent period (e.g., Jacobs et al., 2002; Gottfried, Fleming & Gottfried, 2001), our results similarly showed that levels of expectancy, intrinsic value, and extrinsic value decreased substantially over time, and did so equally for both groups of students.

EVT Components and the Prediction of Effort, Self-Regulation of Learning, and Class Grades

Our second objective was to evaluate the relations between EVT components and students' levels of effort, self-regulation of learning, and class grades in mathematics to systematically assess whether these relations differed across samples of students with low levels of cognitive ability relative to their peers with average-to-high levels of cognitive abilities, and to test whether these relations and group differences would vary across developmental periods. Time specific predictive models first revealed that identical relations applied to both groups of participants. In contrast, these models revealed the presence of few, but important, differences in the relations estimated across early and middle adolescence. More precisely, in early adolescence, extrinsic value positively related to students' effort and self-regulation of learning, but negatively to class grades in mathematics. In contrast, in middle adolescence, these negative relations for extrinsic value were no longer apparent. However, the results obtained during both developmental periods showed no significant relations of intrinsic value with any of the outcomes considered, and revealed that expectancy was the only motivational component to have a positive relation with class grades in mathematics, in addition to having a positive relation with self-regulation of learning. Perhaps even more interesting was the observation that expectancy was the only EVT component to be related to class grades and self-regulation of learning three years later, supporting the robustness of these relations. The findings highlight the importance of bolstering students' academic self-concepts (or expectancies of being able to succeed; e.g., O'Mara, Marsh, Craven, & Debus, 2006) in order to improve class grades.

Longitudinal Prediction and Mediation

Our final objective was to test how well EVT components were able to predict changes in effort, self-regulation of learning, and class grades in mathematics over time, and whether changes in effort and self-regulation of learning were able to mediate the longitudinal relations between EVT components and class grades in mathematics based on Pintrich's (2000, 2004) proposal. Our results first confirmed the significant impact of expectancy on class grades in mathematics and self-regulation of learning, with expectancy identified as the sole positive predictor of increases over time in class grades in mathematics. These results thus clearly support the idea that student's beliefs in their ability to succeed in an academic area (expectancy; operationalized as academic self-concept) is the most influential EVT component to drive improvements in class grades in mathematics for students in contemporary education systems (which is consistent with previous findings of Arens et al., 2017; Guo, Marsh et al., 2015; Guo et al., 2016).

Contrary to our cross-sectional and simple longitudinal findings, intrinsic value was found to predict significant increases over time in efforts, whereas extrinsic value predicted decreases in efforts. Taken together, our results thus suggest that the benefits of interventions aiming to build the intrinsic value of mathematics (seeking to improve students' enjoyment of, and interest in, maths; e.g., Hulleman & Harackiewicz, 2009) may represent a profitable long-term investment. In contrast, developing extrinsic value (seeking to improve awareness of the utility of mathematics; e.g., Harackiewicz, Canning, Tibbetts, Priniski, & Hyde, 2016; Hulleman, Kosovich, Barron, & Daniel, 2017) may only represent a short term fix to bolster students' levels of efforts and self-regulation at a specific point in time that appears less risky to use in middle adolescence given its negative impact on class grades in early adolescence. The results from our final longitudinal mediation models further reinforce the risks of extrinsic value, in showing that early levels of extrinsic values are related to a decrease over time in effort levels. These apparent undesirable effects of extrinsic value are aligned with prior results (Crumpton & Gregory, 2011; Taylor et al., 2014), and suggest that interventions aiming to bolster students' extrinsic valuing of mathematics should be followed by actions aiming to convert this extrinsic valuing into more intrinsic motives.

Finally, and contradicting Pintrich (2000, 2004), no evidence of mediation was found, which could possibly be due to either a lack of mediation effects, or perhaps to the long time lag considered. The latter explanation would suggest that mediation could occur over shorter time periods (e.g., one year), a possibility that would need to be investigated in future studies. Interestingly, our results supported the equivalence of the predictive paths across samples of students with low levels of cognitive ability relative to their peers with average-to-high levels of cognitive abilities. Taken together, this similarity of results obtained across these two samples of students provides very promising evidence supporting the idea that the far more extensive knowledge base built within typically developing populations could realistically be used to guide interventions conducted among students with low levels of cognitive ability.

Limitations

Although this study surpasses the methodological rigor typically utilized in studies of students with low cognitive ability, it is not without shortcomings. First, we relied on self-reports of motivation, effort, and self-regulation of learning which, although this limit is shared with most EVT research, may not reflect actual educational behaviors. In contrast, our measure of class grades was based on objective school reports. This decision was anchored in a desire to rely on a maximally ecologically valid measure of student performance (Pekrun et al., 2014, 2017). A key advantage of class grades is that they represent performance in terms of the math curriculum truly taught in classes and learned by the students and for this reason may be superior to alternative measures in terms of curricular validity. As is typical for class grades however, more detailed information about reliability was not available, apart from an imperfect indicator of one-year test-retest stability. However, in German secondary schools, end-of-year class grades are summative and based on multiple exams conducted throughout the year, which is likely to boost their reliability in comparison to grades obtained on a single exam given the well documented positive relation between the number of indicators and reliability (Streiner, 2003).

Fortunately, because class grades were included in our models as a distal outcome, their limited reliability (Graham, 2015) could not be considered to have a biasing effect on the relations (Ree & Carretta, 2006). However, research has shown class grades and standardized achievement test scores to provide distinct, and complementary, sources of information about students' academic performance (Arens & Morin, 2016; Arens et al., 2015; Marsh et al., 2005, 2015). For this reason, it would be highly informative for future research to assess the extent to which the current results generalize to standardized achievement measures. Likewise, our results are limited to a single academic subject (mathematics), and future research should more thoroughly assess the extent to which they generalize to additional subjects.

Second, recent operationalizations of EVT (e.g., Nagengast et al., 2011) underscore the need to incorporate interaction effects among expectancy and value components in studies of EVT components. Unfortunately, no evidence in support of the proposed EVT interaction was found in any of the models estimated here (see Appendix B of the online supplements for additional details), which could be possibly due to a lack of statistical power which is often an issue in tests of latent interactions (Marsh, Hau, et al., 2013). Still, even with large samples, previous studies generally found such interactions to be small and inconsistent (Guo, Marsh et al., 2015; Guo, Parker et al., 2015; Guo et al., 2017; Nagengast et al., 2011; Trautwein et al, 2012). Third, the PALMA study is designed to focus on regular education classes. Thus, despite being representative of the German state of Bavaria across all three achievement tracks (low, middle, and high), it excluded specialized establishments and thus students with the lowest levels of cognitive abilities for whom the current results may not generalize.

Conclusions and Recommendations

This study provides new knowledge about not only the motivational structure, development, and role of EVT motivation components in the prediction of valuable academic outcomes, it also suggests recommendations for educational practice. Thus, our results first demonstrate that a consideration of academic motivation is at least equally important for students with low levels of cognitive ability, which present comparable levels of motivation as their peers, similar developmental trends, and identical relations with educational outcomes. Importantly, these findings challenge the entrenched notion that students with low levels of cognitive ability tend to be less motivated than their peers with average to high levels of cognitive ability. Our results also revealed that amongst the various relations identified, the most influential factor driving class grades for all students (irrespective of their levels of cognitive ability) is the belief that they can be successful in their studies (i.e., expectancy). Motivational beliefs play a critical role in the development of academic competencies, perseverance, and success, yet teachers and school psychologists do not systematically implement motivation instruction in their work with students with low levels of

cognitive ability (Grigal, Neubart, Moon, & Graham, 2003; Cleary, Gubi, & Prescott, 2010). To advance the academic success of students with low levels of cognitive ability, and advance the goals of inclusive education, our results suggest that schools may need to provide tangible learning support focusing on all components of motivation (Dryer et al., 2016). Motivation is amenable to change (Lazowski & Hulleman, 2016; Wigfield et al., 2015) and thus worthy of consideration for educational agendas. Informed by empirical research, such as this study, educational shifts can build momentum to realize improved outcomes for disadvantaged students.

References

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*, 621-634.

Arens, A.K., & Morin, A.J.S. (2016). Relations between teachers' emotional exhaustion and students' educational outcomes. *Journal of Educational Psychology, 108* (6), 800-813.

Arens, A.K., & Morin, A.J.S. & Waterman, R. (2015). Relations between classroom disciplinary problems and student motivation: Achievement as a potential mediator? *Learning and Instruction, 39*, 184-193.

Atkinson, J.W. (1957). Motivational determinants of risk-taking. *Psychological Reviews, 64*, 359-372.

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

Blair, C., Greenberg, M., & Crnic, K. (2001). Age-related increases in motivation among children with mental retardation and MA- and CA- matched controls. *American Journal on Mental Retardation, 106*, 511-524.

Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.

Borman, G. D., Hewes, G.M., Overman, L.T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*, 125-230.

Byrne, B., Shavelson, R., & Muthén, B. (1989). Testing for the equivalence of covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456-466.

Chen, F.F. (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling, 14*, 464-504.

Cheung, G. W., & Lau, R. (2008). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational Research Methods, 11*, 296-325.

Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255.

Cleary, T.J., Gubi, A., & Prescott, M.V. (2010). Motivation and self-regulation assessments in urban and suburban schools: Professional practices and needs of school psychologists. *Psychology in the Schools, 47*, 985-1002.

Cleary, T.J., & Kitsantas, A. (2017). Motivation and self-regulated learning influences on middle school mathematics achievement. *School Psychology Review, 46*(1), 88-107.

Craven, R.G., Morin, A.J.S., Tracey, D., Parker, P.D., & Zhong, H.F. (2015). *Inclusive Education for Students with Intellectual Disabilities*. Charlotte, NC: Information Age

Crothers L.M., Hughes, T.L., & Morine, K.A. (2008). *Theory and cases in school-based consultation: A resource for school psychologists, school counselors, special educators, and other mental health professionals*. New York: Routledge.

Crumpton, H.E., & Gregory, A. (2011). "I'm not learning": The role of academic relevanccy for low-achieving students. *The Journal of Educational Research, 104*, 42-53.

Dryer, R., Henning, M.A., Tyson, G.A., & Shaw, R. (2016). Academic achievement performance of university students with disability: Exploring the influence of non-academic factors. *International Journal of Disability, Development and Education, 63*, 419-430.

Eccles, J.S. (2009). Who am I am and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*, 78-89.

Eccles, J.S. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development, 35*, 195-201.

Eccles, J.S., Adler, T., Futterman, R., Goff, S., Kaczala, C., Meece, J., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75-145). San Francisco: Freeman.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109-132.

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

Feather, N. (1982). Expectancy-value approaches: Present status and future directions. In N. Feather (Ed.), *Expectations and actions: Expectancy-value models* (pp. 395-420). Hillsdale, NJ: Erlbaum.

Frielink, N., Schuengel, C., & Embregts, P. (2017). Distinguishing subtypes of extrinsic motivation among people with mild to borderline intellectual disability: Distinguishing subtypes of extrinsic motivation. *Journal of Intellectual Disability Research, 61*, 625-636.

Gilmore, L., & Cuskelly, M. (2009). A longitudinal study of motivation and competence in children with Down syndrome. *Journal of Intellectual Disability Research, 53*, 484-492.

Gottfried, A. E., Fleming, J.S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence. *Journal of Educational Psychology, 93*, 3-13.

Göransson, K., Hellblom-Thibblin, T., & Axdorph, E. (2016). A conceptual approach to teaching maths to students with intellectual disability. *Scandinavian Journal of Educational Research, 60*, 182-200.

Graham, S. (2015). Inaugural editorial for the Journal of Educational Psychology. *Journal of Educational Psychology, 107*, 1-2.

Grigal, M., Neubart, D.A., Moon, S.M., & Graham, S. (2003). Self-determination for students with disabilities: Views of parents and teachers. *Exceptional Children, 70*, 97-112.

Guo, J., Nagengast, B., Marsh, H., Kelava, A., Gaspard, H., Brandt, H., . . . Trautwein, U. (2016). Probing the Unique Contributions of Self-Concept, Task Values, and Their Interactions Using Multiple Value Facets and Multiple Academic Outcomes. *AERA Open, 2*(1).

Guo J., Marsh, H.W., Morin, A.J.S., Parker, P.D., & Kaur, G. (2015). Directionality of the associations of high school expectancy-value, aspirations, and attainment: A longitudinal study. *American Educational Research Journal, 52*, 371-402.

Guo, J., Marsh, H.W., Parker, P., Morin, A.J.S., & Yeung A. (2015). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences, 37*, 161-168.

Guo, J., Marsh, H.W., Parker, P., Morin, A.J.S., & Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in sciences: Dimensional comparison processes and expectancy-by-value interactions. *Learning & Instruction, 49*, 81-91.

Guo, J., Parker, P.D., Marsh, H.W., & Morin, A J.S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology, 51*, 1163-1176.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology, 111*, 745–765.

Heckhausen, H. (1991). Motivation and action. New York: Springer.

Heller, K.A., & Perleth. C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Beltz Test GmbH, Göttingen.

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

Huang, C. (2012). Discriminant and criterion-related validity of achievement goals in predicting academic achievement: A meta-analysis. *Journal of Educational Psychology, 104*, 48-73.

Hulleman, C., & Harackiewicz, J. (2009). Promoting Interest and Performance in High School Science Classes. *Science, 326*, 1410–1412.

Hulleman, C., Kosovich, J., Barron, K., & Daniel, D. (2017). Making connections: Replicating and extending the utility value intervention. *Journal of Educational Psychology, 109*, 387-404.

Inclusion International. (2009). *Better education for all.* Salamanca, Spain: Author.

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 through Twelve. *Child Development, 73*, 509–527.

Katz, I., & Cohen, R. (2014). Assessing autonomous motivation in students with cognitive impairment. *Journal of Intellectual and Developmental Disability, 39*, 323–32.

Korpershoek, H. (2016). Relationships among motivation, commitment, cognitive capacities, and achievement in secondary education. *Frontline Learning Research, 4*, 28-43.

Lau, R. S., & Cheung, G. W. (2012). Estimating and comparing specific mediation effects in complex latent variable models. *Organizational Research Methods, 15*, 3-16.

Lauerman, F., Eccles, J. S., & Pekrun, R. (2017). Why do children worry about their academic achievement? An expectancy-value perspective on elementary students' worry about their mathematics and reading performance. *ZDM Mathematics Education, 49*, 339-354.

Lazowski, R.A., & Hulleman, C.S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research, 86*, 602-640.

Liao, H-A., Ferdenzi, A.C., & Edlin, M. (2012). Motivation, self-regulated learning efficacy, and academic achievement among international and domestic students at an urban community college: A comparison. *Community College Enterprise*, 9-38.

Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods, 4*, 192-211.

Marsh, H. (2007). Do university teachers become more effective with experience?. *Journal of Educational Psychology, 99*, 775-790.

Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M., Morin, A. J. S., Abdelfattah, F., Leung, K. C., Xu, M. K., Nagengast, B., & Parker, P. (2013). Factor structure, discriminant and convergent validity of TIMSS math and science motivation measures: A comparison of USA and Saudi Arabia. *Journal of Educational Psychology, 105*, 108-128.

Marsh, H.W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares, & J. McArdle (Eds.), *Contemporary psychometrics. A Festschrift for Roderick P. McDonald*. Mahwah NJ: Erlbaum.

Marsh, H.W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to cutoff values for fit indexes and dangers in overgeneralizing Hu & Bentler's (1999). *Structural Equation Modeling, 11*, 320-341.

Marsh, H.W., Hau, K.-T., Wen, Z., Nagengast, B., & Morin, A.J.S. (2013). Moderation. In T.D. Little (Ed.), *Oxford Handbook of Quantitative Methods* (pp. 361-386), New York: Oxford University.

Marsh, H. W., Kuyper, H., Seaton, M., Parker, P. D., Morin, A. J. S., Moller, J., et al. (2014). Dimensional comparison theory: an extension of the internal/external frame of reference effect on academic self-concept formation. *Contemporary Educational Psychology, 39*, 326-341.

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

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*, 331–353.

Marsh, H. W., Trautwein, U., Lüdtke, O., Koller, O., & Baumert, J. (2005). Academic self-concept, interest, grades and standardized test scores: reciprocal effects models of causal ordering. *Child Development, 76*, 397-416.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577-605.

Millsap, R.E. (2011). *Statistical approaches to measurement invariance*. New York: Routledge.

Morin, A.J.S., Arens, A.K., Maïano, C., Ciarrochi, J., Tracey, D., Parker, P.D., & Craven, R.G. (2017a). Reciprocal relationships between teacher ratings of internalizing and externalizing behaviors in adolescents with different levels of cognitive abilities. *Journal of Youth and Adolescence, 46*, 801-825.

Morin, A.J.S., Arens, A.K., Tracey, D., Parker, P.D., Ciarrochi, J., Craven, R.G., & Maïano, C. (2017b). Self-esteem trajectories and their social determinants in adolescents with different levels of cognitive ability. *American Journal on Intellectual & Developmental Disabilities, 122*, 539-560.

Morin, A.J.S., Marsh, H.W., & Nagengast, B. (2013). Exploratory Structural Equation Modeling. In G.R. Hancock & R.O. Mueller (Eds.), *Structural Equation Modeling: A Second Course, 2nd Edition* (pp. 395-436). Greewich, CT: IAP.

Murayama, K., Pekrun, R., Suzuki, M., Marsh, H. W., & Lichtenfeld, S. (2016). Don't aim too high for your kids: Parental over-aspiration undermines students' learning in mathematics. *Journal of Personality and Social Psychology, 111*, 166-179.

Muthén, L.K., & Muthén, B.O. (2015). *Mplus User's Guide. 7th Ed*. Los Angeles, CA: Muthén & Muthén.

Nagengast, B., Marsh, H.W., Scalas, L.F., Xu, M,K., Hau, K-T, & Trautwein, U. (2011). Who took the "x" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22*, 1058-1066.

O'Mara, A.J., Marsh, H.W., Craven, R.G., & Debus, R.L. (2006). Do self-concept interventions make a difference? *Educational Psychologist, 41*, 181–206.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. Contemporary Educational Psychology, 20, 426-443.

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*, 193-203.

Pekrun, R. (1988). *Emotion, Motivation und Persönlichkeit* [Emotion, motivation and personality]. München/Weinheim: Psychologie Verlags Union.

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. (1993). Facets of students' academic motivation: A longitudinal expectancy-value approach. In M. Maehr & P. Pintrich (Eds.), *Advances in Motivation and Achievement* (Vol. 8, pp. 139-189). Greenwich, CT: JAI Press.

Pekrun, R., Hall, N.C., Goetz, T., & Perry, R.P. (2014). Boredom and academic achievement: Testing a model of reciporcal causation. *Journal of Educational Psychology, 106*, 696-710.

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance. *Child Development, 88,* 1653-1670.

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.

Pintrich, P.R. (2000). An achievement goal theory perspective on issues in motivational terminology, theory and research. *Contemporary Educational Psychology, 25,* 92-104.

Pintrich, P.R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385-407.

Raykov, T., & Marcoulides, G. A. (2004). Using the delta method for approximate interval estimation of parameter functions in SEM. *Structural Equation Modeling, 11*, 621-637.

Rotgans, J.I., & Schmidt, H.G. (2012). The intricate relationship between motivation and achievement: Examining the mediating role of self-regulated learning and achievement-related classroom behaviors. *International Journal of Teaching and Learning in Higher Education, 24*, 197-208.

Schunk, D.H., & Zimmerman, B. (2003). Self-regulation and learning. In W. Reynolds, & G. Miller (Eds.). *Handbook of psychology, Vol. 7, Educational Psychology* (pp. 59-78). New Jersey: Wiley.

Sekhon, J.S. (2011). Multivariate and propensity score matching software with automated balance optimization: The Matching package for R. *Journal of Statistical Software, 42*, 1-52.

Shurr, J., & Bouck, E. (2013). Research on curriculum for students with moderate and severe intellectual disability. *Education and Training in Autism and Developmental Disabilities, 48,* 76–87.

Streiner, D.L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment, 80*, 99-103.

Taylor, G., Jungert, T., & Mageau, G.A., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: The unique role of intrinsic motivation. *Contemporary Educational Psychology, 39,* 342–358.

Tracey, D., Craven, R. G. & Marsh, H.W. (2015). The next frontier – participatory action research to explicate the self-concepts of children with intellectual disabilities. In R. G., Craven, A. Morin, P. Parker, & D. Tracey (Eds.), *Inclusive education for students with intellectual disabilities* (pp. 211-227). Charlotte, NC: Information Age.

Trautwein, U., Marsh. H.W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy-value theory: A latent interaction modelling study. *Journal of Educational Psychology, 104*, 763-777.

Watts, T., Duncan, G., Chen, M., Claessens, A., Davis-Kean, P., Duckworth, K., . . . Susperreguy, M. (2015). The Role of Mediators in the Development of Longitudinal Mathematics Achievement Associations. *Child Development, 86*(6), 1892-1907.

Walberg, H. (1984). Improving the productivity of American schools. *Educational Leadership, 41*, 19–30.

Wigfield, A., & Eccles, J.S. (2000). Expectancy-value theory of motivation. *Contemporary Educational Psychology, 25*, 68-81.

Wigfield, A., & Eccles, J.S. (2002). The development of competence beliefs and values from childhood through adolescence. In A. Wigfield & J.S. Eccles (Eds.), *Development of achievement motivation* (pp. 92-120). San Diego, CA: Academic Press.

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. *Handbook of child psychology and developmental science*, 1-44

Zisimopoulos, D.A. & Galanaki, E.P. (2009). Academic intrinsic motivation and perceived academic competence in Greek elementary students with and without learning disabilities. *Learning Disabilities Research and Practice, 24*, 33-43.

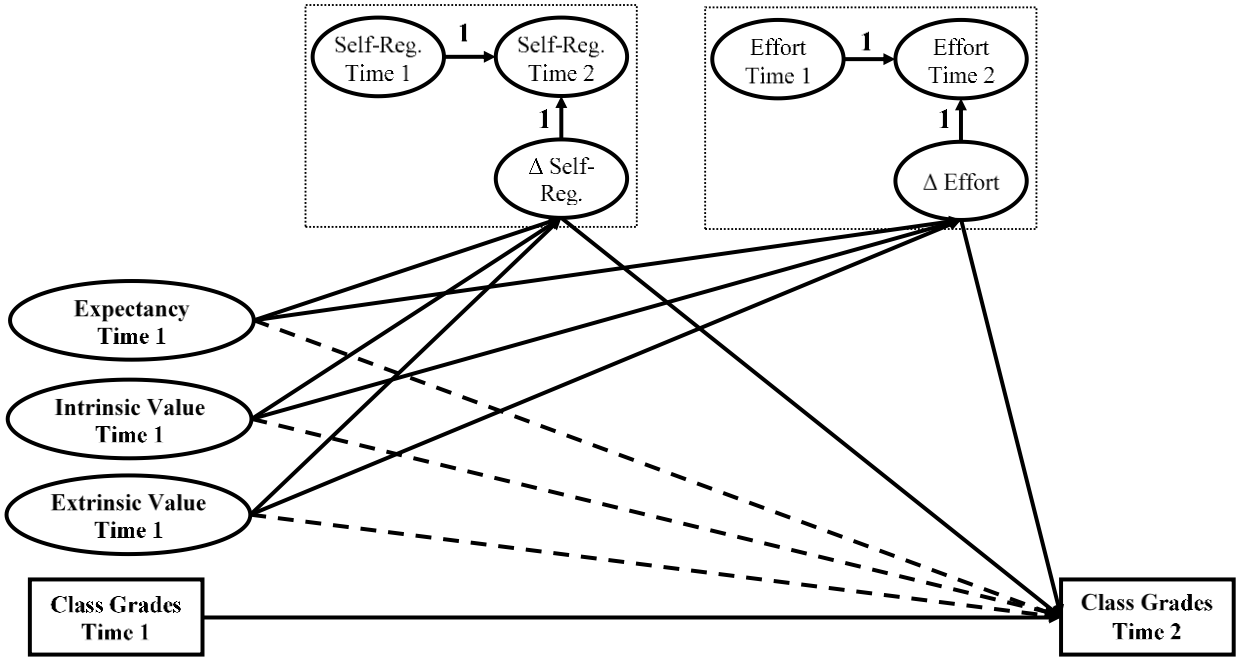

Figure 1. A Priori Longitudinal Mediation Model.
Note. Ovals represent latent variables, rectangles represent observed manifest variables, full arrows represent the a priori mediational paths, dashed arrows represent paths added for tests of partial mediation, specification of the latent change (Δ) variables are encapsulated in dotted squares, the measurement part of the latent variables, as well as the included correlations (between predictors, between latent change factors, and between latent change factors and their Grade 5 components) are left out for greater clarity.

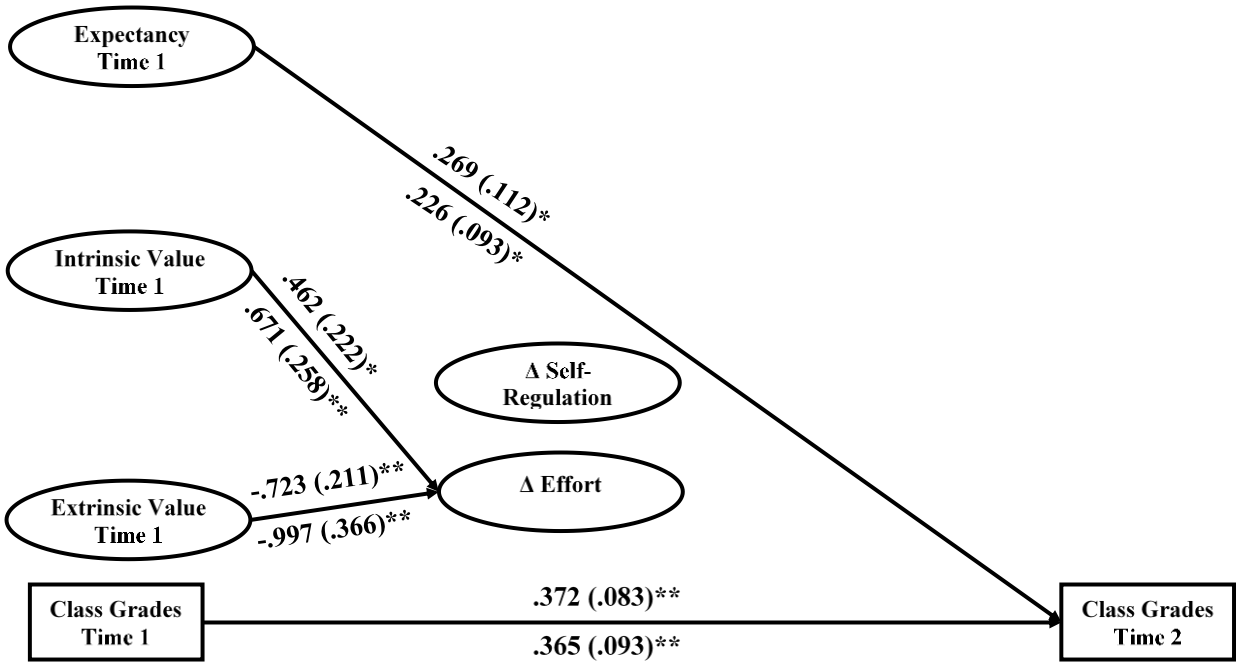

Figure 2. Result from the Final Longitudinal Mediation Model
Note. The standardized coefficients (β) and their standard errors (in parentheses) obtained for participants with low cognitive ability are reported above the arrows, and those obtained for participants with average to high levels of cognitive abilities are reported under the arrows; ovals represent latent variables, rectangles represent observed manifest variables, full arrows represent the final statistically significant predictive paths (* $p \leq .05$; ** $p \leq .01$), the measurement part of the latent variables, the specification of the latent change (Δ) factors, the non-significant paths, and the included correlations (between predictors, between latent change factors, and between latent change factors and their Grade 5 components) are left out for greater clarity.

Table 1
Goodness-Of-Fit Statistics for the Measurement Invariance Models

	χ^2	df	CFI	TLI	RMSEA	CI	$\Delta\chi^2$	Δdf	ΔCFI	ΔTLI	ΔRMSEA
Measurement Invariance across Groups: Time 1 (Grade 5)											
Configural invariance	232.862*	148	.933	.917	.068	.051; .085	---	---	---	---	---
Weak Invariance	248.399*	159	.929	.919	.068	.051; .084	15.142	11	-.004	.002	.000
Weak Invariance with ETEC	249.451*	160	.929	.919	.068	.051; .083	1.012	1	.000	.000	.000
Strong Invariance (ETEC)	272.767*	171	.919	.914	.070	.054; .085	24.430	11	-.010	-.005	.002
Strict Invariance (ETEC)	320.983*	185	.892	.894	.077	.063; .091	47.152*	14	-.027	-.020	.007
Partial Strict Invariance (ETEC)	292.550*	181	.911	.911	.071	.056; .086	19.566	10	-.008	-.003	.001
Latent Variance-Covariance Invariance (ETEC; P. Strict)	302.960*	187	.908	.910	.071	.056; .086	10.542	6	-.003	-.001	.000
Latent Means Invariance (ETEC; P. Strict)	303.978*	190	.909	.913	.070	.055; .084	.619	3	.001	.003	-.001
Measurement Invariance across Groups: Time 2 (Grade 8)											
Configural invariance	238.324*	148	.926	.909	.077	.058; .095	---	---	---	---	---
Weak Invariance	242.801*	159	.931	.921	.072	.053; .089	5.481	11	.005	.012	-.005
Weak Invariance with ETEC	242.811*	160	.932	.923	.071	.052; .088	.122	1	.001	.002	-.001
Strong Invariance (ETEC)	252.135*	171	.933	.929	.068	.049; .085	8.772	11	.001	.006	-.003
Strict Invariance (ETEC)	292.525*	185	.912	.913	.075	.058; .091	38.021*	14	-.021	-.016	.007
Partial Strict Invariance (ETEC)	276.738*	183	.923	.924	.071	.053; .087	23.713	12	-.010	-.005	.003
Latent Variance-Covariance Invariance (ETEC; P. Strict)	290.861*	189	.916	.920	.072	.055; .088	15.025	6	-.007	-.004	.001
Latent Means Invariance (ETEC; P. Strict)	302.481*	192	.909	.914	.075	.058; .090	12.780*	3	-.007	-.006	.003
Longitudinal Invariance (Total Sample)											
Configural invariance	506.325*	321	.930	.917	.042	.035; .048	---	---	---	---	---
Weak Invariance	531.214*	332	.925	.914	.042	.036; .049	26.030*	11	-.005	-.003	.000
Weak Invariance with ETEC	531.303*	333	.925	.915	.042	.035; .049	.043	1	.000	.001	.000
Strong Invariance (ETEC)	557.134*	344	.919	.911	.043	.036; .050	26.052*	11	-.006	-.004	.001
Strict Invariance (ETEC)	572.803*	358	.919	.914	.042	.036; .049	16.733	14	.000	.003	-.001
Latent Variance-Covariance Invariance (ETEC)	592.679*	364	.914	.910	.043	.037; .050	21.223*	6	-.005	-.004	.001
Latent Means Invariance (ETEC)	678.984*	367	.882	.879	.051	.045; .056	110.561*	3	-.032	-.031	.008

Note. * $p \leq .01$; All models are estimated with the Robust Maximum Likelihood (MLR) estimator; χ^2 = robust chi square test of exact fit; df = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; CI = 90% confidence interval for the RMSEA; Δ = change in model fit relative to the previously retained model; $\Delta\chi^2$ were calculated while taking into account the robust scaling correction factors (calculated from models loglikelihoods for greater precision).

Table 2
Results from the Predictive Models

	Effort Time 1			Self-Regulation Time 1			Class Grades Time 1		
Predictors	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)
Time 1									
Expectancy Time 1	-.061 (.114)	-.059 (.111)	-.078 (.146)	.095 (.101)	.107 (.114)	.123 (.134)	.617 (.153)**	.522 (.120)**	.513 (.131)**
Intrinsic Value Time 1	-.020 (.092)	-.023 (.107)	-.033 (.155)	.085 (.095)	.116 (.132)	.146 (.159)	.111 (.144)	.112 (.147)	.121 (.157)
Extrinsic Value Time 1	.850 (.001)**	.932 (.060)**	.934 (.082)**	.574 (.086)**	.735 (.101)**	.642 (.112)**	-.444 (.192)*	-.424 (.172)*	-.317 (.135)*
	Effort Time 2			Self-Regulation Time 2			Class Grades Time 2		
EVT Predictors	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)
Time 2									
Expectancy Time 2	.066 (.100)	.074 (.112)	.090 (.134)	.236 (.090)**	.297 (.120)*	.329 (.119)**	.686 (.101)**	.750 (.114)**	.667 (.093)**
Intrinsic Value Time 2	.175 (.134)	.201 (.156)	.217 (.164)	-.105 (.109)	-.137 (.141)	-.135 (.139)	-.230 (.143)	-.261 (.165)	-.206 (.126)
Extrinsic Value Time 2	.344 (.167)*	.389 (.193)*	.314 (.157)*	.524 (.133)**	.673 (.140)**	.494 (.130)**	.336 (.227)	.374 (.268)	.221 (.127)
	Effort Time 2			Self-Regulation Time 2			Class Grades Time 2		
EVT Predictors	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)
Longitudinal Prediction									
Expectancy Time 1	.080 (.155)	.084 (.165)	.086 (.167)	.327 (.160)*	.425 (.214)*	.353 (.168)*	.474 (.185)**	.483 (.183)**	.372 (.142)**
Intrinsic Value Time 1	.205 (.188)	.257 (.241)	.291 (.265)	-.046 (.199)	-.070 (.305)	-.065 (.286)	.139 (.209)	.168 (.254)	.145 (.216)
Extrinsic Value Time 1	.067 (.265)	.066 (.263)	.069 (.274)	.236 (.268)	.284 (.323)	.247 (.274)	-.367 (.313)	-.348 (.282)	-.279 (.231)
	ΔEffort			ΔSelf-Regulation			Class Grades Time 2		
EVT Predictors	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)	*b* (SE)	$\beta_{L\text{-}CA}$ (SE)	$\beta_{A/H\text{-}CA}$ (SE)
Longitudinal Mediation									
Expectancy	.004 (.242)	.003 (.185)	.004 (.246)	.145 (.213)	.229 (.346)	.161 (.233)	.281 (.117)*	.269 (.112)*	.226 (.093)*
Intrinsic Value	.502 (.242)*	.462 (.222)*	.671 (.258)**	.017 (.227)	.032 (.429)	.024 (.334)	--	--	--
Extrinsic Value	-1.051 (.449)**	-.723 (.211)**	-.997 (.366)**	-.515 (.328)	-.735 (.386)	-.535 (.310)	--	--	--
Class Grades (Time 1)	--	--	--	--	--	--	.358 (.086)**	.372 (.083)**	.365 (.093)**
ΔEffort	--	--	--	--	--	--	.202 (.353)	.253 (.436)	.160 (.281)
ΔSelf-Regulation	--	--	--	--	--	--	-.004 (.497)	-.002 (.300)	-.003 (.360)

Note. * $p \leq .05$; ** $p \leq .01$; the invariance of the predictive paths refers to the unstandardized coefficients (which are thus equivalent across groups); the standardized coefficients are also a function of the level of within-group variability and thus differ slightly across samples; *b* = unstandardized regression coefficient (invariant across groups); $\beta_{L\text{-}CA}$ = standardized regression coefficient obtained in the group of students with low levels of cognitive ability; $\beta_{A/H\text{-}CA}$ = standardized regression coefficient obtained in the group of students with average to high levels of cognitive ability; SE = standard error of the coefficient; Δ = latent change factor; -- = path not estimated in the final model.

Appendix
Correlations and Descriptive Statistics

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Expectancy T1		.585**	.330**	.224*	.357**	.386**	.512**	.450**	.311*	.518**	.550**	.495**
2. Intrinsic T1	.638**		.311**	.370**	.505**	.323**	.286	.425**	.520**	.585**	.282	.528**
3. Extrinsic T1	.342**	.516**		.633**	.567**	.097	-.180	-.120	-.086	.442**	.432**	.133
4. Effort T1	.377**	.454**	.480**		.732**	-.084	-.176	-.162	.171	.171	.365**	.139
5. Self-Reg. T1	.467**	.466**	.413**	.686**		.121	-.063	.012	.201	.224*	.386**	.145
6. Class Grades T1	.461**	.212*	-.044	.224**	.342**		.130	.103	.151	.198	.183	.410**
7. Expectancy T2	.527**	.304**	.230*	.265**	.311**	.389**		.716**	.213*	.444**	.337**	.601**
8. Intrinsic T2	.269**	.395**	.246**	.078	.073	-.070	.515**		.321*	.418**	.307*	.529**
9. Extrinsic T2	.097	.296**	.325**	.168	.140	-.078	.454**	.489**		.477**	.409**	.494**
10. Effort T2	.172	.269**	.127**	.234*	.268**	.036	.305**	.394**	.276**		.599**	.500**
11. Self-Reg. T2	.252*	.279**	.202*	.326**	.305**	.166	.451**	.329**	.388**	.681**		
12. Class Grades T2	.386**	.161**	.102	.231**	.217*	.488**	.671**	.304**	.175*	.212*	.236**	.390**
Mean L-CA	3.410	3.191	3.682	3.733	3.473	3.072	2.839	2.390	3.156	3.261	3.162	2.973
Mean A/H-CA	3.450	3.226	3.751	3.783	3.529	3.907	3.135	2.328	3.282	3.391	3.303	3.706
Variance L-CA	.818	1.102	1.326	.924	.769	.808	.933	1.018	1.441	.865	.680	.760
Variance A/H-CA	.665	1.126	1.008	.473	.537	.906	.969	.858	.891	.602	.627	.917

Note. * $p \leq .05$; ** $p \leq .01$.The results for the students with low levels of cognitive ability (L-CA) are reported above the diagonal, whereas those for the students with average to high levels of cognitive ability (A/H-CA) are reported under the diagonal.