Check for updates

CHILD DEVELOPMENT

Child Development, March/April 2018, Volume 89, Number 2, Pages e138–e156

Changes in the Relation Between Competence Beliefs and Achievement in Math Across Elementary School Years

Anne F. Weidinger and Ricarda Steinmayr
Technical University Dortmund

Birgit Spinath
Heidelberg University

Math competence beliefs and achievement are important outcomes of school-based learning. Previous studies yielded inconsistent results on whether skill development, self-enhancement, or reciprocal effects account for the interplay among them. A development-related change in the direction of their relation in the early school years might explain the inconsistency. To test this, 542 German elementary school students (M = 7.95 years, SD = 0.58) were repeatedly investigated over 24 months from Grade 2 to Grade 4. Math competence beliefs declined and had a growing influence on subsequent math grades. This suggests changes in the dominant direction of the relation from a skill development to a reciprocal effects model during elementary school. Findings are discussed with regard to their theoretical and practical implications.

Students' competence beliefs are not only an important prerequisite of school achievement but also an important end of education in itself. To believe in one's competence is associated with more positive valuing of a domain (e.g., Wigfield et al., 2015) and an increased likelihood that a person will further engage in corresponding activities (e.g., Musu-Gillette, Wigfield, Harring, & Eccles, 2015; Nagy, Trautwein, Baumert, Köller, & Garrett, 2006). Even in elementary school, students have differentiated competence beliefs for different domains like math, reading, and sports (e.g., Eccles, Wigfield, Harold, & Blumenfeld, 1993). Recent studies provided evidence that the development of competence beliefs in math and of other motivation constructs in the math domain early in school is predictive of achievement and choices of math-intensive careers in adolescence and beyond (Ceci, Ginther, Kahn, & Williams, 2014; Musu-Gillette et al., 2015). Given that there is still a need for professionals in the fields of science, technology, engineering, and mathematics (STEM) and that degrees in math-intensive fields are associated with higher status careers (Ma & Johnson, 2008), it is crucial to understand the emergence and change in early math competence beliefs in order to find ways to positively influence their development.

Achievement is considered as a very important factor influencing the emergence and change in competence beliefs because achievement comparisons have been argued to constitute the main determinant of competence beliefs (Skaalvik & Skaalvik, 2002). Students may use different sources of achievement information for these comparisons (see Skaalvik & Skaalvik, 2002). Especially for children at the beginning of their school careers, achievement feedback provided by teachers is the most important source of information about their competence (e.g., Spinath & Spinath, 2005a). Although grades might not be the most valid assessments of students' actual ability, they are the most visible, universally present form of feedback for students and therefore play a central role in the early formation of competence beliefs (see Niepel, Brunner, & Preckel, 2014a). Besides the fact that achievement influences competence beliefs (*skill development effects*), it is well known that competence beliefs in turn influence achievement (*self-enhancement effects*; for a review, see Marsh & Martin, 2011). The majority of studies investigated students beyond elementary school and found reciprocal effects between competence beliefs and achievement in different domains, inter alia math (e.g., Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Niepel, Brunner, & Preckel, 2014b). By contrast, we are aware of only three studies that have

We would like to thank Verena Freiberger for her support in data collection and data preparation in MEGA.

Correspondence concerning this article should be addressed to Anne F. Weidinger, Department of Psychology, Technical University Dortmund, Emil-Figge-Straße 50, 44227 Dortmund, Germany. Electronic mail may be sent to anne.weidinger@tu-dortmund.de.

© 2017 The Authors
Child Development © 2017 Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/2018/8902-0033
DOI: 10.1111/cdev.12806

investigated the causal interplay between math achievement and math competence beliefs in elementary school, and their results are inconsistent for this age group. Two studies only found skill development effects (Skaalvik & Valas, 1999; Viljaranta, Tolvanen, Aunola, & Nurmi, 2014), whereas one also revealed self-enhancement effects (Helmke & van Aken, 1995). We will review this work in more detail below. The inconsistent findings might be explained by a change in the dominant direction of the relation from skill development to reciprocal effects in the elementary school years (see Wigfield & Karpathian, 1991). Given the importance of the development of math competence in elementary school for future educational pathways, it is necessary to capture the early interplay of competence beliefs and achievement in the math domain. The present study is the first that considers children's math competence beliefs and their grades several times within a school year to investigate the hypothesis that the dominant direction of the relation between the constructs changes from children's 8th to their 10th year of life.

Competence Beliefs, Academic Achievement, and Their Relation

Competence beliefs as defined in the present study refer to a person's sense of how good they are at a given activity. This kind of competence belief is frequently investigated under the label of ability self-concept or self-perceived abilities (e.g., Harter, 1982; Herbert & Stipek, 2005; Marsh, 1990; Wigfield et al., 2015) and is typically measured by items such as "How good are you at (*domain or task*)?" (see Wigfield et al., 1997). To evaluate the quality of one's achievement and generate competence beliefs, students use internal and external achievement comparisons (i.e., *frame of references*; Skaalvik & Skaalvik, 2002). In school, the most important source of information that is used for achievement comparisons is the achievement feedback provided by teachers (e.g., Spinath & Spinath, 2005a).

School achievement is most often measured via grades or standardized achievement tests. Because in schools children typically get achievement feedback via grades but not via standardized achievement tests, studies concerned with the effects of feedback like the present one operationalize achievement by grades (see Niepel et al., 2014a). Empirically, closer relations between competence beliefs and achievement are found when achievement is operationalized in terms of grades compared to objective test measures (e.g., Marsh et al., 2005; Möller, Pohlmann, Köller, & Marsh, 2009; Valentine, DuBois, & Cooper, 2004). Although there are many ways in which teachers provide competence-related feedback, grades are more standardized than other forms of feedback like verbal feedback during class; they are omnipresent; and they are readily available for investigation.

Positive associations between students' competence beliefs and their academic achievement have been documented in a large body of research (e.g., Huang, 2011; Möller et al., 2009; Skaalvik & Skaalvik, 2008). The strength of association between competence beliefs and achievement depends, among others, on children's age. Self-concepts of preschool and early elementary school children are on average very positive and only weakly, if at all, associated with external indicators such as grades (e.g., Marsh, 1990; Spinath & Spinath, 2005a) or parents' perceptions of children's ability (e.g., Spinath & Spinath, 2005a; see Stipek & Mac Iver, 1989). When growing older, children's self-evaluations become more negative, a development that is especially pronounced in the math domain (Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Spinath & Spinath, 2005b). At the same time, the positive association between competence beliefs and external achievement criteria becomes increasingly stronger (e.g., Marsh, 1990; Spinath & Spinath, 2005a). This increasingly close association might be explained by either children adapting their competence beliefs to achievement feedback or, opposed to this, children's competence beliefs influencing subsequent achievement or both. An unresolved question concerns the timing of the two processes. Because of the aforementioned change in children's competence beliefs in elementary school as well as the important role of the early development of math competence beliefs for future educational and career choices (Musu-Gillette et al., 2015), elementary school age is a very important stage to investigate.

Direction of the Relation Between Competence Beliefs and Academic Achievement

Since the 1970s, two mechanisms describing the interplay among competence beliefs and achievement have been contrasted (Calysn & Kenny, 1977). According to the *skill development approach*, achievement is predominant over competence beliefs, implying that competence beliefs emerge as a consequence of prior achievement. Contrary to this, the *self-enhancement approach* holds that competence beliefs are predominant over achievement, meaning

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

that changes in competence beliefs will lead to changes in academic achievement. Besides these two unidirectional approaches, a third possibility is a *reciprocal effects model* (REM) in which prior achievement affects subsequent competence beliefs and, at the same time, prior competence beliefs affect subsequent achievement (e.g., Marsh, 1990). To investigate potential causal relations between competence beliefs and academic achievement, specific longitudinal study designs are needed. According to Marsh, Byrne, and Yeung (1999), constructs should be measured at least twice but preferably more frequently, data analysis should be based on structural equation modeling, and concepts should ideally be measured by means of three or more indicators. Finally, a large and diverse sample is essential to justify the use of structural equation models. In the following overview of findings, only studies that comply with these guidelines are considered.

For children beyond elementary school, that is, 10 years and older, a model with reciprocal effects seems to describe the relation between competence beliefs and achievement in math and other domains best (e.g., Marsh et al., 2005; Niepel et al., 2014b; Pinxten, de Fraine, van Damme, & D'Haenens, 2010; for review, see Marsh & Martin, 2011). With children in elementary school (i.e., up to 10 years of age), however, the situation is less clear. Table 1 summarized the studies that investigated the direction of the relation between competence beliefs and achievement in elementary school children (Chapman & Tunmer, 1997; Guay, Marsh, & Boivin, 2003; Helmke & van Aken, 1995; Kurtz-Costes & Schneider, 1994; Marsh, Gerlach, Trautwein, Lüdtke, & Brettschneider, 2007; Skaalvik & Hagtvet, 1990; Skaalvik & Valas, 1999; Viljaranta et al., 2014). Among the eight studies, only three examined the relation in the math domain. Viljaranta et al. (2014) investigated students from Central Finland in Grades 1 (t_1), 2 (t_2), 4 (t_3), and 7 (t_4). At an interval of 1, 2, and 3 years, respectively, they assessed students' math competence beliefs and math achievement (t_1–t_3: teacher ratings, t_4: math grades). Their results support only the skill development model. Helmke and van Aken (1995) assessed math competence beliefs and math achievement (grades and achievement test scores) three times at an interval of 1 year in Grades 2, 3, and 4 in Germany. Contrary to Viljaranta et al. (2014), they found evidence for the REM if grades and test scores were taken as separate indicators of achievement, whereby effects of prior competence beliefs on subsequent achievement were only small. If grades and test scores were aggregated, Helmke and van Aken (1995) found only skill development effects. By contrast, results of Skaalvik and Valas (1999) support the skill development model (SDM) regardless of whether achievement was indicated by both teacher ratings and test scores or whether teacher ratings and test scores were taken as separate indicators of achievement. They investigated students from Grades 3 to 4 (Cohort 3), from Grades 6 to 7 (Cohort 2), and from Grades 8 to 9 (Cohort 1), and assessed math competence beliefs and math achievement twice at an interval of 1 year in each cohort. Taken together, only Helmke and van Aken (1995) found evidence for reciprocal effects in elementary school when investigating relations in the math domain. Findings of studies that investigated the relations in the verbal domain in elementary school are in favor of the SDM (Chapman & Tunmer, 1997; Skaalvik & Valas, 1999; Viljaranta et al., 2014). By contrast, some of the previous studies on elementary school students that assessed competence beliefs concerning school in general (e.g., "do well at class work"; Guay et al., 2003) or sports (Marsh et al., 2007) instead of domain specific for the math or verbal domain found effects of prior competence beliefs on subsequent achievement.

All in all, results for elementary school children are inconsistent and do not conform to the results for older students. One explanation for the heterogeneity of results might be that the research on the relation between competence beliefs and achievement in elementary school was done in many different countries with different educational practices that might impact competence beliefs. For instance, ability tracking varies in its timing and scope between countries (see Wigfield et al., 2015). Another explanation might be the developmental change in domain-specific competence beliefs for academic domains like math that takes place in elementary school.

Developmental and Contextual Changes in Elementary School

It has been argued that the relation among domain-specific competence beliefs and achievement changes during the first years in school (e.g., Skaalvik & Hagtvet, 1990; Skaalvik & Valas, 1999; Wigfield & Karpathian, 1991). This development is interpreted in terms of age-related changes in children's cognitive abilities and in their educational environment (Dweck, 2002; Nicholls, 1978; Stipek & Mac Iver, 1989). Between 7 and 8 years of age, children's competence beliefs become more accurate

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Overview of Previous Studies Investigating the Longitudinal Relation of Competence Beliefs and Academic Achievement in Math and Other Domains in Elementary School (Ordered According to the Youngest Cohort Investigated)

Authors (year): country	Number of cohorts (sample size): grades	Number of waves; time intervals	Domains	Achievement measure	Results in favor of: skill development model (SDM) self-enhancement model (SEM) reciprocal effects model (REM)	Remark
Chapman and Tunmer (1997): New Zealand	One cohort (N = 122): 1, 2, 3	3; 1.5 years (t_1–t_2) and 1 year (t_2–t_3)	Reading	Different standardized reading performance test scores	No significant cross-lagged paths between Grade 1 and Grade 2; SDM between Grade 2 and Grade 3	On average, children were not older than 8 years at the end of the study
Viljaranta et al. (2014): Central Finland	One cohort (N = 216): 1, 2, 4, 7	4; 1 year (t_1–t_2), 2 years (t_2–t_3), and 3 years (t_3–t_4)	**Math** and reading (t_1–t_3) or the Finnish language (t_4)	t_1–t_3: teacher ratings of math and reading performance t_4: grades in math and in the Finnish language	Math: SDM from t_1 to t_4 Reading/Finnish language: SDM from t_1 to t_2 and from t_3 to t_4; no significant cross-lagged path from t_2 to t_3	Domain-specific interest was incorporated into the models
Kurtz-Costes and Schneider (1994): Munich, Germany	One cohort (N = 46): 2, 4	2; 2 years	Mean of different domains	Mean of grades in math and German	REM	Relatively long distance between waves and small sample size
Guay et al. (2003): Quebec City, Canada	Cohort 3 (n = 125): 2, 3, 4 Cohort 2 (n = 147): 3, 4, 5 Cohort 1 (n = 113): 4, 5, 6	3; 1 year	School in general	Teacher ratings for reading, writing and math (aggregated)	REM from t_1 to t_2 SEM from t_2 to t_3 (consistent in three cohorts)	
Helmke and van Aken (1995): Urban and rural regions in and around Munich, Germany	One cohort (N = 697): 2, 3, 4	3; 1 year	**Math**	Grades in math and math achievement test scores	SDM if grades and test scores were aggregated; REM if grades and test scores were taken as separate indicators of achievement	
Marsh et al. (2007): Paderborn, Germany	One cohort (N = 1,135): 3, 4, 6	3; ca. 1.3 years	Sport	t_1: teacher ratings of sport performance t_2–t_3: grades in sport	REM	No academic domain in the narrow sense

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Continued

Authors (year): country	Number of cohorts (sample size): grades	Number of waves; time intervals	Domains	Achievement measure	Results in favor of: skill development model (SDM) self-enhancement model (SEM) reciprocal effects model (REM)	Remark
Skaalvik and Hagtvet (1990): Trondheim, Norway	Cohort 2 (n = 271): 3, 4 Cohort 1 (n = 364): 6, 7	2; 1.5 years	School in general	Teacher ratings of general school achievement	SDM from Grade 3 to Grade 4; REM for Grade 6 to Grade 7	General self-esteem was incorporated into the models
Skaalvik and Valas (1999): Norway	Cohort 3 (n = 494): 3, 4 Cohort 2 (n = 285): 6, 7 Cohort 1 (n = 226): 8, 9	2; 1 year	**Math** and language	Teacher ratings of performance and achievement test scores in math and language	SDM if achievement was indicated by teacher ratings and test scores, and if teacher ratings and test scores were taken as separate indicators (in all three cohorts and in both domains)	Domain-specific interest and investment were incorporated into the models

Note. Studies printed in bold investigated the domain of math. t = measurement occasion.

and realistic (see Dweck, 2002), and thus less positive (e.g., Archambault, Eccles, & Vida, 2010; Spinath & Spinath, 2005b; Spinath & Steinmayr, 2008). This is a result of children's growing capacity to differentiate between ability, effort, and other factors influencing achievement and to use information for inferences about ability. Although the majority of young children does not differentiate ability from other characteristics of a person (e.g., personality, physical appearance; see Dweck, 2002; Harter, 2006), 5- to 7-year-olds are starting to define ability as its own domain so that it can be clearly distinguished from social qualities, and different subdomains of ability (e.g., math, reading, sports) can be differentiated (e.g., Marsh, Craven, & Debus, 1991, 1998; Wigfield et al., 1997). At about 7 to 8 years of age, children are increasingly able to draw inferences about their own and others' ability from failure and success feedback, and to integrate ability-related information from different sources (e.g., task difficulty, effort, temporal and social comparison). Their competence beliefs move into line with external evaluations, for example, teachers perceptions of children's ability (see Dweck, 2002; e.g., Spinath & Spinath, 2005a). These trends continue throughout elementary school so that the majority of 10-year-olds is able to make inferences about their ability from effort and performance experiences (e.g., Ruble, Parsons, & Ross, 1976) and defines ability as a relatively stable trait rather than a set of skills (e.g., Nicholls, 1978; Stipek & Mac Iver, 1989). Students progressively use their ability as an explanation for academic achievement, which is perceived as being constrained by current ability (see Dweck, 2002). Thus, at the end of elementary school (9–10 years), predictions of future success should increasingly base on current competence beliefs.

At the same time, important changes in educational contexts take place: There is an increase in normative and critical evaluations from significant others (e.g., teachers) that leads to a growing concern with relative performance (see Dweck, 2002; Harter, 2006), makes normative achievement comparisons more likely (e.g., Spinath & Spinath, 2005a), and thus results in self-evaluations that contain more self-criticism and are less positive (e.g., Frey & Ruble, 1987; Jacobs et al., 2002). The extent to which these contextual changes affect students' self-evaluations should vary between schools from different countries depending on the educational systems and practices (Wigfield et al., 2015). The German school system is characterized by a relatively early ability tracking, usually after Grade 4.

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Students in Germany receive recommendations for secondary school on the basis of the grades they gain in the first term of Grade 4. Only graduation from some secondary school types (i.e., Gymnasium or a school covering courses on Gymnasium level) enables one to enroll at a university right after school. As there is little movement across tracks once the decision for a specific secondary school type has been made, the recommendation for secondary track is of utmost importance for the students, and in Germany, external evaluations should be particularly salient at the end of elementary school.

Considering these developmental and contextual changes in elementary school, prior achievement experiences, like grades, should have predominance over children's math competence beliefs between the ages of 7 and 8 (i.e., until the middle of Grade 3 in German elementary school) because children increasingly incorporate external performance feedback into their self-evaluations at this time. Between the age of 8 and 10 (i.e., at the end of elementary school), the influence of math competence beliefs on math achievement (i.e., grades) should increase because students at this age should have established competence beliefs for the math domain. Students with realistically high domain-specific competence beliefs should get more involved in domain-specific tasks (e.g., spending more time and effort) than students with lower competence beliefs so that the former should develop better domain-specific competencies than the latter (see Nicholls, 1978).

To test this, competence beliefs and achievement need to be assessed more than twice between the ages of 8 and 10 which was the case in only three of the presented studies (Guay et al., 2003; Helmke & van Aken, 1995; Viljaranta et al., 2014). Results of these studies do not confirm the developmental hypothesis that the influence of competence beliefs on subsequent achievement increases with age. This could be due to the relatively longtime interval between the measurement occasions ($\geq$ 1 year) that might have masked the developmental dynamics between competence beliefs and achievement (see Helmke & van Aken, 1995; Viljaranta et al., 2014). Shorter intervals between measurement occasions might help to capture the developmental phenomena in more detail. The relation between competence beliefs and grades might differ not only between school years but also within a school year as students receive information about their achievement (e.g., grades) throughout the school year. Thus, it is crucial to assess both constructs several times within school years. From a developmental perspective it is reasonable to assess competence beliefs and grades at the beginning, in the middle, and at the end of the school year because achievement feedback accumulates throughout a school year and feedback like grades should become increasingly important the closer the end-year report cards get. There is no study that systematically investigates the direction of the relation between competence beliefs and achievement in elementary school in this developmental time period. An integrative work is crucial for a better understanding of the interplay of self-perceptions and competence development in the critical time period when school-related competence beliefs are established, and children are expected to develop positive beliefs about their abilities (see Chapman & Tunmer, 1997; Guay et al., 2003).

The Present Study

Empirical evidence of potential change in the relation between competence beliefs and achievement in math across the elementary school years is scarce. In particular, there are no studies that investigate elementary school students more than twice within a school year. Therefore, hitherto little is known about changes in the direction of the relation between math competence beliefs and academic achievement in the period of education when competence beliefs are established. With the aim to broaden the knowledge about the longitudinal interplay of math competence beliefs and achievement, our study should fill in this gap. We investigated elementary school students at seven points in time between their 8th and 10th year of life. To capture the expected directive relations between competence beliefs and achievement in math in detail, math-related competence beliefs and math grades were assessed at the beginning, in the middle, and at the end of each school year. The following research questions were addressed:

1. How do students' math competence beliefs change from the end of Grade 2 to the end of Grade 4? With respect to previous research and theoretical considerations, we expect a significant decline in students' competence beliefs over the investigated time period.
2. Does the dominant direction of the relation between students' math competence beliefs and their math grades change from the 8th to the 10th year of life (i.e., from the end of Grade 2 to the end of Grade 4)? With respect to

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

theoretical considerations, we expect that at first achievement will have predominance over competence beliefs so that skill development effects will prevail. As competence beliefs get increasing influence on subsequent achievement (self-enhancement effect), the dominant direction of the relation will change. Besides skill development effects, self-enhancement effects will occur.

Method

Participants and Procedure

Data come from a longitudinal project on the *development of motivation during elementary school* (MEGA, *M*otivations*E*ntwicklung im *G*rund-schul*A*lter). A total of 542 second graders (270 boys and 272 girls) from 27 classes from 11 elementary schools in and around two midsized towns in southern Germany participated in this project. Of the children, 76.7% indicated German as their native language. Comparing the present data with those from the Federal Statistical Office (2009) indicated that the participating children were representative for the population in the federal state where they came from in terms of gender ratio ($\chi^2 = .004$, $df = 1$, $p = .950$) and immigration background ($\chi^2 = .220$, $df = 1$, $p = .639$). Between July 2009 and July 2011, data were collected at seven measurement occasions (see Table 2). Each school year started in the middle of September and ended at the end of July. Children received report cards with their grades at the end of each school year (at the end of July) and in the middle of each school year (at the beginning of February). The first measurement occasion (t_1) took place at the end of Grade 2 when children were on average 7.95 ($SD = .58$) years old. The next six measurement occasions followed at an interval of 4 months each, so that the children were approximately 2 years older at the last measurement occasion (t_7; $M = 9.93$, $SD = .72$). In Germany, students leave elementary school after Grade 4 to visit a specific secondary school type. The teacher gives a recommendation for the Gymnasium (academic track) or for another kind of secondary school (different kinds of vocational tracks) in the fourth grade based on the student's school grades at this time. We investigated students in a federal state where the teacher's recommendation was binding. In 2011, 50.4% of the students in this federal state received a Gymnasium recommendation (Schwarz-Jung, 2013). Only students whose parents provided informed consent participated in the study, which equals 75% of the basic population.

Besides other variables that were assessed in MEGA, the variables of interest were assessed during regular class in students' classrooms at all measurement occasions. All items were read aloud to make sure that all students worked at the same speed. In total, testing took about one lesson (45 min) and was conducted by trained research assistants. While students were working on their questionnaires, teachers filled in a questionnaire. Inter alia, they indicated students' math grades.

Measures

Math Competence Beliefs

The items measuring children's competence beliefs in math had been established in previous studies (Spinath & Steinmayr, 2008). Items are based on the questionnaire to assess ability perceptions according to the *expectancy-value model* (Wigfield et al., 1997). Similar items were used to assess students' competence beliefs in other studies (e.g., Musu-Gillette et al., 2015; Viljaranta et al., 2014). The items used in the present study were, "How good are you at math?" with response choice ranging from 1 (*very good*) to 5 (*very bad*), "How easy is it for you to learn new things in math?," from 1 (*very easy*) to 5 (*very hard*), and "To which group of pupils do you belong in your class in math?," from 1 (*the best*) to 5 (*the worst*). Results of prior studies support the construct validity of this competence beliefs scale (Spinath & Steinmayr, 2008). A confirmatory factor analysis (CFA) of three correlated competence beliefs factors (i.e., competence beliefs for math, German, and school in general) yielded a good model fit for children from the end of second to the end of third grade. Factor loadings of competence beliefs items were high within domains (Spinath & Steinmayr, 2008). To estimate the reliability of the scale in the present study, we calculated McDonald's omega (ω), which varied from $\omega = .883$ to $\omega = .941$ (see Table 2).

Math Grades

At each time of measurement, teachers filled in a questionnaire in which they were asked to report students' current math grades. They received the following instruction: "Please note the current math grade of the student (i.e., the grade that the student would receive if it was time for report cards now)." These grades mainly rely on the student's

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2
Latent Means (M) and Standard Deviations (SD), Reliabilities (McDonald's Omega, ω), Intraclass Correlations (ICC), and Sample Sizes (n) for Competence Beliefs and Math Grades at Seven Measurement Occasions (t)

t	t_1	t_2	t_3	t_4	t_5	t_6	t_7
year-month	2009-07	2009-11	2010-03	2010-07	2010-11	2011-03	2011-07
	Grade 2		Grade 3			Grade 4	
Competence beliefs[a]							
M	3.980	3.915	3.821	3.787	3.879	3.788	3.752
SD	0.737	0.731	0.739	0.751	0.718	0.752	0.730
ω	.883	.896	.941	.907	.919	.920	.928
ICC	.030	.014	.027	.034	.025	.024	.040
n	458	468	462	468	451	449	446
Math grades[b]							
M	4.763	4.712	4.744	4.608	4.592	4.594	4.569
SD	0.881	0.920	0.899	1.008	0.952	0.923	0.910
ICC	.147	.167	.198	.163	.124	.120	.100
n	406	435	452	433	461	418	454
Age							
M	7.946	8.278	8.600	8.904	9.284	9.618	9.926
SD	0.575	0.543	0.605	0.590	0.549	0.743	0.719

Note. $N = 540$ students in 27 classes. Latent means and standard deviations are from the longitudinal CFA model with competence beliefs and grades, where phantom constructs were used to convert covariance information into the latent standard deviations of competence beliefs. [a]Range: 1 to 5. [b]Range: 1 to 6.

performance on written math tests that are taken at several times during the school year. Grades also include the evaluation of oral participation in class. Thus, the math grades considered in our analyses represented an averaged measure of students' preceding performance (i.e., cumulative attainment; see Niepel et al., 2014b), and they were expected to be a good indicator of the performance feedback that the student received from the teacher because they were based on the grades that the students received during the school year. In Germany, grades range from 1 (*excellent*) to 6 (*insufficient/fail*).

For a more intuitive understanding of the scales, all items were recoded so that higher values denoted more positive competence beliefs and better grades.

Statistical Analysis

To model mean-level changes in competence beliefs on the one hand and to examine the expected directive relations between competence beliefs and grades on the other, we estimated latent growth curve models (LGCM) and latent cross-lagged models (CLM). All analyses were computed in SPSS 22.0 and M*plus* 6.12 (Muthén & Muthén, 1998-2011). For parameter estimation, maximum likelihood estimation with robust standard errors (MLR) was used to account for nonindependence of observations resulting from the hierarchical structure of the data (i.e., students nested within classes; "type = complex" option in M*plus* with class membership as a cluster variable). Latent analyses were based on data from 540 students (269 boys and 271 girls) because for two children the information on class membership was missing.

Missing Data

As is common in longitudinal studies, the current study had to deal with missing data related to nonresponse and attrition (see Little, 2013). Missing data resulting from nonresponse occurred at each measurement occasion and ranged from 0.0% (t_2 and t_3) to 0.8% (t_4) for competence beliefs and from 0.2% (t_5) to 11.5% (t_1) for grades. Missing data related to attrition occurred when children missed one measurement occasion. The main reason for nonparticipation was illness. Complete data sets for all seven measurement occasions were available for 264 students (48.7%); 490 students (90.1%) participated at least at five measurement occasions. More detailed information about the number of students participating at each measurement can be found in Table 2. Following recommendations (Graham & Coffman, 2012; Little, 2013), full information

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

maximum likelihood estimation (FIML) was used to account for missing data. To make the assumption of missing-at-random more plausible, we estimated all models using the AUXILIARY option in M*plus*, which is based on Graham's (2003) *saturated correlates approach* (Asparouhov & Muthén, 2008). Within this approach, a set of variables is considered as missing data correlates (i.e., auxiliary variables) in addition to the analysis variables (Graham, 2003). To identify potential correlates of missingness that can serve as auxiliary variables in subsequent analyses, we examined mean differences across missing data patterns as well as correlations between the incomplete analysis variables and other variables in the data set (Enders, 2010). Two variables (i.e., the teacher rating of children's math ability and children's self-reported ability self-concept concerning school in general) were identified as potential correlates of missingness and included as auxiliary variables in all analysis models. Further information on the auxiliary variables, which are not of substantive interest here, can be requested from the corresponding author.

Longitudinal CFA Modeling

First, we estimated a longitudinal CFA model to test for strong factorial invariance of competence beliefs measures across time (i.e., time-invariant factor loadings and intercepts; see Meredith, 1993). Competence beliefs were indicated by three items at each of the seven measurement occasions, the residual variances among the corresponding indicators were allowed to correlate over time, and the effects coded method of identification was used for scale setting (Little, 2013). In a first step, a *configurally invariant model* where only the factor structure was invariant across time was compared to a more parsimonious *weak invariant model* where factor loadings of the manifest indicators were fixed to be invariant across time. If the change in model fit was negligible, we additionally fixed the intercepts of the manifest indicators to be invariant across time (*strong invariant model*). The supposition of strong factorial invariance holds if this additional restriction does not lead to a significant deterioration in model fit. To evaluate invariance constraints, we calculated the change in CFI (ΔCFI) as here the chi-square (χ^2) difference test is too sensitive to trivial differences (Little, 2013). Following guidelines from a simulation study, the supposition of invariance is tenable if $\Delta CFI \leq .01$ (Cheung & Rensvold, 2002).

To estimate latent means and standard deviations of competence beliefs and grades as well as the correlations across time, grades from t_1 to t_7 were included as manifest variables in the longitudinal CFA model. Phantom constructs were used to transform covariance information into estimates of the latent standard deviations of competence beliefs (Little, 2013).

Latent Growth Curve Modeling

Second-order LGCMs were used to fit a model to the mean levels of competence beliefs, so that the intraindividual development of competence beliefs and interindividual differences in this development could be examined (Hypothesis 1). LGCMs allow conclusions on whether a student's competence beliefs significantly changed over time (i.e., change in mean levels), on the shape of this development, and on whether the trend differed between students. In the first step, an unconditional LGCM was estimated (*level and shape model*; Little, 2013). Competence beliefs were considered as first-order factors from t_1 to t_7 indicated by three items each time (see CFA model). The latent intercept and slope were modeled as second-order factors, whereby the following factor loadings were set to model change: All loadings of the first-order factors on the intercept were set to 1. In the unconditional LGCM, the slope factor loadings from t_2 to t_6 (γ_{t2} to γ_{t6}) were freed, whereas the loading of competence beliefs at t_1 (γ_{t1}) was set to 0, and the loading at t_7 (γ_{t7}) was fixed to 1: [0 * * * * * 1]. To evaluate whether there was a linear or nonlinear trend in competence beliefs, in a second step, the unconditional LGCM was compared to a linear LGCM within which the slope factor loadings were fixed to [0 4 8 12 16 20 24] to estimate a linear development. We used the scaled chi-square difference test to evaluate change in model fit (Satorra & Bentler, 2001). If the restriction of the slope factor loadings led to a significant decline in model fit, the less parsimonious unconditional model with freed loadings should be accepted. This would indicate that the development was nonlinear. In this case, we chose the unconditional model to fit the nonlinear trend instead of polynomial models with an additional higher order term such as a quadratic or cubic slope because there is no sufficient theory about the specific elements of a nonlinear trend in competence beliefs (see Little, 2013).

Latent Cross-Lagged Modeling

CLMs enable to investigate the stability of interindividual differences (i.e., relative stability)

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and the relation between two constructs across time when controlling for prior values on these constructs. To investigate the predictive relations among competence beliefs and grades, a multivariate CLM with competence beliefs and grades was set up (Hypothesis 2). Therefore, the cross-time nondirectional associations of the CFA model were converted into direct paths. Only the within-time associations were still modeled as nondirectional covariance relations. The CLM was a more restricted model than the CFA model in that only theoretically based structure relations were specified (i.e., autoregressive and cross-lagged paths between adjacent times of measurement). Phantom constructs were used, first, to test for stationarity of autoregressive, skill development and self-enhancement effects (Little, 2013). That is, standardized autoregressive and cross-lagged paths, respectively, were equated over time, and change in model fit was evaluated calculating the scaled chi-square difference test (Satorra & Bentler, 2001). A significant poorer fit of the constraint model with equated autoregressive, skill development or self-enhancement paths compared to the baseline model within which all paths were freely estimated would indicate that the respective effects changed over time. In this case, pairwise comparisons were performed to examine which paths differed from each other. Second, the skill development and self-enhancement paths between adjacent time points were equated (e.g., the cross-lagged paths between t_1 and t_2) to examine the dominant direction of the relation between competence beliefs and grades. Each constrained model with two equated cross-lagged paths was compared to the unconstrained baseline model within which all cross-lagged paths were freely estimated. If the constraint led to a significant decline in model fit, it would indicate that the respective cross-lagged paths differed from each other and either skill development or self-enhancement effects prevailed.

To evaluate the goodness of fit of all the tested models, we computed the chi-square test statistic, the root mean square error of approximation (RMSEA) along with associated 90% confidence interval, the standardized root mean square residual (SRMR), and the comparative fit index (CFI). According to recent recommendations, the following cutoff criteria for the fit indices were used to evaluate model fit: both the point estimate and the upper confidence interval of RMSEA < .06, SRMR < .08, CFI > .95 (see West, Taylor, & Wu, 2012; Wu & West, 2010).

Results

Longitudinal Measurement Invariance and Descriptive Statistics

Results of the model comparisons performed within the scope of invariance testing can be found in Table 3. There was no substantial decline in model fit when constraining the factor loadings to be invariant over time. An additional constraint of all intercepts to be invariant over time also resulted in a nonsubstantial decline in model fit. Therefore, we can legitimately conclude that strong factorial longitudinal invariance was given (for changes in CFI, see Table 3). Thus, students were interpreting items and using response scales in the same way over time. In all subsequent analyses, strong longitudinal invariance of the math competence beliefs measures was specified.

Besides children's mean age and the intraclass correlations of children's competence beliefs and their grades at all measurement occasions, Table 2 depicts the reliabilities of the competence beliefs scale, which were high. Moreover, Table 2 contains the latent means and standard deviations from the longitudinal CFA model with math competence beliefs and grades. The model showed a good fit to

Table 3
Model Fit Statistics for Testing the Longitudinal Measurement Invariance of Competence Beliefs

Model tested	χ^2 (*df*)	*p*	$\Delta\chi^2_{corr}$(*df*)	*p*	RMSEA	RMSEA 90% CI	SRMR	CFI	ΔCFI	Pass?
Configurally invariant model	103.490 (105)	.523	—	—	< .001	[.000, .022]	.022	1.000	—	—
Weak invariant model	112.859 (117)	.591	9.643 (12)	.647	< .001	[.000, .020]	.032	1.000	.000	Yes
Strong invariant model	152.870 (129)	.074	37.283 (12)	< .001	.019	[.000, .029]	.032	.995	.005	Yes

Note. $N = 540$. Competence beliefs were assessed at seven measurement occasions with three items. Maximum likelihood estimation with robust standard errors (MLR) was used. χ^2 (*df*) = chi-square test statistic with degrees of freedom in parentheses; $\Delta\chi^2_{corr}$ (*df*) = test statistic of the scaled chi-square difference test with degrees of freedom in parentheses; RMSEA = root mean square error of approximation along with its associated confidence interval; SRMR = standardized root mean square residual; CFI = comparative fit index; Configurally invariant model = invariant factor structure; Weak invariant model = invariant factor loadings; Strong invariant model = invariant factor loadings and intercepts.

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 4
Correlations Between Students' Math-Related Competence Beliefs (CB) and Math Grades (MG) at All Measurement Occasions (t)

	CB t_2	CB t_3	CB t_4	CB t_5	CB t_6	CB t_7	MG t_1	MG t_2	MG t_3	MG t_4	MG t_5	MG t_6	MG t_7
CB t_1	.711	.573	.563	.574	.518	.520	.453	.356	.279	.356	.349	.352	.362
CB t_2		.753	.722	.684	.666	.641	.543	.551	.442	.503	.463	.449	.483
CB t_3			.826	.783	.750	.757	.598	.581	.528	.569	.550	.551	.540
CB t_4				.834	.805	.792	.571	.551	.515	.582	.596	.586	.532
CB t_5					.851	.816	.492	.484	.468	.525	.575	.536	.512
CB t_6						.900	.571	.607	.590	.624	.657	.665	.646
CB t_7							.605	.615	.588	.638	.685	.673	.670
MG t_1								.707	.649	.718	.692	.703	.692
MG t_2									.784	.779	.739	.734	.735
MG t_3										.800	.745	.737	.732
MG t_4											.773	.807	.805
MG t_5												.876	.822
MG t_6													.872
MG t_7													—

Note. Results are from the longitudinal CFA model with competence beliefs and grades. All correlations were significant on the $p < .001$ level.

the data, $\chi^2_{(227)} = 277.778$, $p = .012$; RMSEA = .020 [.010, .028], SRMR = .036, CFI = .994. The correlations from the CFA model are depicted in Table 4. All correlations were moderate to high.

Mean-Level Changes in Math Competence Beliefs

To test whether mean levels in competence beliefs changed over time, we ran LGCMs (Hypothesis 1). Results are presented in Table 5. The unconditional model showed a good fit to the data. Constraining the slope factor loadings to [0 4 8 12 16 20 24] to specify a linear trend in competence beliefs led to a significant deterioration in model fit ($\Delta\chi^2 = 14.450$, $\Delta df = 5$, $p = .013$) indicating that the trend was nonlinear. As recommended, we thus used the unconditional LGCM to describe the development of competence beliefs. The latent mean of the intercept and slope factors were significantly different from zero ($p < .001$, see Table 5). The estimated mean of the slope had a negative algebraic sign indicating a significant average decline in mathematical competence beliefs over time. Moreover, the variance of the slope was significantly different from zero pointing to interindividual differences in this decline.

Table 5
Unconditional Second-Order Latent Growth Curve Model (LGCM) With Competence Beliefs in Math

	Parameter (*SE*)[a]
Slope factor loadings	
γ_{t1}	0.00
γ_{t2}	0.248 (0.145)†
γ_{t3}	0.623 (0.125)***
γ_{t4}	0.757 (0.105)***
γ_{t5}	0.731 (0.088)***
γ_{t6}	0.990 (0.056)***
γ_{t7}	1.00
Variances	
$\psi_{Intercept}$	0.411 (0.045)***
ψ_{Slope}	0.321 (0.070)***
Covariance	
$\psi_{Intercept,\ Slope}$	−0.117 (0.038)**
Latent means	
$\mu_{Intercept}$	3.978 (0.043)***
μ_{Slope}	−0.198 (0.058)**

Note. Unstandardized solution. γ = estimated regression of one construct onto another; t = measurement occasion. Model fit: $\chi^2_{(147)} = 203.250$, $p = .002$; RMSEA = .027 [.017, .035]; SRMR = .046; CFI = .989. [a]Maximum likelihood parameter estimates with standard errors and a chi-square test statistic that are robust to nonindependence of observations (MLR) were used. **$p < .01$. ***$p < .001$. $^{\dagger}p = .088$.

The Direction of the Relation Between Competence Beliefs and Achievement in Math

To investigate the direction of the association between students' competence beliefs and grades, we evaluated a CLM that included t_1 to t_7 measures (see Figure 1). Fit indices for this model indicated a good model fit, $\chi^2_{(287)} = 637.998$, $p < .001$; RMSEA = .048 [.043, .053], SRMR = .073, CFI = .956. To test for

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

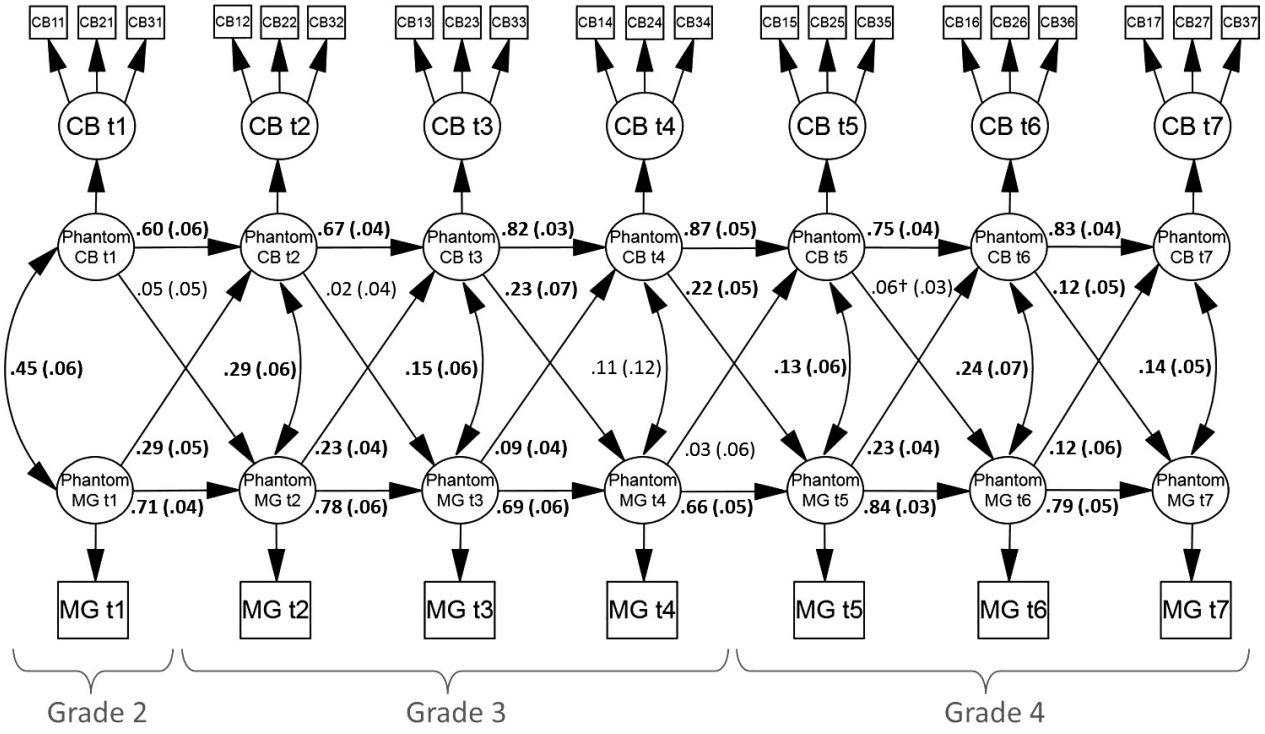


Figure 1. Latent cross-lagged model (CLM) of math-related competence beliefs (CB) and math grades (MG) assessed at seven measurement occasions (*t*). $N = 540$. Standardized solution (standard errors in parentheses). Phantom constructs were used to test for differences in standardized cross-lagged paths across time using scaled chi-square difference test (Little, 2013). Concurrent correlations between phantom constructs at t_2–t_7 represent residual correlations between the endogenous variables. Factor loadings and correlated uniquenesses between all corresponding CB measures collected at subsequent measurement occasions are not shown. Factor loadings and intercepts of the manifest indicators are invariant across time. Bold coefficients are significant ($p < .05$). Model fit: $\chi^2_{(287)} = 637.998$, $p < .001$; RMSEA = .048 [.043, .053], SRMR = .073, CFI = .956. †$p = .070$.

stationarity of the autoregressive effects, autoregressive paths from prior to subsequent competence beliefs and from prior to subsequent grades, respectively, were equated across time (competence beliefs: $\beta = .873$, $SE = .027$, $p < .001$; grades: $\beta = .823$, $SE = .031$, $p < .001$). For competence beliefs, the equation of the autoregressive paths did not result in a significant decline in model fit indicating that relative stabilities of competence beliefs were comparable across time ($\Delta\chi^2 = 2.033$, $\Delta df = 5$, $p = .845$). By contrast, the equation of the autoregressive paths of grades led to a significant decline in model fit indicating that relative stabilities of grades differed across time ($\Delta\chi^2 = 11.305$, $\Delta df = 5$, $p = .046$). Pairwise comparisons revealed that the autoregressive effect between t_1 grade and t_2 grade ($\beta_{t1\rightarrow t2} = .707$) was significantly stronger than the one between t_3 grade and t_4 grade ($\beta_{t3\rightarrow t4} = .687$; $\Delta\chi^2 = 4.749$, $\Delta df = 1$, $p = .029$) and the one between t_4 grade and t_5 grade ($\beta_{t4\rightarrow t5} = .657$; $\Delta\chi^2 = 10.426$, $\Delta df = 1$, $p = .001$). Moreover, the autoregressive effect between t_5 grade and t_6 grade ($\beta_{t5\rightarrow t6} = .838$) was significantly stronger than the one between t_3 grade and t_4 grade ($\beta_{t3\rightarrow t4} = .687$; $\Delta\chi^2 = 5.385$, $\Delta df = 1$, $p = .020$) and the one between t_4 grade and t_5 grade ($\beta_{t4\rightarrow t5} = .657$; $\Delta\chi^2 = 42.607$, $\Delta df = 1$, $p < .001$). All other autoregressive effects did not differ from each other ($\Delta\chi^2 \leq 3.561$, $\Delta df = 1$, $p > .05$).

The critical parameters to evaluate the direction of the association are the cross-paths leading from one construct at a specific measurement occasion to the other at the following measurement occasion (see Figure 1). Consistent with expectations, effects from prior competence beliefs to subsequent grades increased from the end of Grade 2 (t_1) to the beginning of Grade 4 (t_5) so that the first two cross-paths were not significant, whereas the last two were (t_1 to t_2: $\beta = .046$, $p = .327$; t_2 to t_3: $\beta = .017$, $p = .652$; t_3 to t_4: $\beta = .226$, $p = .001$; t_4 to t_5: $\beta = .215$, $p < .001$). Over the same period (t_1 to t_5), an opposite pattern occurred for cross-paths leading from prior math grades to subsequent competence beliefs. Here, the first three cross-paths were

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

statistically significant but not the last one (t_1 to t_2: $\beta = .291$, $p < .001$; t_2 to t_3: $\beta = .231$, $p < .001$; t_3 to t_4: $\beta = .093$, $p = .015$; t_4 to t_5: $\beta = .028$, $p = .663$). Contrary to expectations, the self-enhancement effect declined again during the first half of Grade 4 (competence beliefs t_5 to math grades t_6: $\beta = .062$, $p = .070$; grades t_5 to competence beliefs t_6: $\beta = .230$, $p < .001$). However, skill development and self-enhancement effects were finally comparable at the end of elementary school (competence beliefs t_6 to math grades t_7: $\beta = .121$, $p = .023$; grades t_6 to competence beliefs t_7: $\beta = .119$, $p = .036$).

To test for stationarity of self-enhancement effects, cross-lagged paths from competence beliefs to grades were equated across time ($\beta = .116$, $SE = .018$, $p < .001$). The equation led to a substantial decline in model fit indicating that self-enhancement effects were not time invariant ($\Delta\chi^2 = 14.498$, $\Delta df = 5$, $p = .013$). Pairwise comparisons revealed that the self-enhancement effect between t_3 and t_4 ($\beta_{t3\rightarrow t4} = .226$) was significantly stronger than the one between t_1 and t_2 ($\beta_{t1\rightarrow t2} = .046$; $\Delta\chi^2 = 5.617$, $\Delta df = 1$, $p = .012$), the one between t_2 and t_3 ($\beta_{t2\rightarrow t3} = .017$; $\Delta\chi^2 = 11.750$, $\Delta df = 1$, $p = .001$), and the one between t_5 and t_6 ($\beta_{t5\rightarrow t6} = .062$; $\Delta\chi^2 = 10.155$, $\Delta df = 1$, $p = .001$). Moreover, the self-enhancement effect between t_4 and t_5 ($\beta_{t4\rightarrow t5} = .215$) was significantly stronger than the one between t_2 and t_3 ($\beta_{t2\rightarrow t3} = .017$; $\Delta\chi^2 = 10.043$, $\Delta df = 1$, $p = .002$), and the one between t_5 and t_6 ($\beta_{t5\rightarrow t6} = .062$; $\Delta\chi^2 = 7.098$, $\Delta df = 1$, $p = .008$). All other self-enhancement effects did not differ from each other ($\Delta\chi^2 \leq 3.094$, $\Delta df = 1$, $p > .05$).

To test for stationarity of skill development effects, cross-lagged paths from grades to competence beliefs were equated across time ($\beta = .193$, $SE = .029$, $p < .001$). Again, the equation led to a substantial decline in model fit indicating that skill development effects were not time invariant too ($\Delta\chi^2 = 25.860$, $\Delta df = 5$, $p < .001$). Pairwise comparisons revealed that the skill development effect between t_1 and t_2 ($\beta_{t1\rightarrow t2} = .291$) was significantly stronger than the one between t_2 and t_3 ($\beta_{t2\rightarrow t3} = .231$; $\Delta\chi^2 = 9.386$, $\Delta df = 1$, $p = .002$), the one between t_3 and t_4 ($\beta_{t3\rightarrow t4} = .093$; $\Delta\chi^2 = 22.260$, $\Delta df = 1$, $p < .001$), the one between t_4 and t_5 ($\beta_{t4\rightarrow t5} = .028$; $\Delta\chi^2 = 24.127$, $\Delta df = 1$, $p < .001$), and the one between t_6 and t_7 ($\beta_{t6\rightarrow t7} = .119$; $\Delta\chi^2 = 8.163$, $\Delta df = 1$, $p = .004$). The skill development effect between t_2 and t_3 ($\beta_{t2\rightarrow t3} = .231$) was significantly stronger than the one between t_3 and t_4 ($\beta_{t3\rightarrow t4} = .093$; $\Delta\chi^2 = 10.963$, $\Delta df = 1$, $p = .001$), and the one between t_4 and t_5 ($\beta_{t4\rightarrow t5} = .028$; $\Delta\chi^2 = 6.980$, $\Delta df = 1$, $p = .008$). Finally, the skill development effect between t_5 and t_6 ($\beta_{t5\rightarrow t6} = .230$) was significantly stronger than the one between t_3 and t_4 ($\beta_{t3\rightarrow t4} = .093$; $\Delta\chi^2 = 8.413$, $\Delta df = 1$, $p = .004$) and the one between t_4 and t_5 ($\beta_{t4\rightarrow t5} = .028$; $\Delta\chi^2 = 11.208$, $\Delta df = 1$, $p = .001$). All other skill development effects did not differ from each other ($\Delta\chi^2 \leq 2.790$, $\Delta df = 1$, $p > .05$).

To examine the dominant direction of the relation, the skill development and self-enhancement paths between adjacent time points were equated. Between t_1 and t_2, t_2 and t_3, and t_5 and t_6, this restriction led to a substantial decline in model fit indicating that skill development effects were stronger than self-enhancement effects between these times of measurement ($\Delta\chi^2 \geq 14.686$, $\Delta df = 1$, $p < .001$). This also was the case between t_4 and t_5, whereby here self-enhancement effects were stronger than skill development effects ($\Delta\chi^2 = 4.070$, $\Delta df = 1$, $p = .044$). Between t_3 and t_4, as well as t_6 and t_7 an equation of the two cross-lagged paths did not result in a significant deterioration in model fit indicating that effects were comparable at this time ($\Delta\chi^2 \leq 3.381$, $\Delta df = 1$, $p \geq .066$).

Taken together, we observed changes in the dominant direction of the relation between competence beliefs and grades in math from a skill development to a REM. Consistent with expectations, pattern of results indicates that competence beliefs substantially affected grades not until the middle of Grade 3. However, against Hypothesis 2, self-enhancement effects did only increase until Grade 3 and were temporarily negligible in Grade 4.

Discussion

The primary focus of the present study was to shed further light on the early development of children's competence beliefs in math and the interplay between math competence beliefs and achievement across the elementary school years. The results showed that, on average, children's math competence beliefs became more negative over time, whereby the intraindividual trend was nonlinear and differed between students. Furthermore, our analyses indicated that the direction of the relation between math competence beliefs and grades changed over time. At the end of second and the beginning of third grade, prior achievement was important for later competence beliefs in math, whereas prior math competence beliefs did not significantly impact on subsequent achievement. During Grade 3, students' competence beliefs in math became predictive of change in grades, whereby this pattern changed

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

temporarily at the beginning of Grade 4. Finally, at the end of Grade 4, we observed a reciprocal relation. Taken together, our results provide evidence for the growing importance of students' competence beliefs for their performance development in math.

Development of Math Competence Beliefs

Considering the mean levels, children's competence beliefs in math became more negative over the investigated time period. The finding is consistent with previous longitudinal studies investigating changes in self-perceived abilities of elementary school children (e.g., Jacobs et al., 2002; Spinath & Steinmayr, 2008). Moreover, it conforms to the theory that students have very positive competence beliefs at the beginning of their school trajectories but soon are starting to use external performance information and social comparisons to generate accurate competence beliefs (e.g., Harter, 2006). In this respect, it is plausible that the decline in competence beliefs was stronger for some children than for others in the present study. In a normative context, some students are the high-performing ones who do not need to strongly adjust their positive competence beliefs because objective achievement information (like grades) align with their subjective perception of own competence. By contrast, the low-performing students should adjust their competence beliefs downward more strongly because positive competence beliefs turn out to be untenable in light of the negative feedback these children receive. However, if the decline in competence beliefs was only due to the increasing realism in self-evaluations, the competence beliefs values of the older children should be spread more broadly across the scale than the scores of the younger children, which should be clustered at the scale's high end (see Eccles et al., 1993). As the variance in competence beliefs did not increase over time, the results suggest that the increasing realism in self-evaluations is not the only process leading to a decline in competence beliefs across the elementary school years (see also Eccles et al., 1993). Moreover, our findings indicate that the overall trend in competence beliefs was nonlinear. Intraindividual changes in competence beliefs were less pronounced in the last months of each school year. One reason might be that fewer new topics are presented at the end of a school year and, simultaneously, students have already received a lot of feedback on their math performance in the current school year. Thus, children should be relatively familiar with the learning material and well aware of their strengths and weaknesses so that changes in mean levels of competence beliefs were less likely at the end of a school year.

Considering the stability, results indicate that interindividual differences in competence beliefs were very stable from the end of Grade 2 to the end of Grade 4. Thus, changes in rank orders seem to be unlikely as early as in elementary school. That previous studies found weaker stabilities of math competence beliefs in this age group (Helmke & van Aken, 1995; Skaalvik & Valas, 1999; Viljaranta et al., 2014) is not surprising given the larger distance between the times of measurement in these studies. The combination of both, a decline in mean levels and high relative stabilities found in the present study, puts children with low competence beliefs in the first school years particularly at risk (see Gottfried, Fleming, & Gottfried, 2001; Musu-Gillette et al., 2015). The increased awareness of their own math competencies compared to those of their peers helps students to make appropriate choices in school even if this awareness results in relatively low math competence beliefs. However, as Stipek (2011) stated, the relative standing tells the student little about which level of math competence the student could achieve when putting effort into math. Feeling incompetent compared to others bears the risk to give up too quickly. Although the decline in math competence beliefs was relatively modest in our study, it should thus not be ignored. The latter becomes obvious when taking a closer look at the early relation between competence beliefs and achievement.

Direction of the Relation Between Early Competence Beliefs and Achievement in Math

The early predominance of achievement over competence beliefs in math observed in our study is consistent with the skill development approach, according to which children form their achievement-related self-perceptions in response to experienced success and failure (e.g., Chapman & Tunmer, 1997; Helmke & van Aken, 1995; Skaalvik & Hagtvet, 1990). Like the overall decline in children's math competence beliefs described above, this strengthens the supposition that children adopt their self-evaluations to the external feedback they receive and generate accurate domain-specific competence beliefs (Dweck, 2002). A special characteristic of the present sample is that German students receive their first report cards with grades on it at the end of Grade 2. The implementation of grades should enable direct achievement comparisons

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

between students and, thus, intensify these comparisons and contribute to the impact achievement had on subsequent math competence beliefs in our study. The two previous studies that also investigated the relation between competence and achievement from Grades 2 to 3 also found a stronger skill development than self-enhancement effect in reading and math, respectively (Chapman & Tunmer, 1997; Helmke & van Aken, 1995). However, although the self-enhancement effect was small, it was significant in the study by Helmke and van Aken when they considered math grades or test scores as achievement indicator. One explanation might be that the time interval between the two measurement occasions was 1 school year instead of 4 month as in our study. Thus, Helmke and van Aken could not test the relations between the constructs within a year.

After 1 year, when children were on average 8 to 9 years old, math competence beliefs affected subsequent math grades in our study. The finding is in accordance with the developmental hypothesis that after domain-specific competence beliefs have been established, they exert effects on achievement because children increasingly use current domain-specific competence beliefs to make predictions about future success in this domain (e.g., Dweck, 2002), which in turn should affect their achievement-related behavior (see Wigfield et al., 2015). Self-enhancement tendencies prevailed when children were on average 9 years old because the impact of grades on subsequent competence beliefs diminished until the end of Grade 3. The latter might be due to the decrease in the relative stability of math grades from the end of second to the end of third grade. That previous studies investigating math competence beliefs and achievement in elementary school did not find a change in the association between the constructs (Helmke & van Aken, 1995; Skaalvik & Valas, 1999; Viljaranta et al., 2014) might actually be due to the relatively long distance between measurement occasions in these studies. Our results suggest that longtime intervals between measurement occasions (e.g., 1 year) mask developmental changes in and effects of domain-specific competence beliefs in elementary school. Given that meta-analyses with students of different age groups did not find any moderating effects of the delay between measurement occasions on the relation between competence beliefs and achievement (Huang, 2011; Valentine et al., 2004), this might only be an issue in elementary school when children are establishing domain-specific competence beliefs.

Aside from the study by Helmke and van Aken (1995), previous studies on elementary school students only found effects of prior competence beliefs on subsequent achievement when they assessed competence beliefs concerning school in general or sports (Guay et al., 2003; Marsh et al., 2007) instead of domain-specific competence beliefs for the math or verbal domain (Chapman & Tunmer, 1997; Skaalvik & Valas, 1999; Viljaranta et al., 2014). An explanation could be that the relative stabilities of general academic and sports competence beliefs, respectively, were somewhat higher than those of verbal and math competence beliefs across 1 year or more in elementary school. Maybe children have already generated beliefs about their competence in school in general and in nonacademic domains like sports before they establish domain-specific competence beliefs for academic domains like math and reading, so that competence beliefs in the math and verbal domain are less stable over 1 year in elementary school. Although the meta-analysis by Huang (2011) revealed that the relation between competence beliefs and achievement did not differ significantly depending on whether constructs were assessed concerning school in general or domain-specific, self-enhancement effects at least tended to be stronger in studies that assessed the constructs concerning school in general (see Huang, 2011). The latter might be especially an issue in elementary school, when domain-specific competence beliefs first need to be established.

Against expectations, effects of prior competence beliefs on subsequent grades no longer increased after Grade 3. That self-enhancement effects were temporarily negligible at the beginning of Grade 4, might be due to characteristics of the German school system: As mentioned earlier, students in Germany receive recommendations for secondary school on the basis of the grades they gain in the first term of Grade 4. As these recommendations for secondary track are highly important for the students, achievement comparisons might have increased and have led to more unsteady competence evaluations, so that the predictive power of competence beliefs decreased. At the same time, the subjective importance of grades might have increased at the beginning of Grade 4 so that grades had a stronger effect on self-evaluations. However, we cannot test this hypothesis because we did not assess the personal relevance grades had for children and whether this relevance varied over time. Because of the supposed effects of the school system, further studies are necessary that separate age effects from school system effects.

As expected, the relation between competence beliefs and grades was reciprocal at the end of

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

elementary school. This confirms the hypothesis that motivation and academic attainment are becoming a coherent self-stabilizing system at the end of elementary school (see Dweck, 2002). Accordingly, results of studies on German students after the transition to secondary school also support the reciprocal effects model for the math domain (e.g., Niepel et al., 2014b).

Taken together, findings indicate that the interplay of academic achievement and competence beliefs in math shows a developmental trend from skill development to reciprocal effects because of an increase in self-enhancement effects in Grade 3. Thus, math competence beliefs exert effects on achievement as early as in elementary school. Moreover, the present study suggests the following. First, especially in the elementary school years, domain-specific competence beliefs should be assessed several times within school years to provide a complete picture of their development, effects, and determinants. Second, studies on the relation between competence beliefs and achievement should distinguish between general academic competence beliefs and domain-specific competence beliefs, particularly in the educational phase when children are establishing domain-specific competence beliefs. Finally, besides individual characteristics of the child (e.g., age, achievement level), also contextual factors (e.g., school system) might be relevant for the development of math competence beliefs and should be investigated in future studies.

Limitations and Future Directions

Some limitations for the generalization of the present data should be pointed out. First, we investigated only one domain, namely mathematics. It is important to evaluate the generalizability of the findings in other domains as well (e.g., reading).

Second, our study is limited to a sample of students from Germany. The German school system is characterized by a relatively early ability tracking (generally after Grade 4), so that achievement pressure might be higher for 9-year-olds in Germany than for their peers in other countries. Nevertheless, also in other countries grades represent an important form of teacher feedback, and effects of grades on students' competence beliefs were found (e.g., Chen, Yeh, Hwang, & Lin, 2013; Viljaranta et al., 2014). In terms of generalization, findings of the present study should be replicated with students from other countries.

Third, we used only one indicator for math achievement at each measurement occasion, namely the current math grade. Using multiple indicators could enhance the reliability of the achievement measure. Besides this, there is an ongoing debate on how students' achievement should be measured and whether the type of achievement indicator (grades, objective performance tests, teacher ratings) impacts on the causal pattern between achievement and competence beliefs (e.g., Helmke & van Aken, 1995; Pinxten et al., 2010). Therefore, it is advisable that future studies measure grades as one achievement variable among others.

Fourth, future studies should also include other sources of feedback. For example, it might be interesting to look at verbal communications that accompany the announcement of test results by teachers or the comments parents make when they learn about such results.

Fifth, we did not include other variables in our analyses, which might have an influence on the relation between competence beliefs and school grades. Although longitudinal designs allow conclusions on temporal precedence, causal interpretations should be offered tentatively (see Biesanz, 2012). Besides student variables like goal orientations, also more distal context variables like socializer's beliefs and behaviors might play a role in this context (see *expectancy-value theory*; Wigfield & Eccles, 2000). However, experimental studies in which potential confounding factors were controlled by random assignment of students also revealed that competence beliefs and achievement are causally linked (e.g., Möller, 2005).

Finally, strictly speaking, item-level data were categorical in nature. We nevertheless preferred MLR over weighted least squares estimation techniques because we chose a model-based approach to handle missing data (i.e., FIML). The latter is incompatible with the use of weighted least squares estimation techniques (for weighted least squares estimation in the presence of missing data, see Asparouhov & Muthén, 2010). As all variables considered in our analyses can be classified to be approximately normally distributed, and the observed data have many categories (i.e., at least five ordered categories) substantial bias in fit indices, parameter estimates, or standard errors when using maximum likelihood estimation techniques are unlikely (see Finney & DiStefano, 2013).

Conclusion and Practical Implications

The present study supports the hypothesis that changes in self-enhancement processes can be explained by developmental changes that take place in elementary school age. This shows that a developmental perspective can further our

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

knowledge about processes that are investigated in the field of educational psychology, and it underlines the benefit of an integrative approach. However, our results are not only of theoretical interest but are also important to guide educational practices and interventions. This is especially important given that achievement and career choices at the end of school trajectories depend on developmental changes in children's motivation earlier in school (Ceci et al., 2014; Musu-Gillette et al., 2015). Negative or low domain-specific competence beliefs are usually associated with lower levels of effort and lesser perseverance in the face of challenge, as well as less ambitious goal setting, motivation, and interest (Denissen, Zarrett, & Eccles, 2007; Wigfield et al., 2015). Our results indicate that this development has its starting point in the early stages of competence beliefs formation and interventions might be most effective if they start that early. That the predominant direction among math competence beliefs and achievement differs at different ages has implications for potential interventions aiming at improving competence beliefs and achievement. Especially at the beginning of elementary school, it seems to be important to provide students with realistic, differentiated, and individual performance feedback to foster realistic self-evaluations. Moreover, for young children, for whom the skill development approach prevails, the most effective way in which to promote achievement and competence beliefs in math is to enhance achievement. If achievement is improved, math competence beliefs will become more positive as a by product of better achievement. When children get older and the self-enhancement approach becomes more important, interventions should also foster math competence beliefs because, as a consequence, achievement will become better, too. Because of reciprocal effects, competence beliefs and achievement in math more and more become a self-stabilizing system. Therefore, interventions for children after the initial phase of competence beliefs formation need to impact on both achievement and competence beliefs in math at the same time to be effective. When students have established domain-specific competence beliefs, the focus, for example, should be on their beliefs about the nature of ability (ability is mutable vs. unchangeable; see Blackwell, Trzesniewski, & Dweck, 2007) because, henceforth, (negative) effects of (low) competence beliefs on performance development are more likely. An underestimation of own abilities might be less detrimental when a child is convinced that ability can be altered.

References

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

Asparouhov, T., & Muthén, B. (2008). *Auxiliary variables predicting missing data. Technical appendix*. Los Angeles, CA: Muthén & Muthén.

Asparouhov, T., & Muthén, B. (2010). *Weighted least squares estimation with missing data. Technical report*. Los Angeles, CA: Authors.

Biesanz, J. C. (2012). Autoregressive longitudinal models. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 459–471). New York, NY: Guilford Press.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*, 246–263. doi:10.1111/j.1467-8624.2007.00995.x

Calysn, R. J., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievements? *Journal of Educational Psychology*, *69*, 136–145. doi:10.1037/0022-0663.69.2.136

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, *15*, 75–141. doi:10.1177/1529100614541236

Chapman, J. W., & Tunmer, W. E. (1997). A longitudinal study of beginning reading achievement and reading self-concept. *British Journal of Educational Psychology*, *67*, 279–291. doi:10.1111/j.2044-8279.1997.tb01244.x

Chen, S.-K., Yeh, Y.-C., Hwang, F.-M., & Lin, S. S. J. (2013). The relationship between academic self-concept and achievement: A multicohort–multioccasion study. *Learning and Individual Differences*, *23*, 172–178. doi:10.1016/j.lindif.2012.07.021

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255. doi:10.1207/S15328007SEM0902_5

Denissen, J. J. A., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development*, *78*, 430–447. doi:10.1111/j.1467-8624.2007.01007.x

Dweck, C. S. (2002). The development of ability conceptions. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 57–88). San Diego, CA: Academic Press.

Eccles, J. S., Wigfield, A., Harold, R., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*, 840–847. doi:10.2307/1131221

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

Federal Statistical Office. (2009). *Statistical yearbook 2009 for the federal republic of Germany*. Wiesbaden, Germany: Author.

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Finney, S. J., & DiStefano, C. (2013). Nonormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 269–314). Charlotte, NC: Information Age Publishing.

Frey, K. S., & Ruble, D. N. (1987). What children say about classroom performance: Sex and grade differences in perceived competence. *Child Development*, *58*, 1066–1078. doi:10.2307/1130547

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology*, *93*, 3–13. doi:10.1037//0022-0663.93.1.3

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling*, *10*, 80–100. doi:10.1207/S15328007SEM1001_4

Graham, J. W., & Coffman, D. L. (2012). Structural equation modeling with missing data. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 277–295). New York, NY: Guilford Press.

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*, 124–136. doi:10.1037/0022-0663.95.1.124

Harter, S. (1982). The perceived competence scale for children. *Child Development*, *53*, 87–97. doi:10.2307/1129640

Harter, S. (2006). The self. In N. Eisenberg (Vol. Ed.), *Handbook of child psychology* (Vol. III, 6th ed., pp. 505–570). New York, NY: Wiley.

Helmke, A., & van Aken, M. A. G. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school: A longitudinal study. *Journal of Educational Psychology*, *87*, 624–637. doi:10.1037/0022-0663.87.4.624

Herbert, J., & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Applied Developmental Psychology*, *26*, 276–295. doi:10.1016/j.appdev.2005.02.007

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology*, *49*, 505–528. doi:10.1016/j.jsp.2011.07.001

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

Kurtz-Costes, B. E., & Schneider, W. (1994). Self-concept, attributional belief, and school achievement: A longitudinal analysis. *Contemporary Educational Psychology*, *19*, 199–216. doi:10.1006/ceps.1994.1017

Little, T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford Press.

Ma, X., & Johnson, W. (2008). Mathematics as the critical filter: Curricular effects on gendered career choices. In H. M. G. Watt & J. S. Eccles (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 55–83). Washington, DC: American Psychological Association.

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology*, *82*, 646–656. doi:10.1037/0022-0663.82.4.646

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and revised recommendations. *Educational Psychologist*, *34*, 155–167. doi:10.1207/s15326985ep3403_2

Marsh, H. W., Craven, R. G., & Debus, R. (1991). Self-concepts of young children 5 to 8 years of age: Measurement and multidimensional structure. *Journal of Educational Psychology*, *83*, 377–392. doi:10.1037/0022-0663.83.3.377

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

Marsh, H. W., Gerlach, E., Trautwein, U., Lüdtke, O., & Brettschneider, W.-D. (2007). Longitudinal study of preadolescent sport self-concept and performance: Reciprocal effects and causal ordering. *Child Development*, *78*, 1640–1656. doi:10.1111/j.1467-8624.2007.01094.x

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*, 59–77. doi:10.1348/000709910X503501

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*, 397–416. doi:10.1111/j.1467-8624.2005.00853.x

Meredith, W. M. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, *58*, 525–543. doi:10.1007/BF02294825

Möller, J. (2005). Paradoxical effects of praise and criticism: Social, dimensional, and temporal comparisons. *British Journal of Educational Psychology*, *75*, 275–295. doi:10.1348/000709904X24744

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, *79*, 1129–1167. doi:10.3102/0034654309337522

Musu-Gillette, L., Wigfield, A., Harring, J., & Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educational Research and Evaluation*, *21*, 343–370. doi:10.1080/13803611.2015.1057161

Muthén, L. K., & Muthén, B. O. (1998-2011). *Mplus (Version 6.12)* [Computer software]. Los Angeles, CA: Authors.

Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

intrinsic value. *Educational Research and Evaluation*, *12*, 323–345. doi:10.1080/13803610600765687

Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Development*, *49*, 800–814. doi:10.2307/1128250

Niepel, C., Brunner, M., & Preckel, F. (2014a). Achievement goals, academic self-concept, and school grades in mathematics: Longitudinal reciprocal relations in above average ability secondary school students. *Contemporary Educational Psychology*, *39*, 301–313. doi:10.1016/j.cedpsych.2014.07.002

Niepel, C., Brunner, M., & Preckel, F. (2014b). The longitudinal interplay of students' academic self-concepts and achievements within and across domains: Replicating and extending the reciprocal internal/external frame of reference model. *Journal of Educational Psychology*, *106*, 1170–1191. doi:10.1037/a0036307

Pinxten, M., de Fraine, B., van Damme, J., & D'Haenens, E. (2010). Causal ordering of academic self-concept and achievement: Effects of type of achievement measure. *British Journal of Educational Psychology*, *80*, 689–709. doi:10.1348/000709910X493071

Ruble, D. N., Parsons, J. E., & Ross, J. (1976). Self-evaluative responses of children in an achievement setting. *Child Development*, *47*, 990–997. doi:10.2307/1128435

Satorra, A., & Bentler, P. M. (2001). A scaled chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. doi:10.1007/BF02296192

Schwarz-Jung, S. (2013). Nach dem Wegfall der verbindlichen Grundschulempfehlung: Übergänge auf weiterführende Schulen zum Schuljahr 2012/13 [After the omission of the binding Basic school recommendation: Transitions on Secondary schools for the school year 2012/13]. *Statistisches Monatsheft Baden-Württemberg*, *4*, 22–27.

Skaalvik, E. M., & Hagtvet, K. A. (1990). Academic achievement and self-concept: An analysis of causal predominance in a developmental perspective. *Journal of Personality and Social Psychology*, *58*, 292–307. doi:10.1037/0022-3514.58.2.292

Skaalvik, E. M., & Skaalvik, S. (2002). Internal and external frames of reference for academic self-concept. *Educational Psychologist*, *37*, 233–244. doi:10.2466/pr0.94.2.619-632

Skaalvik, E. M., & Skaalvik, S. (2008). Self-concept and self-efficacy in mathematics: Relation with mathematics motivation and achievement. In F. M. Olsson (Ed.), *New developments in the psychology of motivation* (pp. 105–128). Hauppauge, NY: Nova Science.

Skaalvik, E. M., & Valas, H. (1999). Relations among achievement, self-concept, and motivation in mathematics and language arts: A longitudinal study. *The Journal of Experimental Education*, *67*, 135–149. doi:10.1080/00220979909598349

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

Spinath, B., & Spinath, F. M. (2005b). Longitudinal analysis of the link between learning motivation and competence beliefs among elementary school children. *Learning and Instruction*, *15*, 87–102. doi. 10.1016/j.learninstruc.2005.04.008

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

Stipek, D. (2011). Comment: The logic and meaning of declining perceptions of academic competence beliefs. In F. E. Weinert & W. Schneider (Eds.), *Individual development from 3 to 12: Findings from the Munich longitudinal study* (pp. 222–226). New York, NY: Cambridge University Press.

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

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*, 111–133. doi:10.1207/s15326985ep3902_3

Viljaranta, J., Tolvanen, A., Aunola, K., & Nurmi, J.-E. (2014). The developmental dynamics between interest, self-concept of ability, and academic performance. *Scandinavian Journal of Educational Research*, *58*, 734–756. doi:10.1080/00313831.2014.904419

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York, NY: Guilford Press.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, *25*, 68–81. doi:10.1006/ceps.1999.1015

Wigfield, A., Eccles, J. S., Fredricks, J., Simpkins, S., Roeser, R., & Schiefele, U. (2015). Development of achievement motivation and engagement. In R. Lerner (Series ed.), M. Lamb & C. Garcia Coll (Vol. Eds.), *Handbook of child psychology and developmental science* (Vol. 3, 7th ed., pp. 657–700). New York, NY: Wiley.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A., Freedman-Doan, K., & Blumenfeld, P. C. (1997). Changes in children's competence beliefs and subjective task values across the elementary school years: A three-year study. *Journal of Educational Psychology*, *89*, 451–469. doi:10.1037/0022-0663.89.3.451

Wigfield, A., & Karpathian, M. (1991). Who am I and what can I do? Children's self-concepts and motivation in achievement situations. *Educational Psychologist*, *26*, 233–261. doi:10.1207/s15326985ep2603&4_3

Wu, W., & West, S. G. (2010). Sensitivity of fit indices to misspecification in growth curve models. *Multivariate Behavioral Research*, *45*, 420–452. doi:10.1080/00273171.2010.483378

14678624, 2018, 2, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12806 by Turkey Cochrane Evidence Aid, Wiley Online Library on [02/10/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License