British Journal of Educational Psychology (2017)

The British Psychological Society

www.wileyonlinelibrary.com

Math grades and intrinsic motivation in elementary school: A longitudinal investigation of their association

Anne F. Weidinger[1a*], Ricarda Steinmayr[1a] and Birgit Spinath[2]
[1]Technical University Dortmund, Germany
[2]Heidelberg University, Heidelberg, Germany

Background. It is often argued that the negative development of intrinsic motivation in elementary school strongly depends on the presence of school grades because grades represent extrinsic consequences and achievement feedback that are supposed to influence intrinsically motivated behaviour. However, only a few studies have tested this hypothesis.

Aims. Therefore, we investigated the role of school grades in inter- and intra-individual changes in elementary school students' intrinsic motivation from when grades were first introduced until the end of elementary school, when students in Germany receive recommendations for a secondary school type on the basis of their prior performance in school.

Sample. A sample of 542 German elementary school students (t_1: M = 7.95 years, SD = 0.57) was followed for 2 years from the end of Grade 2 to the end of Grade 4.

Methods. At seven measurement occasions, children's math grades and their domain-specific intrinsic motivation were assessed.

Results. Latent growth curve models showed differences in trajectories of intrinsic motivation across students rather than uniform development. Moreover, students' trajectories of grades and intrinsic motivation were only weakly associated. A latent cross-lagged model revealed that reciprocal effects between the two constructs over time were small at best.

Conclusions. Contrary to theoretical considerations, our results indicate that negative performance feedback in the form of grades does not necessarily lead to a decrease in intrinsic motivation. This calls into question the common opinion that a perception of being less competent, as reflected by poor grades, is responsible for weakening students' intrinsic motivation.

Students' school-related intrinsic motivation is commonly regarded as a main determinant of academic achievement, engagement, and school functioning in general (e.g., Corpus & Wormington, 2014; Wigfield *et al.*, 2015). Examining intrinsic motivation (IM) as early as in elementary school is important because IM in the early school years may have far-reaching consequences for initial and future school success (e.g., Gottfried, 1990; Musu-Gillette, Wigfield, Harring, & Eccles, 2015). Research on the development of students'

[a]The first two authors contributed equally.

**Correspondence should be addressed to Anne F. Weidinger, Emil-Figge-Straße 50, Dortmund, Germany (email: anne.weidinger@tu-dortmund.de).*

DOI:10.1111/bjep.12143

school-related IM has consistently documented that, although very positive at the beginning of elementary school, IM quickly begins to decline and continues to do so until the end of compulsory education (e.g., Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002; Spinath & Steinmayr, 2012). Therefore, researchers as well as educators are interested in the underlying mechanisms that lead to the weakening of learning-related motivation in order to find ways to support or reactivate the initially high motivation for learning that is found in children at the beginning of their school trajectories.

A great variety of reasons have been discussed when it comes to explaining the observed decline (for an overview, see Wigfield *et al.*, 2015). It is often argued that the negative development of IM in elementary school strongly depends on the presence of school grades (e.g., Deutsch, 1979) as grades represent extrinsic consequences (see *self-determination theory*; Ryan & Deci, 2000) and previous achievement outcomes (see *expectancy-value model*; Eccles & Wigfield, 2002) that should influence intrinsically motivated behaviour. Indeed, a large body of research has steadily documented relations between IM and grades (e.g., Corpus, McClintic-Gilbert, & Hayenga, 2009; Gottfried, 1990). To clarify the role of grades for change in students' IM, however, longitudinal studies are needed. Moreover, it is crucial to assess the constructs several times a year (see Viljaranta, Tolvanen, Aunola, & Nurmi, 2014). First, children receive grades not only at the end but also throughout the school year, and the relation between grades and IM could vary within a school year. Second, because of changes in the school environment throughout a student's school career, it is likely that the importance of grades – and thus their relation to IM – varies across school years. More precisely, grades should be of great relevance right after their implementation (usually at the end of Grade 2 in German schools). Furthermore, in tracked academic school systems, grades are especially important at the end of elementary school when students receive recommendations for a specific secondary school type on the basis of their prior performance in school. Investigating children in the period when grades are highly salient is thus a promising approach for broadening knowledge about the developmental dynamics of grades and IM.

We are aware of one longitudinal study in which elementary school students' IM and grades were assessed more than twice a year, and the reciprocal relation between them was investigated. However, Weidinger, Spinath, and Steinmayr (2015) found only minor effects of grades on change in students' IM in the domain of German. It remains an open question whether they found only weak relations because they did not consider grades immediately after their implementation. Moreover, the findings should be cross-validated in other domains. Therefore, we investigated whether and how changes in students' IM are related to students' grades in math right after grades are implemented in elementary school.

Intrinsic motivation – definition and development

IM is defined as intrinsic task value according to the *expectancy-value model* (Eccles & Wigfield, 2002). Accordingly, an activity has high intrinsic value and, thus, is understood as intrinsically motivated when it is performed for no sake but its own reward, and the activity is accompanied by positive emotions and is perceived as highly satisfying (Wigfield & Eccles, 1992). Even though intrinsic task values are not the only reason for learning, affective liking and enjoyment can be considered to be the most desirable states for learners because learning happens as a side effect of engaging in a pleasurable activity (see White, 1959). Research has found that children's intrinsic task values are domain-specific even in the first grade (e.g., Eccles, Wigfield, Harold, & Blumenfeld, 1993). In particular, for elementary school-aged children, intrinsic task values play an outstanding role as these children do not yet distinguish other task values such as utility or importance (e.g., Eccles *et al.*, 1993; Wigfield & Eccles, 1992).

At the beginning of elementary school, students' IM is usually very high, but it declines over the school years (e.g., Gottfried, Fleming, & Gottfried, 2001; Jacobs *et al.*, 2002; Spinath & Spinath, 2005; Wigfield *et al.*, 1997). Previous studies on the development of IM in the verbal and math domain in elementary school found a linear motivational decline (Aunola, Leskinen, & Nurmi, 2006; McElvany, Kortenbruck, & Becker, 2008). The question of why students lose their IM for school-related activities has yet to be satisfactorily addressed in empirical research. With regard to suppositions made by motivation theories, we investigated whether and how the onset of grades and their trajectories in elementary school are associated with changes in IM.

Why and how are grades supposed to affect intrinsic motivation?

Researchers have postulated that an increasingly extrinsic and controlling school atmosphere, an increase in achievement pressure and competition between students, and a progressive negative effect of extrinsic consequences on IM contribute to the decline in IM with age (e.g., Gottfried *et al.*, 2001; Harter, 1981; Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006). School grades, 'the basic currency of our educational system' (Deutsch, 1979, p. 393), might play a decisive role in this development. Because of their symbolic value, grades represent performance feedback and enable a direct social comparison between children (see Stipek & Mac Iver, 1989).

Regarding the potential impact of grades on students' IM, the *cognitive evaluation theory* (CET; Deci & Ryan, 1985; Ryan & Deci, 2000) provides a useful framework. This subtheory within *self-determination theory* posits that IM in achievement settings depends on the fulfilment of two primary needs: the need for competence and the need for autonomy. According to CET, external incentives (e.g., tangible rewards, pressured evaluations, directives) undermine IM because individuals perceive their behaviour as externally regulated instead of self-determined (*undermining effect*; Deci & Ryan, 1985). However, experimental studies have shown that this undermining effect does not occur when the external incentive comprises positive competence feedback (for an overview, see Deci, Koestner, & Ryan, 1999). It was argued that if the informative aspect of the positive feedback outweighed the controlling aspect of the external incentive, IM should not diminish because the positive competence feedback should satisfy the person's need for competence (see Deci & Ryan, 1985). Applying these considerations to the educational context, school grades can be seen as an external incentive (Deutsch, 1979) that comprises competence-related information. Students with good grades receive positive performance feedback, which should satisfy their need for competence so that they should not lose their IM. By contrast, students with poor grades receive negative performance feedback, which should undermine feelings of competence and therefore lead to a decrease in IM. Vice versa, it is also plausible that IM influences later achievement, for example, mediated via task choice (Eccles & Wigfield, 2002).

Some studies on older students found at least small effects of grades on changes in IM and of IM on changes in grades (Guay, Ratelle, Roy, & Litalien, 2010; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Wang, 2012). Moreover, the effect of students' first report card grades in secondary school on changes in students' school engagement was mediated via their affective reactions to their grades (Poorthuis *et al.*, 2015). Although effects on IM were not tested, negative affect elicited by poor grades should be incompatible with IM as IM is characterized by positive feelings during task processing (White, 1959). We are aware of only two studies that longitudinally investigated the reciprocal relation between school grades and IM in elementary school while controlling for prior grades and motivation. Using a series of hierarchical regressions, Corpus *et al.* (2009) found small reciprocal effects between grades and school-related IM over 1 school year in six cohorts (third to eighth grade), whereby analyses were not performed separately for each cohort. Thus, no specific conclusions can be drawn for elementary school children. Weidinger *et al.* (2015) found only minor effects of students' grades in German on changes in domain-specific IM, whereby both constructs were assessed six times at an interval of 4 months between Grades 3 and 4. The effects might have been somewhat stronger if the constructs had been assessed at the end of Grade 2 when the children received grades for the first time in their school career. If there is an effect of grades on IM, the chances of finding such an effect in this period of education are very high due to the pronounced importance of grades. Moreover, it is necessary to cross-validate findings in another domain. These limitations should be addressed in this study.

The present investigation

Theory suggests that grades have an impact on IM and, accordingly, are responsible for the often-found motivational decline in elementary school. We are not aware of any longitudinal study that investigated the impact of grades on change in IM in elementary school immediately after the onset of grades. To fill this gap and expand knowledge on the determinants of school-related IM, we investigated children's math grades and their domain-specific IM from Grade 2 to Grade 4. Data stem from the same project as the data analysed in the study by Weidinger *et al.* (2015). However, we focused on another domain, namely math, and on a larger time period that included the second grade. Given the above-mentioned theoretical considerations, this is a useful and necessary extension. The following hypotheses were addressed:

Hypothesis 1: On average, students' IM in math will show a linear decline from the end of Grade 2 to the end of Grade 4 (1a), and there will be significant interindividual differences in the decline over the investigated time period (1b).

Hypothesis 2: The change in students' math grades will be related to the change in students' IM in math: students whose grades deteriorate will show a steeper decline in IM over the investigated time period.

Hypothesis 3: The relation between math grades and IM in math will be reciprocal, so that grades will predict changes in IM (3a), and IM will predict changes in grades (3b). Theoretical considerations point to a reciprocal relation, whereas previous findings call it into question (Weidinger *et al.*, 2015).

Because children receive grades not only at the end but also throughout the school year, effects might be masked or misinterpreted when grades and IM are assessed only at the beginning and at the end of a school year (Corpus *et al.*, 2009). Therefore, we explored whether the effects would vary according to the frequency of the assessment of the construct within a school year (i.e., the time interval between measurements).

Method

Sample

A total of 542 students (270 boys and 272 girls) participated in a longitudinal study on the development of motivation during elementary school (MEGA; MotivationsEntwicklung im GrundschulAlter). Students came from 27 classes from 11 German elementary schools in and around two mid-sized towns in southern Germany. Only students whose parents provided informed consent participated in the study (75% of the target sample). The study began in July 2009 and comprised seven measurement occasions at an interval of 4 months each (see Table 1). The first measurement occasion (t_1) took place at the end of Grade 2 when children were on average 7.95 years old (SD = 0.58). At the last measurement occasion (t_7), the children were approximately 2 years older (M = 9.93, SD = 0.72). Weidinger *et al.* (2015) analysed data from the same students and teachers, but focused on another domain for which data was assessed only from t_2 onwards.

Procedure

The study took place during a regular math class (45 min). Students were asked to provide self-reports of their IM in math (among other variables). All items were read aloud to ensure that all of the students worked at the same speed. While students were working on their questionnaires, teachers filled in a questionnaire in which they indicated, interalia,

Table 1. Latent means (*M*) and standard deviations (*SD*), Reliabilities (McDonald's Omega, ω), Intraclass correlations (ICC), and Sample sizes (*n*) for Intrinsic motivation and math grades at seven measurement occasions (*t*)

	t_1 2009–07	t_2 2009–11	t_3 2010–03	t_4 2010–07	t_5 2010–11	t_6 2011–03	t_7 2011–07
	Grade 2	Grade 3			Grade 4		
Intrinsic motivation[a]							
M	3.945	3.972	3.929	3.786	3.860	3.804	3.647
SD	0.994	0.865	0.911	0.993	0.970	0.890	0.968
ω	.834	.832	.919	.844	.865	.870	.872
ICC	.025	.030	.018	.056	.050	.027	.032
n	458	468	462	468	451	449	446
Math grades[b]							
M	4.766	4.717	4.745	4.607	4.593	4.595	4.568
SD	0.881	0.922	0.900	1.006	0.951	0.924	0.911
ICC	.143	.147	.192	.163	.114	.116	.101
n	406	435	452	433	461	418	454

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

students' math grades. The students of each class had the same math teacher in Grade 2, 3, and 4.

Measures

Intrinsic motivation in math

The three items measuring children's IM in math were established in previous studies (e.g., Spinath & Steinmayr, 2008) and are based on the questionnaire to assess intrinsic values according to the *expectancy-value model* (Wigfield *et al.*, 1997). The following items were read to the children: 'How much do you like mental arithmetic/solving math problems/doing calculations?' Response choices ranged from 1 (*really a lot*) to 5 (*not at all*). To estimate the reliability of the scale, we calculated McDonald's Omega (ω), which varied from $\omega = .832$ to $\omega = .919$ (see Table 1).

Math grades

Teachers were asked to provide students' math grades for each measurement occasion ('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)'). Grades in elementary school rely heavily on performance on written tests, which are taken several times during the school year, but the quality of oral participation in class is also considered. The math grades considered in our analyses thus represented an average measure of students' preceding performance (see Niepel, Brunner, & Preckel, 2014), and they were based on the grades that the students received during the school year. Previous studies showed that students are well aware of their math grades (Dickhäuser & Plenter, 2005) and that even in elementary school correlations between actual math grades and those reported by the students are strong (Schneider & Sparfeldt, 2016). In Germany, grades range from 1 (*excellent*) to 6 (*insufficient/fail*).

For a more intuitive understanding, the IM scale and grades were recoded so that higher scores denoted a higher IM and better grades, respectively.

Analysis

All analyses were computed in SPSS 22.0 and M*plus* 6.12 (Muthén & Muthén, 1998–2011). As students were nested within classes, maximum likelihood estimation with robust standard errors (MLR) was used for parameter estimation ('type = complex' option in M*plus* with class membership as a cluster variable). Latent analyses were based on data from 540 students (two students were missing class membership).

Missing data

As is common in longitudinal studies, the current investigation had to deal with missing data related to attrition and non-response (see Little, 2013). Missing data related to attrition occurred when children missed one or more measurement occasions, mainly because of illness. Missing data resulting from non-response occurred at each measurement occasion and ranged from 0.0% (t_1, t_2, and t_3) to 0.6% (t_4) for intrinsic motivation and from 0.2% (t_5) to 11.5% (t_1) for grades. Complete data sets for all seven measurement occasions were available for 264 students (48.7%), whereas 490 students (90.1%) participated in at least five measurement occasions. Table 1 contains the number of students participating at each measurement. To maximize the power of the analyses, we accounted for missing data by applying full information maximum likelihood (FIML) estimations (Graham & Coffman, 2012; Little, 2013). To make the assumption of missing-at-random more plausible, we considered a set of variables as missing data correlates (i.e., the teacher ratings of children's math ability) in addition to the analysis variables (Graham, 2003; AUXILIARY option in M*plus*, Asparouhov & Muthén, 2008). Further information on the auxiliary variables, which are not of substantive interest here, can be requested from the corresponding author.

Longitudinal confirmatory factor analysis (CFA) modelling

In a preliminary analysis, we tested whether strong factorial invariance of IM measures over time was given (i.e., time-invariant factor loadings and intercepts, Meredith, 1993). Therefore, we estimated a longitudinal CFA model with IM 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). To evaluate invariance constraints, we calculated the change in CFI (ΔCFI) as here the scaled chi-square (χ^2) difference test is too sensitive to trivial differences (Little, 2013). Following guidelines, the supposition of invariance is tenable if $\Delta CFI \leq .01$ (Cheung & Rensvold, 2002).

To estimate latent means and standard deviations of IM as well as the latent correlations between IM and grades 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 IM (Little, 2013).

Latent growth curve modelling (LGCM)

Latent growth curve models were used to examine the developmental changes in students' IM and grades in math (Hypotheses 1 and 2). This method allowed us to describe an overall pattern of change in the construct of interest over time and individual differences in these trajectories (Duncan & Duncan, 2004). First, two linear LGCMs (for IM and grades) were estimated. IM was considered as first-order factor from t_1 to t_7 indicated by three items each time (see CFA model). The latent intercept and slope were modelled as second-order factors. To model a change, the loadings of the first-order factors on the intercept were set to 1. The slope factor loadings from t_1 to t_7 (γ_{t1}–γ_{t7}) were fixed to [0 1 2 3 4 5 6] to estimate a linear development. Thus, the mean of the slope factor and its algebraic sign indicate the overall magnitude of the change and its direction, whereas the slope variance indicates interindividual differences in change (Hypothesis 1). We used the same approach to model change in math grades, except that grades were included as manifest variable at each time point and the growth factors were first-order factors.

Second, the two LGCMs were combined in a bivariate LGCM, and the correlations between the latent intercept and slope factors were estimated. Thus, it was possible to examine the co-occurrence of age-related changes in grades and IM (Hypothesis 2).

Latent cross-lagged modelling (CLM)

We used bivariate latent cross-lagged models to investigate the relation between grades and IM over time while controlling for prior values on these constructs (Hypothesis 3). First, a CLM with IM and grades assessed at seven time points was estimated (4-Month

Interval Model). Therefore, the cross-time non-directional associations of the CFA model were converted into direct paths (i.e., autoregressive and cross-lagged paths between adjacent times of measurement). Only the within-time associations were still modelled as non-directional covariance relations. Second, we estimated a multivariate CLM with cross-lagged and autoregressive paths between IM and grades considering only three measurement occasions (t_1, t_4, t_7) at an interval of 1 year (1-Year Interval Model) to examine whether the pattern of results would be the same.

To evaluate the goodness of fit of all latent models, we computed the χ^2 test statistic, the comparative fit index (CFI), the root mean square error of approximation (RMSEA) along with its associated confidence interval, and the standardized root mean square residual (SRMR). We used the following cut-off values to evaluate goodness of fit: CFI > .95, RMSEA < .06, and SRMR < .08 (see West, Taylor, & Wu, 2012; Wu & West, 2010).

Results

Measurement invariance, descriptive statistics, and correlations

The preliminary analyses demonstrated strong factorial invariance in the IM measures over time (see Table 2). Thus, in all subsequent analyses, strong longitudinal invariance was specified.

Table 1 contains the latent means and standard deviations from the longitudinal CFA model with strong invariance specified. On average, children's IM was relatively high at all measurement occasions. Moreover, Table 1 depicts the intraclass correlations of IM and grades, as well as the reliabilities of the IM scales, which were high. The correlations from the CFA model with IM and grades are depicted in Table 3. Test–retest correlations were high and statistically significant.

Developmental changes in intrinsic motivation and math grades

Results of the linear LGCMs are presented in Table 4. Both models showed at least a satisfactory fit to the data. The latent mean of the slope of the IM model was negative and significantly different from zero ($\mu_{\text{Slope}} = -0.050$, $SE = 0.010$, $p < .001$). In line with Hypothesis 1a, students' IM in math thus underwent a linear change for the worse from the end of Grade 2 to the end of Grade 4. The variance of the slope was significantly different from zero ($\psi_{\text{Slope}} = 0.015$, $SE = 0.002$, $p < .001$), indicating that there were substantial interindividual differences among students in motivational decline (Hypothesis 1b).

The same was true for math grades. The negative latent mean ($\mu_{\text{Slope}} = -0.034$, $SE = 0.008$, $p < .001$) and the variance of the slope ($\psi_{\text{Slope}} = 0.006$, $SE = 0.001$, $p < .001$) were statistically significant. On average, students' grades changed for the worse. However, there were significant interindividual differences in the change in grades.

To investigate whether the trajectories of students' IM were related to their math grade trajectories (Hypothesis 2), we inspected the correlation between the slopes of the constructs in the bivariate LGCM, which showed a good fit to the data, $\chi^2_{(381)} = 489.503$, $p < .001$, CFI = .978, RMSEA = .032, 90% CI [.026, .037], SRMR = .065. Contrary to Hypothesis 2, the slopes were only weakly and not significantly correlated ($r = .227$, $SE = 0.131$, $p = .082$). Therefore, the downward trend in IM was only weakly associated with a decline in students' math grades.

Table 2. Model fit statistics for testing the longitudinal measurement invariance of intrinsic motivation

Model tested	χ^2 (*df*)	*p*	$\Delta\chi^2_{\mathrm{corr}}$ (*df*)	*p*	RMSEA	RMSEA 90% CI	SRMR	CFI	ΔCFI	Pass?
Configurally invariant model	149.985 (105)	.003	–	–	.028	.017, .038	.034	.991	–	–
Weak invariant model	165.368 (117)	.002	14.927 (12)	.245	.028	.017, .037	.036	.991	.000	Yes
Strong invariant model	187.707 (129)	.001	21.898 (12)	.039	.029	.019, .038	.036	.989	.002	Yes

Note. $N = 540$. Intrinsic motivation was assessed at seven measurement occasions with three items. Results are based on the effects coded method of identification. The residual variances among the corresponding indicators were allowed to correlate over time. Maximum likelihood estimation with robust standard errors (MLR) was used. χ^2 (*df*) = chi-square test statistic with degrees of freedom in parentheses; $\Delta\chi^2_{\mathrm{corr}}$(*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.

Table 3. Correlations between students' intrinsic motivation (IM) and math grades (Gr) at all measurement occasions (t)

	IM t_2	IM t_3	IM t_4	IM t_5	IM t_6	IM t_7	Gr t_1	Gr t_2	Gr t_3	Gr t_4	Gr t_5	Gr t_6	Gr t_7
IM t_1	.682	.550	.485	.507	.470	.473	.196	.135	.109	.152	.173	.151	.162
IM t_2		.671	.668	.592	.546	.549	.253	.266	.199	.238	.210	.203	.223
IM t_3			.742	.715	.675	.662	.282	.241	.200	.255	.226	.236	.245
IM t_4				.844	.751	.746	.275	.251	.212	.265	.266	.245	.204
IM t_5					.776	.775	.261	.230	.186	.253	.290	.253	.257
IM t_6						.856	.256	.270	.207	.250	.298	.289	.287
IM t_7							.285	.314	.261	.306	.371	.334	.326
Gr t_1								.712	.648	.713	.688	.703	.689
Gr t_2									.785	.778	.738	.735	.737
Gr t_3										.800	.748	.739	.735
Gr t_4											.772	.810	.807
Gr t_5												.875	.823
Gr t_6													.874
Gr t_7													–

Note. Results are from the longitudinal CFA model with intrinsic motivation and grades. For $rs \geq |.109|$, $p < .05$; for $rs \geq |.151|$, $p < .01$; and for $rs \geq |.186|$, $p < .001$.

Reciprocal relation between intrinsic motivation and math grades over time

To investigate the reciprocal relation between IM and math grades (Hypothesis 3), we evaluated a CLM that included autoregressive and cross-lagged paths between the constructs (4-Month Interval Model; see Figure 1). Goodness-of-fit indices indicated an acceptable model fit. There were significant small concurrent correlations only at t_1, t_2, and t_5 (see Figure 1). Contrary to Hypothesis 3a, grades had significant small effects on changes in IM only from t_1 to t_2 ($\beta = .130$, $SE = 0.036$, $p < .001$) and from t_6 to t_7 ($\beta = .079$, $SE = 0.027$, $p = .003$). Moreover, except for a small significant effect from t_3 to t_4 ($\beta = .105$, $SE = 0.034$, $p = .002$), IM did not predict any changes in students' math grades (Hypothesis 3b). Thus, we found no substantial reciprocal effects between math grades and IM.

To examine whether the pattern of results changed when only three measurement occasions at an interval of 1 year (t_1, t_4, and t_7) were incorporated into the model, we evaluated a 1-Year Interval Model that included autoregressive and cross-lagged paths between the constructs (see Figure 2). Goodness-of-fit indices indicated a good model fit. Whereas both cross-lagged paths from grades to students' IM were small but significant, we did not find significant cross-lagged paths from IM to grades (see Figure 2). Thus, there were small substantial effects of math grades on change in IM over 1 year, whereas the effects of grades on changes in IM over 4 months were negligible.

Discussion

The primary focus of the current investigation was to shed further light on the development of students' IM in math in the early school years right after the implementation of grades as grades are often considered to be responsible for students' motivational decline (see Gottfried *et al.*, 2001; Wigfield *et al.*, 2006). The results showed

Table 4. Unconditional latent growth curve model considering students' intrinsic motivation, and math grades, respectively

	Second-order model: Intrinsic motivation		First-order model: Math grades	
	Parameter (*SE*)[a]		Parameter (*SE*)[a]	
Slope factor loadings	γ_{t1}	0.00	λ_{t1}	0.00
	γ_{t2}	1.00	λ_{t2}	1.00
	γ_{t3}	2.00	λ_{t3}	2.00
	γ_{t4}	3.00	λ_{t4}	3.00
	γ_{t5}	4.00	λ_{t5}	4.00
	γ_{t6}	5.00	λ_{t6}	5.00
	γ_{t1}	6.00	λ_{t7}	6.00
Variances				
$\psi_{\text{Intercept}}$	0.606 (0.064)***		0.630 (0.085)***	
ψ_{Slope}	0.015 (0.002)***		0.006 (0.001)***	
Covariance				
$\psi_{\text{Intercept, Slope}}$	−0.025 (0.011)*		−0.004 (0.008)	
Latent means				
$\mu_{\text{Intercept}}$	4.004 (0.050)***		4.762 (0.071)***	
μ_{Slope}	−0.050 (0.010)***		−0.034 (0.008)***	
Model fit				
χ^2 (*df*)	266.760 (152)***		54.495 (23)***	
CFI	.978		.973	
RMSEA [90% CI]	.037 [.030, .045]		.051 [.033, .068]	
SRMR	.057		.114	

Note. Unstandardized solution. t = measurement occasion; γ = structural coefficient; λ = loading coefficient; χ^2 (*df*) = chi-square test statistic with degrees of freedom in parentheses; CFI = comparative fit index; RMSEA = root mean square error of approximation along with its associated confidence interval; SRMR = standardized root mean square residual.
[a]Maximum likelihood parameter estimates with standard errors and a chi-square test statistic that are robust to non-independence of observations (MLR) were used.
*$p < .05$; ***$p < .001$.

that there were substantial differences in trajectories between students instead of a uniform decline in IM. Contrary to expectations, a change in math grades was only weakly associated with a change in IM. Furthermore, effects of grades on change in IM were small at best.

Grades and students' intrinsic motivation trajectories

Consistent with the literature, our analyses showed a significant linear decrease in math-related IM (e.g., Aunola *et al.*, 2006; Spinath & Spinath, 2005; Wigfield *et al.*, 1997). There were also significant differences in trajectories across students, suggesting that the downward trend in IM was more pronounced for some children and less so for others (see Spinath & Steinmayr, 2008). The finding that the development of students' math grades was only weakly related to the decline in students' IM is in line with what was found for the domain of German in elementary school (Weidinger *et al.*, 2015). That we replicated prior findings in another domain even when considering the constructs right after the

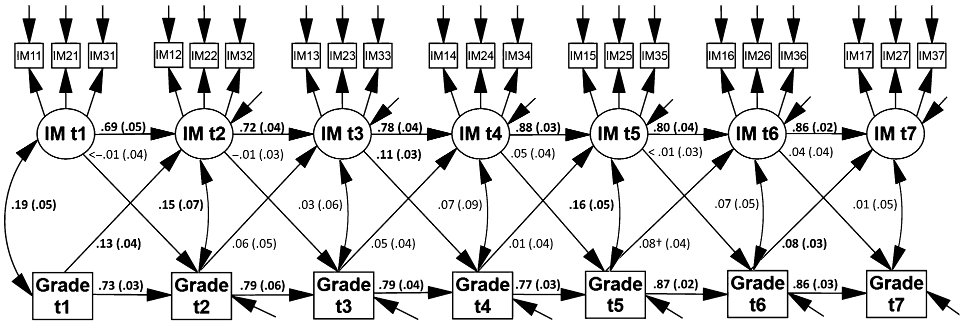

Figure 1. 4-Month Interval Model: Latent cross-lagged model of math-related intrinsic motivation (IM) and math grades (Grade) assessed at seven measurement occasions (t). N = 540. Standardized solution (standard errors in parentheses). Concurrent correlations between IM and Grade at t_2–t_7 represent residual correlations between the endogenous variables. Factor loadings and correlated uniquenesses between all corresponding IM measures collected at subsequent measurement occasions are not shown. Factor loadings and intercepts of the manifest IM indicators are invariant across time. Bold coefficients are significant ($p < .05$). Model fit: $\chi^2_{(287)} = 696.562$; $p < .001$; CFI = .947; RMSEA = .051, 90% CI [.047, .056]; SRMR = 0.090. †p = .067.

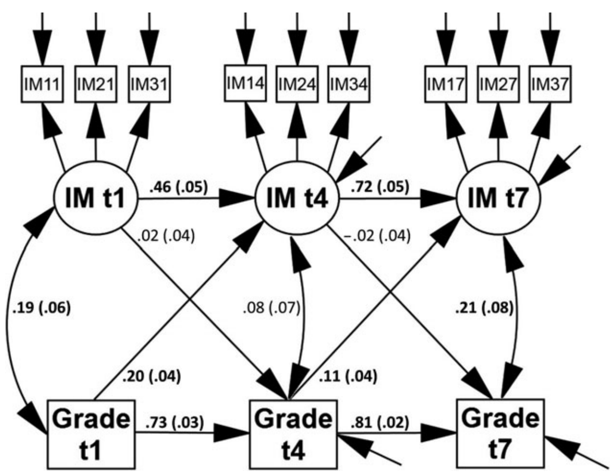

Figure 2. 1-Year Interval Model: Latent cross-lagged model of math-related intrinsic motivation (IM) and math grades (Grade) assessed at three measurement occasions (t_1, t_4, t_7). N = 540. Standardized solution (standard errors in parentheses). Concurrent correlations between IM and Grade at t_4 and t_7 represent residual correlations between the endogenous variables. Factor loadings and correlated uniquenesses between all corresponding IM measures collected at subsequent measurement occasions are not shown. Factor loadings and intercepts of the manifest IM indicators are invariant across time. Bold coefficients are significant ($p < .05$). Model fit: $\chi^2_{(45)} = 93.742$; $p < .001$; CFI = .980; RMSEA = .045, 90% CI [.032, .058]; SRMR = 0.050.

implementation of grades is especially remarkable and extends extant research. However, the finding contradicts the supposition that extrinsic consequences and performance feedback in the form of grades undermine intrinsically motivated behaviour (see Deci & Ryan, 1985; Eccles & Wigfield, 2002).

An explanation for the weak association between change in grades and change in IM could be that, on average, there was only weak change in grades over time and only weak interindividual differences in change. One explanation for the relatively stable mean levels of grades might be that grading – at least to some extant – is carried out in reference to normative criteria (*grading on the curve*; e.g., Guskey, 2000). If this result can be replicated in samples from other countries as well, it will be even more unlikely that changes in grades are a source of changes in IM. One might argue that even if it is not the change in grades that affects children's IM, it is still possible that unchanging poor grades are detrimental for IM. However, the latter is unlikely considering the weak concurrent relations between IM and grades found in the present study.

The results from the cross-lagged models also indicate that effects of grades on change in children's math-related IM are, if existent at all, very small in size. Again, our results mirror the findings from the domain of German (Weidinger *et al.*, 2015). Moreover, our results match findings from studies that investigated effects of teachers' ratings of children's performance (Skaalvik & Valas, 1999; Viljaranta *et al.*, 2014) and standardized test performance (Garon-Carrier *et al.*, 2016) on children's IM in math, as these studies also found no substantial effects at the end of elementary school. It is not only the high relative stabilities of grades and IM that account for our results as studies with comparable construct stabilities found at least small effects of prior achievement on the later IM of older children (i.e., from the end of elementary school onwards; e.g., Corpus *et al.*, 2009; Pinxten, Marsh, De Fraine, Van Den Noortgate, & Van Damme, 2014). Furthermore, influences linking IM and performance before the introduction of grades (e.g., teachers' feedback) can also widely be ruled out because grades and IM were only slightly associated at t_1.

However, the present findings on effects of grades contradict those for secondary school students (e.g., Marsh *et al.*, 2005; Poorthuis *et al.*, 2015). Maybe performance feedback in the form of grades becomes more important for students' motivation after elementary school because self-worth and self-perceptions are shaped more by normative factors (social comparisons) in adolescence than in childhood (see Harter, 1998; Jacobs, Bleeker, & Constantino, 2003) and because of the changes in the school environment that occur in secondary school (see Wigfield & Eccles, 1994). For example, public evaluations of achievement increase in secondary school (Harter, Whitesell, & Kowalski, 1992), which facilitates social achievement comparisons and competition between students (see Wigfield & Eccles, 1994), which in turn should increase the salience of grades and their effects on motivation. In line with this, Poorthuis *et al.* (2015) found that students who expected their classmates to perform very well (high-performance norms) experienced relatively high levels of negative affect when receiving poor grades. Thus, students' reactions to grades seem to depend at least in part on the perceived importance of their own performance as the perceived importance of performance should be positively associated with higher performance pressure and increased social comparisons. This is in line with the two small effects found in our study. Grades slightly impacted change in IM in math after grades were introduced at the end of Grade 2 and after the recommendations for the school transition in Grade 4 were made. When grades become extremely salient, they seem to have at least a small impact on students' math-related IM even in elementary school. This differentiated picture of the relation between the constructs would not have come to light whether the constructs had been assessed only once a year. Thus, we recommend that researchers investigate change in IM several times during the school year to better understand what impacts this change.

Also contrary to studies on older students (Marsh *et al.*, 2005; Steinmayr & Spinath, 2009; Wang, 2012) but in line with results from the German domain in elementary school (Weidinger *et al.*, 2015), we did not find substantial effects of prior IM on later grades in math. In spite of its overall decline, IM remained relatively high at all time points, whereas math-related IM is typically not that positive in secondary school (e.g., Spinath & Steinmayr, 2012). Perhaps a ceiling effect took place, meaning that IM had no further effects on grades.

On the basis of our results, we can legitimately conclude that grades seem to play only a minor role in the decline in students' math-related IM in elementary school. When we found an effect, it was only in phases when grades were extremely salient (i.e., after being introduced and after the recommendation for subsequent schools had been made). Therefore, the decline in IM might be attributed to other possibly causative factors. Previous studies showed that parents' perceptions of their child's ability (Fredricks & Eccles, 2002) and parents' motivational practices (i.e., task-intrinsic practices; Gottfried, Marcoulides, Gottfried, & Oliver, 2009) contributed to within-person changes in students' IM. Further potential moderators are factors that proved to be important for IM development such as students' achievement goals (Haimovitz, Wormington, & Corpus, 2011), the teacher–student relationship (Opdenakker, Maulana, & den Brok, 2012), classroom values, teacher enthusiasm (Frenzel, Goetz, Pekrun, & Watt, 2010), and verbal feedback during class (Jurik, Gröschner, & Seidel, 2014; Rakoczy, Klieme, Bürgermeister, & Harks, 2008).

Strengths, limitations, and future directions

The fact that we considered the longitudinal interplay between IM and grades in an educational phase when grades are highly relevant to students (i.e., right after grades are implemented and when they are used for secondary school recommendations) is particularly suitable for drawing conclusions about the role of grades in change in IM. However, our findings must be interpreted in the context of several limitations. First, we investigated only students from Germany. In most German federal states, the transition to secondary school takes place after Grade 4. Thus, it is important to evaluate the generalizability of the findings in other countries where children leave elementary school later. Second, we analysed data from the same project as Weidinger *et al.* (2015). Results should be replicated with other teachers and students. Third, we investigated only one aspect of adaptive motivational orientation in school, namely the enjoyment and pleasure one gains from doing math. Thus, our results are limited to this specific aspect of motivation. Finally, we investigated only grades as one source of achievement-related feedback. However, students are exposed to different kinds of performance feedback in everyday school life. Besides verbal feedback from teachers and peers during class, students receive implicit information on how good their school performance is (e.g., who is offered help or who is asked to answer a difficult question; see Stipek & Mac Iver, 1989). In line with these ideas, even young children seem to be quite good at evaluating their actual performance without receiving explicit feedback (see Viljaranta, Lerkkanen, Poikkeus, Aunola, & Nurmi, 2009). In this context, it would be interesting to see how teachers' judgements of students' ability impact a student's motivational development (see Urhahne, 2015). This is particularly important with regard to intervention programs as teachers could be taught how several behaviours and responses might affect their students' IM.

Conclusion and practical implications

The fact that results found in the German domain (Weidinger *et al.*, 2015) could be replicated in math in the present study is a further hint that IM is not so much a result of competence feedback from grades than of other factors. Furthermore, our results showed that IM is very high in elementary school and its decline is not as strong as one might expect even in phases that put much pressure on students such as when transition recommendations are made. Finally, it shows that the relation between change in IM and school grades should be investigated within time spans shorter than 1 year because effects might otherwise be overlooked or misinterpreted.

Our findings have important practical implications for educators and educational and developmental psychologists working in school settings. The finding that grades are only weakly related to the negative development of students' IM is counterintuitive and contradicts the widespread notion that grades, especially poor grades, are detrimental to the pleasure one gains from doing school-related task. It is important to know this because teachers might feel as if they are doing harm when grading their students. The results indicate that students' IM for school-based learning does not necessarily suffer in the face of suboptimal grades.

Acknowledgements

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

References

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

Aunola, K., Leskinen, E., & Nurmi, J.-E. (2006). Developmental dynamics between mathematical performance, task motivation, and teachers' goals during the transition to primary school. *British Journal of Educational Psychology*, *76*, 21–40. doi:10.1348/000709905X51608

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

Corpus, J. H., McClintic-Gilbert, M. S., & Hayenga, A. O. (2009). Within-year changes in children's intrinsic and extrinsic motivational orientations: Contextual predictors and academic outcomes. *Contemporary Educational Psychology*, *34*, 154–166. doi:10.1016/j.cedpsych.2009.01.001

Corpus, J. H., & Wormington, S. V. (2014). Profiles of intrinsic and extrinsic motivations in elementary school: A longitudinal analysis. *The Journal of Experimental Education*, *82*, 480–501. doi:10.1080/00220973.2013.876225

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*, 627–668. doi:10.1037/0033-2909.125.6.627

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum.

Deutsch, M. (1979). Education and distributive justice: Some reflections on grading systems. *American Psychologist*, *34*, 391–401. doi:10.1037/0003-066X.34.5.391

Dickhäuser, O., & Plenter, I. (2005). "Letztes Halbjahr stand ich zwei". Zur Akkuratheit selbst berichteter Noten ["Last term I had a B". On the accuracy of self-reported school grades]. *Zeitschrift für Pädagogische Psychologie*, *19*, 219–224. doi:10.1024/1010-0652.19.4.219

Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy*, *35*, 333–363. doi:10.1016/S0005-7894(04)80042-X

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109–132. doi:10.1146/annurev.psych.53.100901.135153

Eccles, J. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*, 830–847. doi:10.2307/1131221

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

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, *20*, 507–537. doi:10.1111/j.1532-7795.2010.00645.x

Garon-Carrier, G., Boivin, M., Guay, F., Kovas, Y., Dionne, G., Lemelin, J.-P., . . . Tremblay, R. E. (2016). Intrinsic motivation and achievement in mathematics in elementary school: A longitudinal investigation of their association. *Child Development*, *87*, 165–175. doi: 10.1111/cdev.12458

Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology*, *82*, 525–538. doi:10.1037/0022-0663.82.3.525

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

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., & Oliver, P. H. (2009). A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation. *Journal of Educational Psychology*, *101*, 729–739. doi:10.1037/a0015084

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., Ratelle, C. F., Roy, A., & Litalien, D. (2010). Academic self-concept, autonomous academic motivation, and academic achievement: Mediating and additive effects. *Learning and Individual Differences*, *20*, 644–653. doi:10.1016/j.lindif.2010.08.001

Guskey, T. R. (2000). Grading policies that work against standards ..and how to fix them. National Association of Secondary School Principals. *NASSP Bulletin*, *84*, 20–27.

Haimovitz, K., Wormington, S. V., & Corpus, J. H. (2011). Dangerous mindsets: How beliefs about intelligence predict motivational change. *Learning and Individual Differences*, *21*, 747–752. doi:10.1016/j.lindif.2011.09.002

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology*, *17*, 300–312. doi:10.1037/0012-1649.17.3.300

Harter, S. (1998). The development of self-representations. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development*, 5th ed. (pp. 553–617). New York, NY: Wiley.

Harter, S., Whitesell, N., & Kowalski, P. (1992). Individual differences in the effects of educational transitions on young adolescents' perceptions of competence and motivational orientation. *American Educational Research Journal*, *29*, 777–807. doi:10.2307/1163407

Jacobs, J. E., Bleeker, M. M., & Constantino, M. J. (2003). The self-system during childhood and adolescence: Development, influences, and implications. *Journal of Psychotherapy Integration*, *13*, 33–65. doi:10.1037/1053-0479.13.1.33

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

Jurik, V., Gröschner, A., & Seidel, T. (2014). Predicting students' cognitive learning activity and intrinsic learning motivation: How powerful are teacher statements, student profiles, and gender? *Learning and Individual Differences*, *32*, 132–139. doi:10.1016/j.lindif.2014.01.005

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

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

McElvany, N., Kortenbruck, M., & Becker, M. (2008). Lesekompetenz und Lesemotivation Entwicklung und Mediation des Zusammenhangs durch Leseverhalten [Reading literacy and reading motivation: Their development and the mediation of the relationship by reading behavior]. *Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology*, *22*, 207–219. doi:10.1024/1010-0652.22.34.207

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

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: Muthén and Muthén.

Niepel, C., Brunner, M., & Preckel, F. (2014). 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

Opdenakker, M.-C., Maulana, R., & den Brok, P. (2012). Teacher—Student interpersonal relationships and academic motivation within one school year: Developmental changes and linkage. *School Effectiveness and School Improvement*, *23*, 95–119. doi:10.1080/09243453.2011.619198

Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., & Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology*, *84*, 152–175. doi:10.1111/bjep.12028

Poorthuis, A. M. G., Juvonen, J., Thomaes, S., Denissen, J. J. A., de Castro, B. O., & van Aken, M. A. G. (2015). Do grades shape students' school engagement? The psychological consequences of report card grades at the beginning of secondary school. *Journal of Educational Psychology*, *107*, 842–854. doi:10.1037/edu0000002

Rakoczy, K., Klieme, E., Bürgermeister, A., & Harks, B. (2008). The interplay between student evaluation and instruction. *Zeitschrift für Psychologie/Journal of Psychology*, *216*, 111–124. doi:10.1027/0044-3409.216.2.111

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68–78. doi:10.1037/0003-066X.55.1.68

Schneider, R., & Sparfeldt, J. R. (2016). Zur (Un-)Genauigkeit selbstberichteter Zensuren bei Grundschulkindern [The accuracy of self-reported grades in elementary school]. *Psychologie in Erziehung und Unterricht*, *63*, 48–59. doi:10.2378/peu2016.art05d

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

Spinath, B., & Spinath, F. M. (2005). 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

Spinath, B., & Steinmayr, R. (2012). The role of competence beliefs and goal orientations for change in intrinsic motivation. *Journal of Educational Psychology*, *104*, 1135–1148. doi:10.1037/a0028115

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*, 80–90. doi:10.1016/j.lindif.2008.05.004

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

Urhahne, D. (2015). Teacher behavior as a mediator of the relationship between teacher judgment and students' motivation and emotion. *Teaching and Teacher Education*, *45*, 73–82. doi:10.1016/j.tate.2014.09.006

Viljaranta, J., Lerkkanen, M.-K., Poikkeus, A.-M., Aunola, K., & Nurmi, J.-E. (2009). Cross-lagged relations between task motivation and performance in arithmetic and literacy in kindergarten. *Learning and Instruction*, *19*, 335–344. doi:10.1016/j.learninstruc.2008.06.011

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

Wang, M.-T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interest. *Child Development*, *48*, 1643–1657. doi:10.1037/a0027247

Weidinger, A. F., Spinath, B., & Steinmayr, R. (2015). Zur Bedeutung von Grundschulnoten für die Veränderung von Intrinsischer Motivation und Fähigkeitsselbstkonzept in Deutsch [The role of grades in changes in intrinsic motivation and ability self-concept in German]. *Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology*, *29*, 193–204. doi:10.1024/1010-0652/a000160

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.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*, 297–333. doi:10.1037/h0040934

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, *12*, 265–310. doi:10.1016/0273-2297(92)90011-P

Wigfield, A., & Eccles, J. S. (1994). Children's competence beliefs, achievement values, and general self-esteem: Change across elementary and middle school. *Journal of Early Adolescence*, *14*, 107–138. doi:10.1177/027243169401400203

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., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2006). Development of achievement motivation. In W. Damon & R. M. Lerner (Series Eds.) & N. Eisenberg (Volume Ed.), *Handbook of child psychology*, *Vol. III*, 6th ed. (pp. 933–1002). 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

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

Received 27 June 2016; revised version received 9 November 2016