The Longitudinal Interplay of Students' Academic Self-Concepts and Achievements Within and Across Domains...

Article *in* Journal of Educational Psychology · April 2014

DOI: 10.1037/a0036307

3 authors:

Christoph Niepel
University of Luxembourg

Martin Brunner
Universität Potsdam

Franzis Preckel
Universität Trier

Journal of Educational Psychology

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

Christoph Niepel, Martin Brunner, and Franzis Preckel

Online First Publication, April 7, 2014. http://dx.doi.org/10.1037/a0036307

Journal of Educational Psychology

0022-0663/14/$12.00 DOI: 10.1037/a0036307

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

Christoph Niepel
University of Trier

Martin Brunner
Free University of Berlin and Berlin-Brandenburg Institute for School Quality, Berlin, Germany

Franzis Preckel
University of Trier

Students' cognitive and motivational profiles have a large impact on their academic careers. The development of such profiles can partly be explained by the reciprocal internal/external frame of reference model (RI/E model). The RI/E model predicts positive *and* negative longitudinal effects between academic self-concepts and achievements within and across 2 academic domains (i.e., the mathematics and verbal domains). In the present study, we replicated the RI/E model in 2 samples and extended it by simultaneously investigating the longitudinal associations of academic self-concepts and achievements in 3 academic domains (i.e., mathematics, German as a native language, English as a foreign language). We examined 2 domains across 4 measurement occasions in 2 independent student samples who were in Grades 5–8 (Study 1: $N = 1{,}529$) or Grades 5–7 (Study 2: $N = 639$). In a 3rd study, we examined a subsample of the 2nd sample ($N = 465$) in 3 domains. Results demonstrated support for the RI/E model for 2 as well as 3 academic domains. We found positive reciprocal effects of academic self-concepts and achievements within a domain, positive reciprocal effects between achievements across domains, negative effects of achievements on subsequent cross-domain self-concepts, negative effects of academic self-concepts on subsequent cross-domain achievements, and some support for negative effects of academic self-concepts on subsequent cross-domain self-concepts. Furthermore, we found that RI/E model effects were of comparable size across time. To conclude, the RI/E model provides significant insights into the development of distinct motivational and cognitive profiles and, thus, of becoming either a math or verbal person. Implications for research and educational practice are discussed.

Keywords: reciprocal internal/external frame of reference model, achievement, mathematics self-concept, verbal self-concept, longitudinal data

Thinking positively about one's academic abilities is essential for promoting desirable educational outcomes (see, e.g., Marsh, 2006; Marsh & Martin, 2011). A plethora of studies have suggested that positive academic self-concepts—as mental representations of one's academic abilities—have beneficial effects on a broad range of educational outcomes, such as students' achievement, academic effort, or academic emotions (see, e.g., Goetz, Preckel, Zeidner, & Schleyer, 2008; Ireson & Hallam, 2009; Marsh, 1991; Marsh et al., 2008; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Valentine, DuBois, & Cooper, 2004).

Most previous research on academic self-concepts has concentrated on the mutual relations between academic self-concepts and academic achievements (e.g., Huang, 2011). More specifically, an ever-growing body of (a) longitudinal research on reciprocal and mutually reinforcing effects between academic self-concepts and academic achievements within a single academic domain (i.e., *reciprocal effects model* [REM]) has been accompanied by a growing body of (b) cross-sectional research on the *internal/external frame of reference model* (I/E model) that predicts (among others) negative contrast effects of achievement in one domain (e.g., mathematics) on academic self-concept in another domain (e.g., German as a verbal subject). Importantly, these two models were typically studied in isolation until Möller, Retelsdorf, Köller, and Marsh (2011) presented their pioneering work. In their study, Möller et al. integrated the two models in the *reciprocal internal/external frame of reference model* (RI/E model) to *simul-*

Christoph Niepel, Department of Psychology, University of Trier; Martin Brunner, Department of Education and Psychology, Free University of Berlin, and Berlin-Brandenburg Institute for School Quality, Berlin, Germany; Franzis Preckel, Department of Psychology, University of Trier.

The data used in this study came from two longitudinal projects. The Study 1 data stem from Franzis Preckel's longitudinal study, which focused on motivational and self-concept development in secondary school students. This longitudinal study is supported by the Ministry of Education, Science, Adolescence, and Culture of Rhineland-Palatinate. The Study 2 and Study 3 data stem from the PULSS Project conducted by Wolfgang Schneider, Eva Stumpf, Franzis Preckel, and Albert Ziegler, a longitudinal study on school achievement and academic development at the beginning of secondary education. The project is funded by the Bavarian Ministry for Education and Culture; by the Ministry for Culture, Youth, and Sports of the Federal State of Baden-Wuerttemberg; and by the Karg-Foundation, Frankfurt/Main, Germany.

Correspondence concerning this article should be addressed to Christoph Niepel, University of Trier, Department of Psychology, 54286 Trier, Germany. E-mail: niepel@uni-trier.de

taneously examine (a) reciprocal effects and (b) contrast effects between academic self-concepts and achievements in different academic domains in the long run. Of note, by considering the development of academic self-concept *and* achievement within *and* across different academic domains, the RI/E model provides insights into the joint development of students' cognitive and motivational profiles. Thus, the RI/E model is important for promoting a better understanding of the developmental tasks and academic trajectories that depend on such profiles. This involves forming one's self, selecting specialized courses, or deciding on a career choice.

Given its immanent importance for educational theory, research, and practice, further research on the RI/E model is necessary. To date, however, there have been no replications of the RI/E model. Furthermore, some predictions derived from the RI/E model need to be subjected to empirical scrutiny. Notably, the potential negative long-term effect of academic self-concept on cross-domain achievement needs further clarification (see next paragraph). The present study therefore aimed to further empirically examine the RI/E model by analyzing the mutual longitudinal relations between achievements and self-concepts in mathematics and German (as a verbal subject). Importantly, to carefully explore the generalizability of the RI/E model, we analyzed two large independent samples of data that both included four waves of data collection: Study 1 included Grades 5 to 8, and Study 2 included Grades 5 to 7. Capitalizing on this broad empirical data base, we were also able to extend the RI/E model by simultaneously investigating academic self-concepts and achievements in mathematics and a native (i.e., German) as well as a foreign (i.e., English) language. This significant extension of the RI/E model across *three* different academic domains has not been conducted before. Additionally, we were interested in the question of whether the strengths of the relations between academic self-concepts and achievements differ in different phases of students' school careers (e.g., new learning groups in the first months of secondary school vs. stable learning groups in higher grades). If yes, this would indicate phase-dependent developmental differences in students' formation of cognitive and motivational profiles. To our knowledge, this question had not been investigated yet.

The Reciprocal Internal/External Frame of Reference Model

Recently, Möller and colleagues (2011) proposed a new approach by merging two established and widely studied academic self-concept models, (a) the REM and (b) the I/E model into the RI/E model. By doing so, they extended the REM to more than one academic domain and the I/E model longitudinally (Möller et al., 2011).

The REM depicts mutual relations between academic self-concepts and academic achievements, suggesting that higher achievements foster self-concepts (when previous self-concepts are controlled; cf. *skill development*; Calsyn & Kenny, 1977) and that higher self-concepts enhance achievements (when previous attainment is controlled; cf. *self-enhancement*; Calsyn & Kenny, 1977; for an overview, see Marsh & Craven, 2006; Marsh & Martin, 2011; see also Guay, Marsh, & Boivin, 2003). Thus, the REM postulates three positive long-lasting effects that are—amongst others—illustrated in Figure 1 by presenting the entire RI/E model framework in two academic domains. These effects consist of: (a) *skill development effects* (positive effects of achievement on subsequent academic self-concept within one domain; indicated by α in Figure 1), (b) *self-enhancement effects* (positive effects of academic self-concept on subsequent achievement within one domain; indicated by β in Figure 1), and (c) *autoregressive effects* (indicated by ε in Figure 1).

The I/E model (e.g., Möller, Pohlmann, Köller, & Marsh, 2009) also postulates positive skill development effects but additionally postulates negative contrast effects of achievement on subsequent academic self-concepts across domains (*internal frame of reference effects*; indicated by γ in Figure 1). Thus, high achievement in one domain (e.g., mathematics achievement) is assumed to exert a negative influence on academic self-concept in a contrasting domain (e.g., verbal self-concept). This internal frame of reference effect thus describes dimensional comparison effects (Möller & Marsh, 2013; see also Marsh, 2006; Möller & Köller, 2001a; Pohlmann & Möller, 2009). Of note, most research on the I/E model has drawn on cross-sectional data (e.g., Marsh, 2006; Möller et al., 2009) and has not tested the predictions longitudinally (see Brunner et al., 2010, for an exception).

The predictions of the RI/E model are largely based on those of both single models. Yet the RI/E model predictions should also hold longitudinally. Furthermore, they should even hold after controlling for baseline measures within *and* across the verbal and mathematics domains (i.e., achievements and academic self-concepts)—going beyond conventional REM and I/E model research (Möller et al., 2011). Figure 1 presents all assumptions of the RI/E model (Möller et al., 2011). As is true for the REM and I/E models, the RI/E model predicts (a) skill development effects (indicated by α), (b) self-enhancement effects (indicated by β), (c) autoregressive effects (indicated by ε), and (d) internal frame of reference effects (indicated by γ). Furthermore, by extending the rationale of the I/E model longitudinally, the RI/E model predicts (e) positive cross-domain effects of achievements on subsequent achievements in the contrasting domain (i.e., positive *reciprocal cross-domain effects* between achievements; indicated by ζ) and (f) weak or null effects of academic self-concepts on subsequent noncorresponding self-concepts across domains (i.e., the prior verbal self-concept only marginally or does not affect the subsequent mathematics self-concept and vice versa; indicated by η). However, a new prediction emerges by combining the REM with the I/E model: (g) Prior academic self-concept should *negatively* affect subsequent achievement in the contrasting domain (e.g., mathematics self-concept should negatively affect subsequent German achievement) when prior achievements are controlled for (indicated by δ in Figure 1). Such negative *cross-domain contrast effects of academic self-concepts* have been neglected in academic self-concept research to date. Prior academic self-concept research has focused almost exclusively on the positive effects of academic self-concept (e.g., Marsh & Martin, 2011), considering self-concept to be "a 'hot' variable that makes good things happen" (Marsh & Craven, 2006, p.158). As a complement, the RI/E model sheds light on a potential dark side of academic self-concepts.

Möller and colleagues (2011) identified five longitudinal studies that tested effects of academic self-concepts on contrasting domains (Köller, Klemmert, Möller, & Baumert, 1999; Marsh & Köller, 2004; Marsh & Yeung, 1997, 1998; Möller & Köller, 2001b). These studies delivered somewhat inconsistent results for

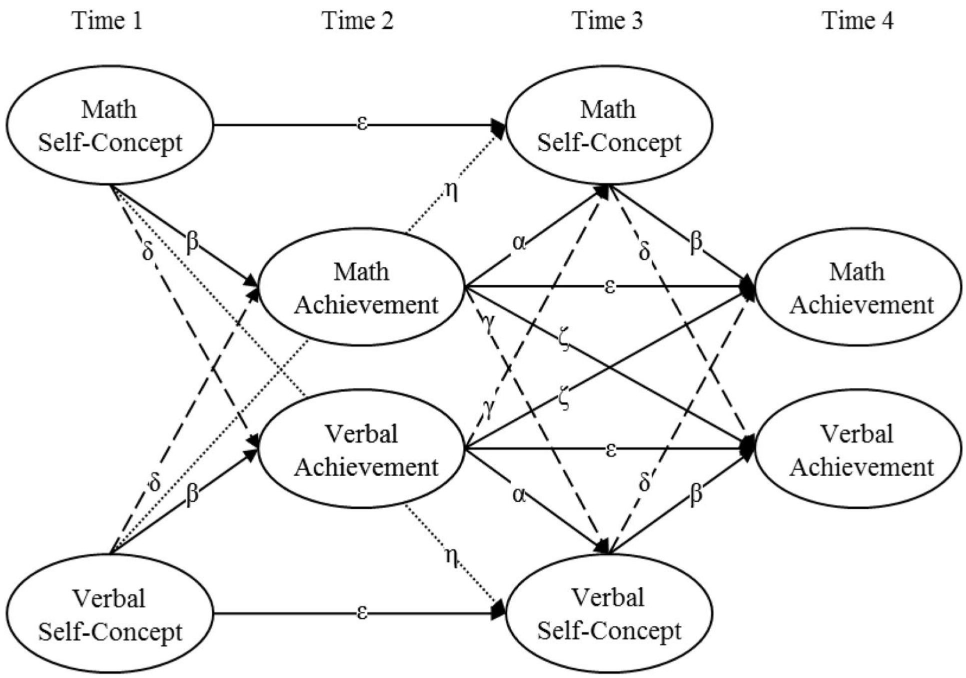

Figure 1. Seven predictions derived from the reciprocal internal/external frame of reference model: (α) skill development effects, (β) self-enhancement effects, (γ) internal frame of reference effects, (δ) cross-domain contrast effects of academic self-concepts, (ε) autoregressive effects, (ζ) reciprocal cross-domain effects between achievements, and (η) null effects of the domain-specific self-concept on the cross-domain self-concept. Predicted positive effects are depicted by continuous lines, negative effects by dashed lines, and null effects by dotted lines.

negative cross-domain contrast effects of academic self-concepts by suggesting either null relations or very small negative effects of academic self-concept in one domain on achievement in another domain.

In their study, Möller et al. (2011) examined data from a sample of 1,508 German secondary school students to analyze the RI/E model framework using three waves of data collection from Grades 5 to 8. Overall, findings supported the seven core predictions of the RI/E model (as mentioned above). Yet, compared to other postulated effects, support for negative effects of academic self-concepts on subsequent achievements across domains (i.e., cross-domain contrast effects of academic self-concepts) was considerably weaker: Möller et al. found only indirect support for negative cross-domain contrast effects of academic self-concepts. Although none of these effects reached statistical significance (at $p < .05$), restricting the corresponding paths to zero resulted in significant decreases in model fit (i.e., for chi-square difference testing via model comparison). In sum, these results point to the need for more research on this effect in particular and on the RI/E model in general. To date, the model has not been replicated.

Extending the RI/E Model to Mathematics, a Native Language, and a Foreign Language

Most I/E model research has contrasted mathematics with a verbal domain (e.g., German as a verbal subject), supporting negative cross-domain contrast effects between achievements and academic self-concepts (i.e., internal frame of reference effects). However, I/E model research juxtaposing two domains that are typically perceived as closely related—such as two verbal domains (i.e., German and English) or two numerical or science domains (i.e., mathematics and physics)—has delivered rather mixed results. For example, Möller, Streblow, Pohlmann, and Köller (2006) extended the I/E model to two verbal and two science domains (i.e., mathematics, physics, German, and English). They identified positive cross-sectional correlations between English achievement and German self-concept, between mathematics achievement and physics self-concept, and between physics achievement and mathematics self-concept, respectively. In line with this, Xu et al. (2013) recently found no internal frame of reference effects among the native (i.e., Chinese) and foreign (i.e., English) languages of students, but they found negative contrast effects of mathematics achievement on Chinese and English self-concepts. In addition, Chinese and English achievements were negatively related to mathematics self-concept. By contrast, Chiu (2012) found negative internal frame of reference effects between mathematics and science. Of note, all of these findings are based on cross-sectional data. For longitudinal data, Brunner et al. (2010) reported negative longitudinal internal frame of reference effects from Grades 6 to 8 between French and German as verbal subjects in a Luxembourgish sample (of note, both languages are vital in Luxembourg; see Brunner et al., 2010, for a discussion). Taken together, more longitudinal research on internal frame of reference effects between closely related domains is needed.

In the present study, we investigated mathematics as well as two verbal domains, namely, German (i.e., students' native language) and English (i.e., students' first foreign language in secondary school) within an extended RI/E model. The extension of the RI/E model across these domains leads—as stated above—automatically to a *longitudinal* examination of internal frame of reference effects between mathematics and German, mathematics and English, and German and English.

Stability of Effects Across Time

Whether the relative impact of the postulated effects of the RI/E model varies or shows stability across time is an open question. One could assume that the impact of academic self-concepts on corresponding achievements (i.e., self-enhancement effects) increases during transition periods such as when students enter secondary school and, therewith, when they enter a new learning environment in which they have to cope with uncertainty regarding their academic standing (Preckel, Götz, & Frenzel, 2010; Wagner, 2001). On the other hand, skill development effects may be more powerful in learning environments that are perceived as rather stable (Helmke, 1992). By examining students from the beginning of secondary school (i.e., Grade 5) to Grade 8 (Study 1) and from the beginning of Grade 5 to Grade 7 (Study 2), we were able to test potential developmental differences in the strength of self-enhancement and skill development effects—and the other effects postulated by the RI/E model. Variations in the magnitudes of the observed effects over time could indicate that these developmental processes depend on phases in secondary school during which the students' learning environment becomes increasingly stable across time. In the methodological literature, the stability of effects is called *stationarity* (Kenny, 1975; T. D. Little, Preacher, Selig, & Card, 2007). In the present study, we examined the stationarity of RI/E model effects in an important transition phase within students' educational careers. To our knowledge, this has never been done before.

Research Aims

The present study pursued two vital research aims. Bonett (2012) stated, "replication evidence is the gold standard by which scientific claims are evaluated" (p. 410), yet only about 1% of psychology publications are actually replications (Makel, Plucker, & Hegarty, 2012; for further arguments on the importance of replications, see, e.g., Pashler & Harris, 2012). Accordingly, our first important research aim was to replicate the RI/E model because the RI/E model has not been replicated to date. Second, we aimed to significantly extend the RI/E model in two directions: (a) We investigated not only the domains of mathematics and native language but also a domain that is closely related to the latter—that is, English as a foreign language (replication-extension study; Bonett, 2012). (b)This study is the first to analyze the temporal stability of RI/E model predictions across time. To these ends, we conducted three studies that capitalized on large data sets from two independent samples of German secondary school students.

Replication

Our study design contained some features that are important to have in a replication study, thus allowing us to extend the generalizability of prior RI/E model results to a different study population and across varying time lags. We were also able to broaden the empirical body of RI/E model research by considering more points of measurement (see, e.g., Bonett, 2012; Shadish, Cook, & Campbell, 2001). First, to complement Möller and colleagues' (2011) sample, which comprised students from different types of secondary schools, Studies 1 to 3 in the present study focused only on students attending the top track of the highly selective German secondary school system (i.e., the German *Gymnasium*). In comparison to the other scholastic tracks, the Gymnasium differs with respect to school quality and learning environment such that instruction in the Gymnasium is more cognitively demanding and activating (Becker, Lüdtke, Trautwein, Köller, & Baumert, 2012). We examined the generalizability of the RI/E model by focusing on a selected subgroup of secondary school students to learn whether the effects found for a heterogeneous sample of the German student body would also hold for students in the Gymnasium, who comprise a more homogeneous subgroup (differences in results would indicate a moderation effect due to specific demographic features of this subgroup). Second, Study 1 on the one hand and Studies 2 and 3 on the other hand varied in their time lags with regard to when performance feedback was given and when academic self-concepts were assessed (differences in results would indicate an effect of time-lag length). Third, Möller et al. examined a panel study with three waves of measurement, whereas all of the three studies in the present article contained *four* waves of measurement.

Extension: Three Domains and Tests of Stationarity

As outlined above, prior examinations of relations between academic self-concepts and achievements in two verbal domains (i.e., research on the I/E model) have revealed rather mixed findings. In addition, most prior findings were derived from cross-sectional data. We therefore investigated the RI/E model predictions longitudinally and across the domains of (a) mathematics and German (Studies 1 and 2) and (b) mathematics, German, and English (Study 3). Furthermore, the present study explored the longitudinal stability (i.e., stationarity) of the effects postulated by the RI/E model (i.e., in Studies 1, 2, and 3).

Note that we are using a terminology that is typically applied in self-concept research in which causal effects are postulated to describe the mutual relations between achievements and academic self-concepts. In the following sections of the present article—for the sake of conceptual clarity—the reader should note that we sometimes refer to these terms (as introduced above; e.g., self-enhancement effects or skill development effects) to label the resulting longitudinal associations among variables, but we do so without directly implying causality.

Study 1

Method

Sample and procedure. Sample 1 was obtained from four waves of an ongoing longitudinal study focusing on self-concept development in secondary school. Data were obtained from a sample of 1,529 students (46.57% female) in 43 classes at five different top-track schools (i.e., the Gymnasium) in two federal states of Germany (i.e., Rhineland-Palatinate and Bavaria). The students in the sample were examined across a 40-month period at the beginning of Grade 5 (on average, 10 weeks after the students' first day in secondary school; Time 1) and shortly after the midterm reports in Grade 5 (Time 2), Grade 6 (Time 3), and Grade 8 (Time 4), respectively. Students stayed with the same peers throughout this time period, and they were all enrolled in the same type of mathematics and German classes. Students' mean age was 10.69 years ($SD = 0.44$) at Time 1, 11.02 years ($SD = 0.44$) at Time 2, 12.04 years ($SD = 0.44$) at Time 3, and 14.04 years ($SD = 0.46$) at Time 4. A self-report questionnaire was used to collect data on students' achievements (i.e., at Times 2, 3, and 4) and to assess academic self-concepts (i.e., at all four times of measurement) in mathematics and German. The questionnaire was administered during regular class sessions, and students' participation was voluntary. Written parental consent was obtained for all participating children.

Variables and measures.

Academic self-concepts. Academic self-concepts in mathematics and German were assessed by six items based on a German short version of the Self-Description Questionnaire (SDQ; Kunter et al., 2002). The SDQ, originally developed by Marsh (1990), is considered to be one of the best self-concept instruments available (e.g., Byrne, 2002) as research has amply confirmed its factor structure, reliability, and criterion-related validity (e.g., Byrne, 1996; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2006). We adapted the approach applied by Trautwein, Lüdtke, Marsh, Köller, and Baumert (2006) and used three of the best items (as determined by item-total correlations) in the total scale per self-concept (i.e., the mathematics and German self-concepts). Specifically, the items selected for German/mathematics self-concepts were "I learn German/mathematics quickly," "I have always done well in German/mathematics," and "German/mathematics is one of my best subjects." Students responded to these items on a 5-point rating scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) with higher scores indicating a higher academic self-concept.

Academic achievements. Students' achievements were assessed by self-reported grades in mathematics and German as obtained from the midterm reports in Grade 5 (i.e., students' first report in secondary school), Grade 6, and Grade 8, respectively. These self-reported grades represent cumulative attainment in mathematics and German from the preceding school term. Students reported their grades shortly after they received their midterm reports at Times 2, 3, and 4, respectively. Although meta-analytic research suggests that self-reported grades can be positively biased to a certain extent (Kuncel, Credé, & Thomas, 2005), recent research from German-speaking countries has demonstrated that self-reported school grades are not subject to systematic reporting biases (Dickhäuser & Plenter, 2005). Hence, these self-reported grades can be considered to represent reliable and valid indicators of students' mathematics and German achievements, respectively. In Germany, a 6-point grading system is used. For ease of interpretation in the present study, we scored grades such that higher numbers indicate better mathematics and German achievement (i.e., 1 = *unsatisfactory* to 6 = *excellent*).

Data analysis. We examined the RI/E model by means of structural equation modeling (SEM) using a four-wave latent-variable cross-lagged panel design (T. D. Little et al., 2007). Herein, academic self-concepts were specified as latent variables with three indicators each, and academic achievements (i.e., midterm records in mathematics and German) were modeled as manifest single-indicator variables. For the latent variables, we allowed the residuals of corresponding indicators to correlate across the points of measurement (i.e., correlated uniquenesses; T. D. Little et al., 2007). Figure 2 illustrates the resulting SEM. Hence, we investigated our research aims by considering three subsequent time spans: (a) the time span of approximately 4 months from the beginning of secondary school in the fifth grade (i.e., Time 1) to the beginning of the second semester of the fifth grade (i.e., Time 2), (b) the time span of 1 year from the beginning of the second semester of the fifth grade (i.e., Time 2) to the beginning of the second semester of the sixth grade (i.e., Time 3), and (c) the time span of 2 years from the beginning of the second semester of the sixth grade (i.e., Time 3) to the beginning of the second semester of the eighth grade (i.e., Time 4). Note that these observed time intervals were not of the same length. Because the magnitudes of the coefficients' absolute values will shift as the lags become shorter or longer, the observed magnitudes of the path coefficients were not directly comparable across these three time spans (Oud & Delsing, 2010). However, the recorded midterm grades were *always* provided to the students about 2 weeks before they answered the questionnaire on academic self-concepts. Thus, these lags were of comparable length. Figure 2 illustrates corresponding paths of comparable length by identical letters. Accordingly, the panel design applied in Study 1 allowed us to compare the magnitude *and,* therewith, test the stability (i.e., stationarity; Kenny, 1975) of (a) skill development effects and (b) internal frame of reference effects. To test the assumption of stationarity, we set all corresponding paths to equality across time (as presented in Figure 2; i.e., all *a* paths, *b* paths, *c* paths, and *d* paths, respectively). This approach enabled us to explore whether skill development effects and internal frame of reference effects changed across time. The tenability of the equality constraints could be tested via nested model comparisons (T. D. Little et al., 2007).

As imperatively required for studying longitudinal RI/E model effects and their stationarity, we previously tested whether the applied latent variables captured the same target constructs across time (i.e., we tested measurement invariance across time). We applied a stepwise strategy to test three levels of measurement invariance: (a) configural invariance (i.e., invariance in the pattern of zero and nonzero loadings), (b) metric invariance (i.e., additionally, invariance in the size of factor loadings), and (c) scalar invariance (i.e., additionally, invariance in the intercepts of the manifest variables). At least metric measurement invariance was required to interpret RI/E model effects in our SEM (for further information, see Widaman & Reise, 1997; see also Meredith, 1993; Meredith & Horn, 2001).

All data analyses were conducted with Mplus 7 (Muthén & Muthén, 1998–2012b). As some distributions of item scores differed slightly from multivariate normality, we used the Mplus option for maximum likelihood estimation to compute standard errors and fit statistics that were robust (MLR; see Muthén & Muthén, 1998–2012a) against these mild violations of normality (e.g., Kaplan, 2009). As is typical in longitudinal research, complete data on all measures across all waves of measurement were not available for all students. Specifically,

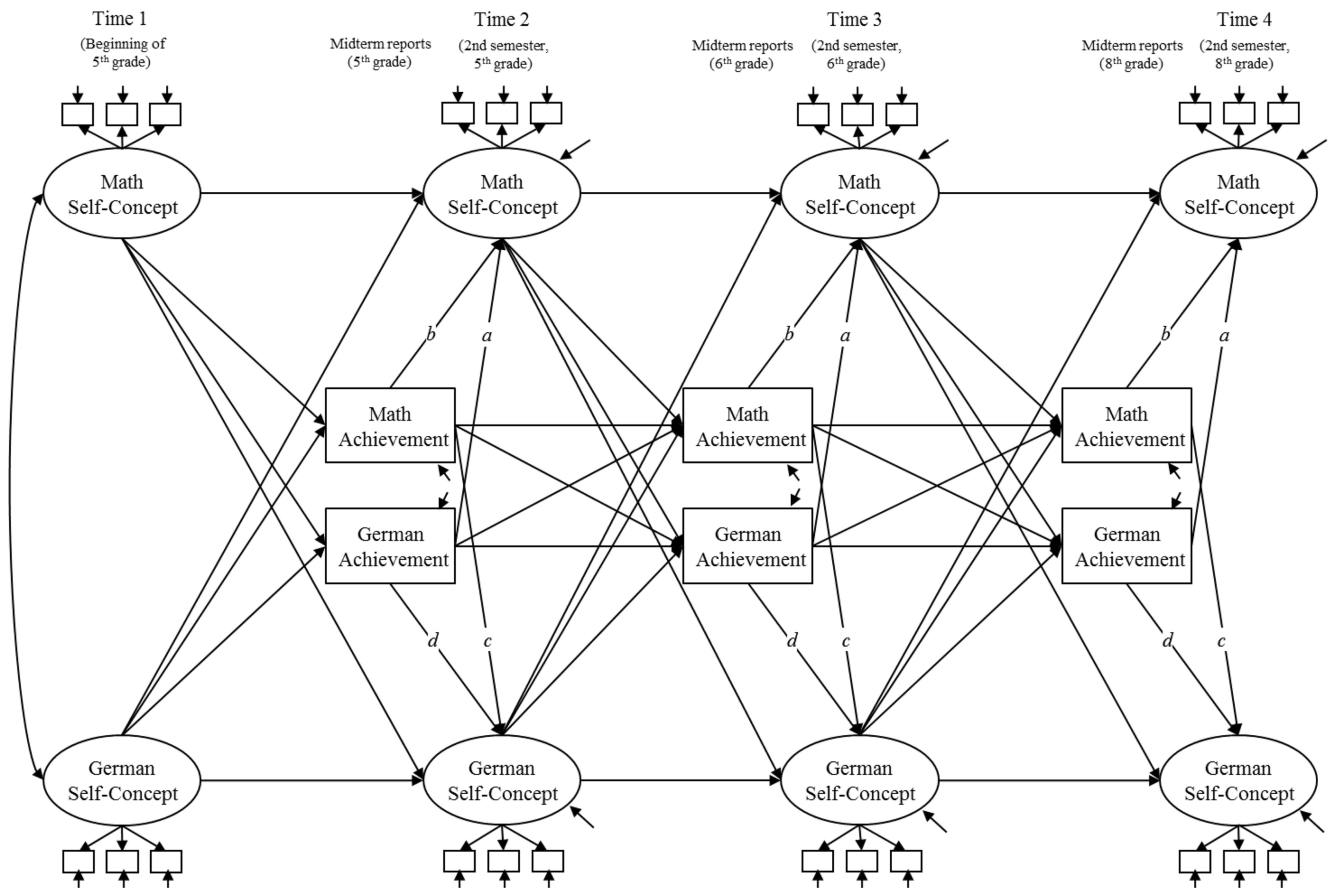

Figure 2. Study 1: Reciprocal internal/external frame of reference model in the German (as verbal subject) and mathematics domains. The *a, b, c,* and *d* paths represent lags of the same length. For the sake of clarity, correlational paths between residuals were omitted from the path diagram.

35% to 70% of the students provided data to estimate a certain variance or covariance at any time of measurement. Applying full-information maximum likelihood, the MLR estimator allowed us to deal efficiently with data that were missing in this way (R. J. A. Little, 1995; McArdle, 1994; Muthén & Muthén, 2006; see also McArdle & Grimm, 2010), thus yielding (a) higher statistical power to study the RI/E model (Collins, Schafer, & Kam, 2001) and (b) parameter estimates that had less bias than estimates gleaned from traditional pairwise or listwise deletion methods (e.g., Graham, 2009).[1]

Students were nested within classes in our data. As the number of clusters in our sample (i.e., 43 classes) was not sufficient to use the complex analysis feature in Mplus (i.e., a correction method to account for the nested data structure; for more information, see the *Mplus User's Guide,* Muthén & Muthén, 1998–2012a), we followed Cronbach's (1976) advice and standardized (M = 0.00, SD = 1.00) all variables within school classes to remove clustering effects (the same procedure was also chosen by Trautwein, Lüdtke, Köller, & Baumert, 2006).

To evaluate model fit, the following criteria were used to indicate adequate fit: (a) the normed version of the chi-square statistic (Bollen, 1989), (b) the root-mean-square error of approximation, (c) the comparative fit index, and (d) the standardized root-mean-square residual (Hu & Bentler, 1999). To compare two nested models (i.e., to test for measurement invariance and stationarity), we computed chi-square difference tests using the Satorra-Bentler-scaled chi-square (Satorra & Bentler, 2001).

Results

Preliminary analysis.

Descriptive statistics. Table 1 presents correlations, means, standard deviations, and internal consistencies for all measures at each wave of measurement. All measures revealed good internal consistencies (αs between .84 and .88). In line with prior I/E model research, Table 1 shows moderate positive correlations between the academic achievement variables (average correlation r = .39, range r = .29 to .50) and low correlations between academic

[1] To illustrate the robustness of our results, we performed the analyses for Study 1 a second time using the listwise deletion technique without any missing values. There were virtually no differences between the results of Study 1 as presented in this article (using full-information maximum likelihood; see the Results section) and Study 1 using the listwise deletion technique.

Table 1
Study 1: Observed Correlations (r), Means (M), Standard Deviations (SD), and Reliabilities (Cronbach's α) for Mathematics and German Self-Concept Scales (MSC/GSC) and Mathematics and German Achievements (Teacher-Assigned School Grades; MACH/GACH)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	*M*	*SD*	α
1. MSC Time 1	—														3.91	0.94	.85
n	1,068																
2. MSC Time 2	.67**	—													3.73	1.01	.85
n	928	1,068															
3. MSC Time 3	.51**	.61**	—												3.43	1.13	.87
n	842	850	1,038														
4. MSC Time 4	.42**	.50**	.64**	—											3.09	1.19	.88
n	556	559	567	844													
5. GSC Time 1	.12**	.08*	−.02	−.08	—										3.74	0.95	.87
n	1,062	925	838	555	1,062												
6. GSC Time 2	.07*	.07*	−.03	−.09*	.58**	—									3.55	0.98	.85
n	923	1,058	844	557	920	1,060											
7. GSC Time 3	.03	.00	.06*	−.02	.45**	.55**	—								3.40	1.01	.84
n	840	846	1,031	568	836	840	1,035										
8. GSC Time 4	.01	.03	.03	−.02	.35**	.40**	.50**	—							3.30	1.02	.84
n	550	554	563	833	549	551	564	836									
9. MACH 5	.31**	.54**	.46**	.39**	.02	.03	.08*	.11**	—						4.47	0.84	
n	917	1,048	839	555	913	1,041	836	550	1,056								
10. MACH 6	.24**	.38**	.59**	.50**	.01	.10**	.12**	.14**	.59**	—					4.25	0.95	
n	840	848	1,009	571	837	842	1,006	566	840	1,034							
11. MACH 8	.24**	.29**	.43**	.63**	−.04	.01	.07	.07	.45**	.59**	—				4.00	1.06	
n	556	559	565	820	555	556	566	812	555	570	838						
12. GACH 5	.02	.05	.09*	.06	.27**	.52**	.48**	.40**	.37**	.37**	.32**	—			4.31	0.79	
n	922	1,052	846	557	918	1,044	843	552	1,051	846	557	1,059					
13. GACH 6	.01	.07*	.14**	.12**	.21**	.36**	.56**	.38**	.38**	.50**	.36**	.63**	—		4.26	0.83	
n	839	848	1,009	569	836	842	1,006	564	841	1,027	568	846	1,034				
14. GACH 8	.11**	.10*	.20**	.19**	.16**	.26**	.36**	.56**	.29**	.43**	.50**	.51**	.59**	—	4.17	0.84	
n	557	561	566	824	556	558	567	816	557	571	834	559	569	842			

Note. All statistics were derived from the original measures (i.e., before standardization within classes); all measures ranged from 1 to 5 except mathematics and verbal achievement: 1 (*worst*) to 6 (*best*).
* $p < .05$. ** $p < .01$.

self-concept variables across domains (average correlation $r = .02$, range $r = -.09$ to .12).

Measurement invariance. Table 2 documents the results of tests of measurement invariance for both constructs across time. An examination of model fit indices and difference-testing statistics between more and less restricted models suggested that a model with scalar invariance provided a reasonable approximation to the data. Thus, subsequent analyses were conducted on models that imposed scalar invariance.

Testing the RI/E model. In the following two sections, we present the results of the four-wave cross-lagged panel design to examine the RI/E model. We began by probing the assumption of stationarity because, when stationarity holds, it allows for a more parsimonious description of the manifold empirical relations as obtained for the RI/E model.

Stationarity of effects. All *a, c,* and *d* paths (see Figure 2) could be successfully restricted to equality across time. Table 3 presents the model fit statistics; all models had acceptable to good fit. However, the first *b* path (i.e., from fifth-grade mathematics achievement to the subsequent mathematics self-concept at Time 2) was significantly different in magnitude than the second and third *b* paths, which, for their part, showed similar magnitudes across time and were, therefore, set equal (see Table 3). This indicates that skill development effects (i.e., the *b* and *d* paths) were similar in size over time except for the first *b* path (i.e., within the domain of mathematics). Furthermore, all internal frame of reference effects (i.e., the *a* and *c* paths) fit the assumption of stationarity across time. Thus, further analyses were conducted on a model that imposed stationarity across time (i.e., Model C; see Table 3).

RI/E model effects. Figure 3 illustrates the results for testing the RI/E model predictions (i.e., standardized path coefficients; for information on model fit, see Table 3, i.e., Study 1, Model C).

Skill development effects and self-enhancement effects. In both domains, the results provided support for skill development and self-enhancement effects across time. Within the same domain, prior achievements showed significant and positive relations with subsequent academic self-concepts, whereas prior academic self-concepts showed significant and positive relations with subsequent achievements. Considering the self-enhancement paths in the first time period in both domains (i.e., between Time 1 and the midterm reports in the fifth grade), the absolute magnitude of the resulting path coefficients was descriptively slightly higher than in the subsequent waves. This pattern may be due to missing baseline measures (either for academic self-concepts or for achievements in any domain).

Autoregressive effects. All autoregressive paths across time were positive and significant, pointing to an overall moderate to high level of temporal stability of academic self-concepts and achievements in both domains (see Figure 3).

Internal frame of reference effects. The results provided robust support for negative longitudinal internal frame of reference effects across domains. All associations between achievements and subsequent academic self-concepts across domains differed significantly from zero. The absolute magnitudes of the resulting coefficients were descriptively higher for paths from mathematics achievement to the subsequent verbal self-concept than from verbal achievement to the subsequent mathematics self-concept (see Figure 3).

Reciprocal cross-domain effects between achievement measures. Prior academic achievement in one domain showed a clear positive association with subsequent achievement in the other domain and vice versa. This pattern was found for both time periods (i.e., between the fifth and sixth grades as well as between the sixth and eighth grades).

Table 2
Results of Tests of Measurement Invariance for Mathematics Self-Concept (MSC) and Verbal Self-Concepts (German: GSC, English: ESC) Across Four Points of Measurement in Studies 1, 2, and 3

Invariance	*N*	χ^2 (*df*)	χ^2/df	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (*df*)
Study 1							
Configural MSC	1,519	86.134 (30)**	2.871	.989	.035 (.027–.044)	.031	
Configural GSC	1,515	47.641 (30)*	1.588	.996	.020 (.008–.030)	.022	
Baseline (simultaneous)	1,522	530.629 (188)**	2.822	.968	.035 (.031–.038)	.039	
Metric	1,522	543.893 (200)**	2.719	.968	.034 (.030–.037)	.040	11.432 (12)
Scalar	1,522	547.691 (212)**	2.583	.969	.032 (.029–.036)	.040	0.245 (12)
Study 2							
Configural MSC	636	50.135 (30)*	1.671	.993	.032 (.015–.048)	.024	
Configural GSC	636	41.953 (30)	1.398	.992	.025 (.000–.042)	.023	
Baseline (simultaneous)	636	327.335 (188)**	1.741	.971	.034 (.028–.040)	.048	
Metric	636	340.627 (200)**	1.703	.971	.033 (.027–.039)	.048	12.481 (12)
Scalar	636	342.295 (212)**	1.615	.973	.031 (.025–.037)	.048	0.138 (12)
Study 3							
Configural MSC	465	48.563 (30)*	1.619	.991	.036 (.015–.055)	.027	
Configural GSC	465	31.029 (30)	1.034	1	.009 (.000–.036)	.024	
Configural ESC	465	29.785 (30)	.993	1	.000 (.000–.035)	.024	
Baseline (simultaneous)	465	864.023 (474)**	1.823	.932	.042 (.038–.046)	.052	
Metric	465	879.975 (492)**	1.789	.933	.041 (.037–.046)	.052	14.820 (18)
Scalar	465	886.598 (510)**	1.738	.935	.040 (.035–.044)	.052	4.457 (18)

Note. CFI = comparative fit index; RMSEA = root-mean-square error of approximation; CI = confidence interval; SRMR = standardized root-mean-square residual.
* $p < .05$. ** $p < .01$.

Table 3
Results for Testing the Assumption of Stationarity (i.e., Studies 1, 2, and 3)

Model	N	χ^2 (df)	χ^2/df	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (df)
Study 1							
Model A (baseline)	1,529	960.210 (348)**	2.759	.959	.034 (.031–.036)	.046	
Model B (a to d paths restricted)[a, c]	1,529	982.435 (359)**	2.737	.959	.034 (.031–.036)	.046	22.562 (11)*
Model C (a to d paths restricted)[b, c]	1,529	968.809 (358)**	2.706	.959	.033 (.031–.036)	.046	9.407 (10)
Study 2							
Model A (baseline)	639	574.680 (348)**	1.651	.963	.032 (.027–.037)	.052	
Model B (a to l paths restricted)[a, c]	639	590.880 (363)**	1.628	.962	.031 (.027–.036)	.054	16.607 (15)
Study 3							
Model A (baseline)	465	1,343.769 (816)**	1.641	.930	.037 (.034–.041)	.053	
Model B (corresponding paths restricted)[a, c]	465	1,391.745 (852)**	1.634	.928	.037 (.033–.040)	.056	48.617 (36)

Note. CFI = comparative fit index; RMSEA = root-mean-square error of approximation; CI = confidence interval; SRMR = standardized root-mean-square residual.
[a] All corresponding paths were constrained to be equal across time (see Figure 2 for Study 1 and Figure 4 for Study 2). [b] Corresponding a (first, second, and third), b (only second and third), c (first, second, and third), and d (first, second, and third) paths were constrained to be equal across time (see Figure 2). [c] The following corresponding cross-sectional residual correlations were constrained to be equal across time to test the assumption of stationarity: domain-specific self-concept's residual with cross-domain self-concept's residual (at Times 2, 3, and 4) and achievement's residual with cross-domain achievement's residual (of Grades 6 and 8).
* $p < .05$. ** $p < .01$.

Cross-domain effects between academic self-concept measures. Here, we found a slightly different pattern of results for verbal and mathematics self-concepts. When controlling for baseline measures, both paths between prior verbal self-concepts (i.e., at Times 2 and 3) and subsequent mathematics self-concepts (i.e., at Times 3 and 4) were negative and statistically significant, thus indicating rather small but significant contrast effects across 1- and 2-year periods, respectively. Of note, when no baseline measure was considered, the negative path coefficient representing the effect of verbal self-concept at Time 1 on mathematics self-concept at Time 2 did not reach significance. However, regarding the associations between prior mathematics self-concepts and subsequent verbal self-concepts, the results suggested a null effect: No relation was statistically significant.

Cross-domain contrast effects of academic self-concepts on achievements. First, all cross-domain associations between prior academic self-concept and subsequent achievement were found to be negative, yet two paths were statistically significant: (a) between verbal self-concept and subsequent mathematics achievement for the third time span (i.e., between Time 3 and eighth-grade achievement) and (b) between mathematics self-concept and subsequent verbal achievement for the second time span (i.e., between Time 2 and sixth-grade achievement; see Figure 3).

Study 2

In Study 2, we analyzed whether the findings from Study 1 could be (a) replicated in another sample and (b) generalized across different time lags between performance feedback and the assessment of academic self-concepts. Additionally, Study 2 was based on a wave of data collection in the seventh grade (i.e., Time 4; instead of the eighth grade as in Study 1). Furthermore, the panel design of Study 2 opened up additional opportunities for examining the stability of RI/E model associations across time. In the following Method section, we mainly highlight the ways in which Study 2 differed from Study 1.

Method

Sample and procedure. The data from Sample 2 were obtained from four waves of measurement from a longitudinal study that focused on student development in secondary school. The sample comprised data from 639 students (44.44% female) in 26 classes at seven different top-track schools (i.e., Gymnasium) in two federal states of Germany (i.e., Baden-Wuerttemberg and Bavaria). Students in the sample were examined across a 29-month period: at the beginning of Grade 5 (5 to 6 weeks after the students' first day in secondary school; Time 1), at the end of Grade 5 (Time 2), at the end of Grade 6 (Time 3), and in the middle of Grade 7 (Time 4). Students stayed with the same peers throughout the study, and they were all enrolled in the same type of mathematics, German, and English classes. Their mean age was 10.70 years (SD = 0.43) at Time 1, 11.27 years (SD = 0.42) at Time 2, 12.26 years (SD = 0.43) at Time 3, and 13.11 years (SD = 0.43) at Time 4. Measurement instruments were administered during regular classes, and students' participation was voluntary. Written parental consent was obtained for all participating children.

Variables and measures.

Academic self-concepts. To assess mathematics and German self-concepts, six items based on the SDQ (as in Study 1; see above) were selected. To assess mathematics self-concept, we used the same three items as in Study 1. For the assessment of German self-concept, three different SDQ items were applied: "I learn things quickly in German class," "I'm hopeless in German class" (the item was reverse scored in all further analyses), and "I get good marks in German." These three items were already implemented in the PISA study (Organisation for Economic Co-operation and Development [OECD], 2001). Students responded to these items on a 5-point rating scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) with higher scores indicating higher academic self-concept.

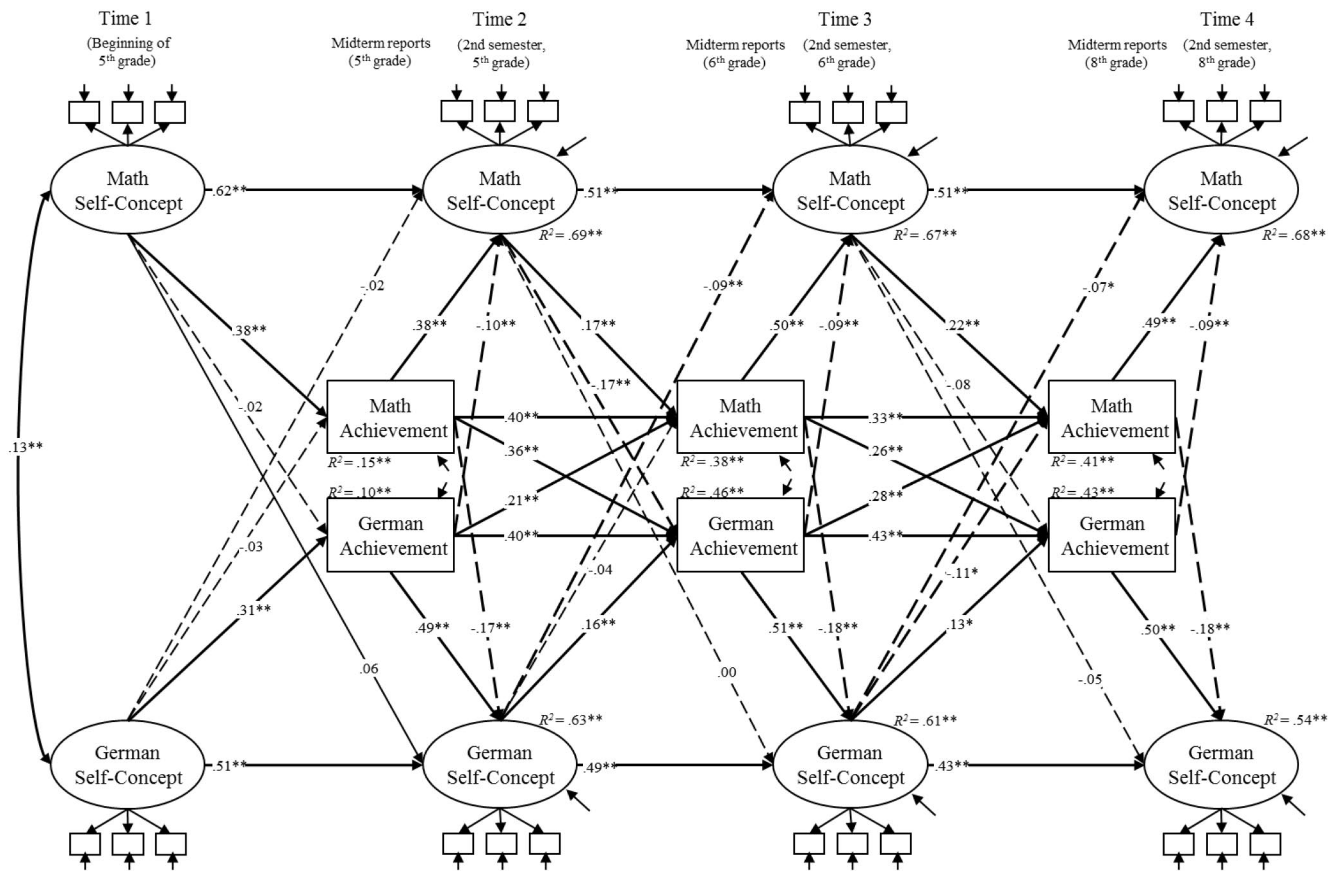

Figure 3. Study 1: Reciprocal internal/external frame of reference model in the German (as verbal subject) and mathematics domains. For the sake of clarity, correlational paths between residuals were omitted from the path diagram. Standardized model parameters are shown. * $p < .05$. ** $p < .01$.

Academic achievements. Students' achievements in mathematics and German were assessed by teacher-assigned grades taken *directly* from the students' midterm records from Grade 5, Grade 6, and Grade 7 (rather than by self-reported grades as in Study 1). Again, we scored grades such that higher numbers indicate better mathematics and German achievement (i.e., 1 = *unsatisfactory* to 6 = *excellent*).

Data analysis. We used the same SEM approach as in Study 1. Herein, academic self-concept measures were specified as latent measures with three indicators each and academic achievement as a single-indicator manifest measure. Correlated uniquenesses between corresponding items across time were allowed. Figure 4 illustrates the model. Of note, the time spans between the waves of measurement differed from those in Study 1. Corresponding paths that have time spans of equal length are indicated by identical letters (considered for testing the assumption of stationarity). In Study 2, we investigated three subsequent time spans of (a) 7 months (i.e., from the beginning of the Gymnasium to the end of the fifth grade), (b) 1 year (i.e., from the end of the fifth grade to the end of the sixth grade), and (c) 10 months (i.e., from the end of the sixth grade to the middle of the seventh grade). As students always received their performance feedback at the end of the first term in Grades 5, 6, and 7 (i.e., from midterm records), the time lags between the performance feedback and the subsequent assessments of academic self-concepts at Times 2 and 3 were about 3 months, and the time lags between assessments of self-concepts at Times 2 and 3 and subsequent academic achievements were about 9 months. The time lags between adjacent performance feedback in Grades 5, 6, and 7 were all of equal length. The assumption of stationarity was tested for (a) skill development effects within the fifth and sixth grades (i.e., *b* and *d* paths; all paths are illustrated in Figure 4), (b) self-enhancement effects (i.e., *e* and *f* paths), (c) autoregressive effects of achievements (i.e., *h* and *l* paths), (d) internal frame of reference effects (i.e., *a* and *c* paths), (e) cross-domain contrast effects of academic self-concepts (i.e., *g* and *j* paths), and (f) reciprocal cross-domain effects between achievements (i.e., *k* and *i* paths).

Similar to Study 1, we (a) standardized our measures within classes prior to our main analyses, (b) tested for measurement invariance, and (c) tested the assumption of stationarity. Overall, we used the same statistical rationale (as described in the Method section of Study 1). Concerning all academic self-concept measures, 64% to 96% of the students provided data to estimate a certain variance or covariance at any time of measurement. The proportion of data present for the German and mathematics achievement measures in the fifth and sixth grades was between 29% and 57% (i.e., missing by design because

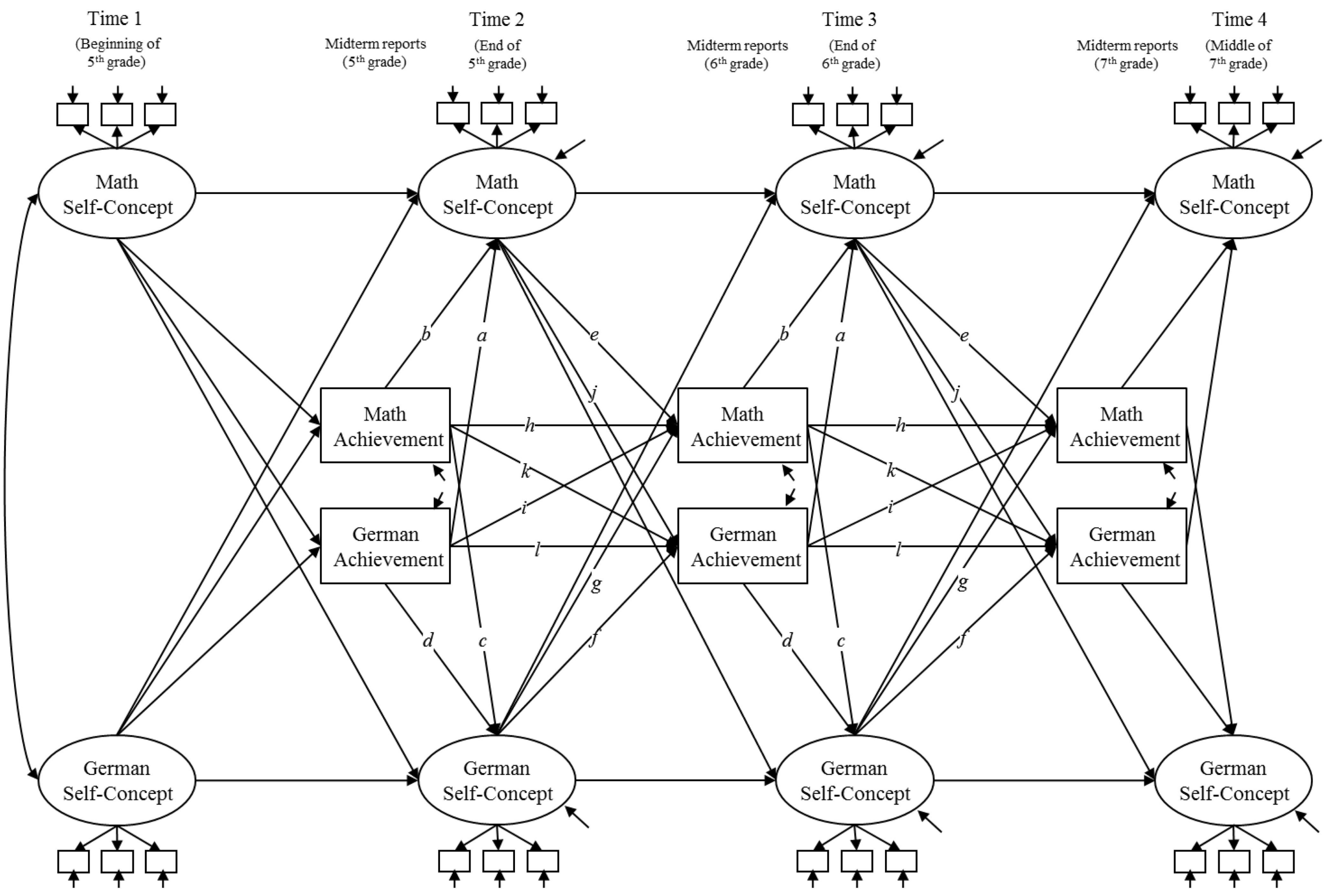

Figure 4. Study 2: Reciprocal internal/external frame of reference model in the German (as verbal subject) and mathematics domains. The *a, b, c, d, e, f, g, h, i, j, k,* and *l* paths represent lags of the same length. For the sake of clarity, correlational paths between residuals were omitted from the path diagram.

teacher-assigned school grades in Grades 5 and 6 were collected in only a subsample of the examined classes).[2] For the seventh-grade achievement measures, we had complete data from 43% to 79% of the students.

Results

Preliminary analyses.

Descriptive statistics. Table 4 illustrates the correlations, means, standard deviations, and internal consistencies for all measures at each wave of measurement. Cronbach's α was satisfactory to good (between .71 and .86). Overall, the correlations were similar to those from Study 1 (cf. Table 1) and in line with prior I/E model findings (average correlations between the academic self-concept variables: $r = .03$, range $r = -.04$ to .10; average correlation between the academic achievement variables: $r = .33$, range $r = .17$ to .47).

Measurement invariance. Table 2 shows the results of the tests of measurement invariance. A model with scalar invariance provided a reasonable approximation to the empirical data. Therefore, the following analyses were computed on models that imposed scalar invariance.

Testing the RI/E model.

Stationarity of effects. Results for tests of the assumption of stationarity are presented in Table 3. All corresponding paths (i.e., *a, b, c, d, e, f, g, h, i, j, k,* and *l* paths; see Figure 4) could be restricted to equality without losses in model fit. For all effects under investigation, these results clearly indicated stationarity across time.

RI/E model effects. Results of the specified RI/E model predictions are shown in Figure 5 (i.e., standardized path coefficients; for information on model fit, see Table 3, i.e., Study 2, Model B).

Skill development effects and self-enhancement effects. The results provided empirical support for the predicted skill development and self-enhancement effects in both domains. Within each domain and consistently across time, prior achievements showed positive relations with subsequent academic self-concepts. Prior academic self-concepts, in turn, showed positive relations with subsequent achievements.

Autoregressive effects. As shown in Figure 5, the results supported autoregressive effects and suggested a moderate to high level of temporal stability for all included variables.

Internal frame of reference effects. All path coefficients between prior mathematics achievements and subsequent German self-concepts were significant and negative across time, thus sup-

[2] In Grade 5, we obtained achievement data from 18 classes, in Grade 6 from 13 classes (total: 26 classes).

Table 4
Study 2: Observed Correlations (r), Means (M), Standard Deviations (SD), and Reliabilities (Cronbach's α) for Mathematics and German Self-Concept Scales (MSC/GSC) and Mathematics and German Achievements (Teacher-Assigned School Grades; MACH/GACH)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	*M*	*SD*	α
1. MSC Time 1	—														3.97	0.98	.86
n	619																
2. MSC Time 2	.55**	—													3.68	1.05	.85
n	586	596															
3. MSC Time 3	.48**	.57**	—												3.44	1.07	.85
n	522	515	539														
4. MSC Time 4	.46**	.53**	.64**	—											3.45	1.08	.86
n	441	431	437	456													
5. GSC Time 1	.03	.01	.00	−.04	—										4.03	0.76	.75
n	619	586	522	441	619												
6. GSC Time 2	−.02	.07	.04	.08	.51**	—									3.86	0.84	.77
n	586	596	515	431	586	596											
7. GSC Time 3	.00	.08	.10*	−.02	.32**	.46**	—								3.76	0.81	.71
n	522	515	539	437	522	515	539										
8. GSC Time 4	−.02	.07	.07	.08	.33**	.51**	.46**	—							3.84	0.82	.76
n	441	431	437	456	441	431	437	456									
9. MACH 5	.29**	.49**	.34**	.38**	.09	.08	−.02	.00	—						4.69	0.74	
n	353	341	323	283	353	341	323	283	364								
10. MACH 6	.26**	.35**	.50**	.44**	.04	.12	.09	.00	.66**	—					4.31	0.92	
n	283	275	259	215	283	275	259	215	187	294							
11. MACH 7	.23**	.30**	.39**	.56**	.00	.08	−.01	.02	.48**	.67**	—				4.06	0.90	
n	483	470	455	414	483	470	455	414	340	276	502						
12. GACH 5	−.08	.02	−.12*	−.07	.32**	.39**	.21*	.24**	.37**	.18*	.17**	—			4.52	0.73	
n	353	341	323	283	353	341	323	283	364	187	340	364					
13. GACH 6	−.04	−.02	.02	.03	.25**	.26**	.33**	.29**	.31**	.47**	.32**	.53**	—		4.33	0.76	
n	283	275	259	215	283	275	259	215	187	294	276	187	294				
14. GACH 7	−.05	.06	.07	.08	.18**	.26**	.18**	.46**	.31**	.39**	.41**	.42**	.55**	—	4.33	0.74	
n	483	470	455	414	483	470	455	414	340	276	502	340	276	502			

Note. All statistics were derived from the original measures (i.e., before standardization within classes); all measures ranged from 1 to 5 except mathematics and verbal achievement: 1 (*worst*) to 6 (*best*).
* $p < .05$. ** $p < .01$.

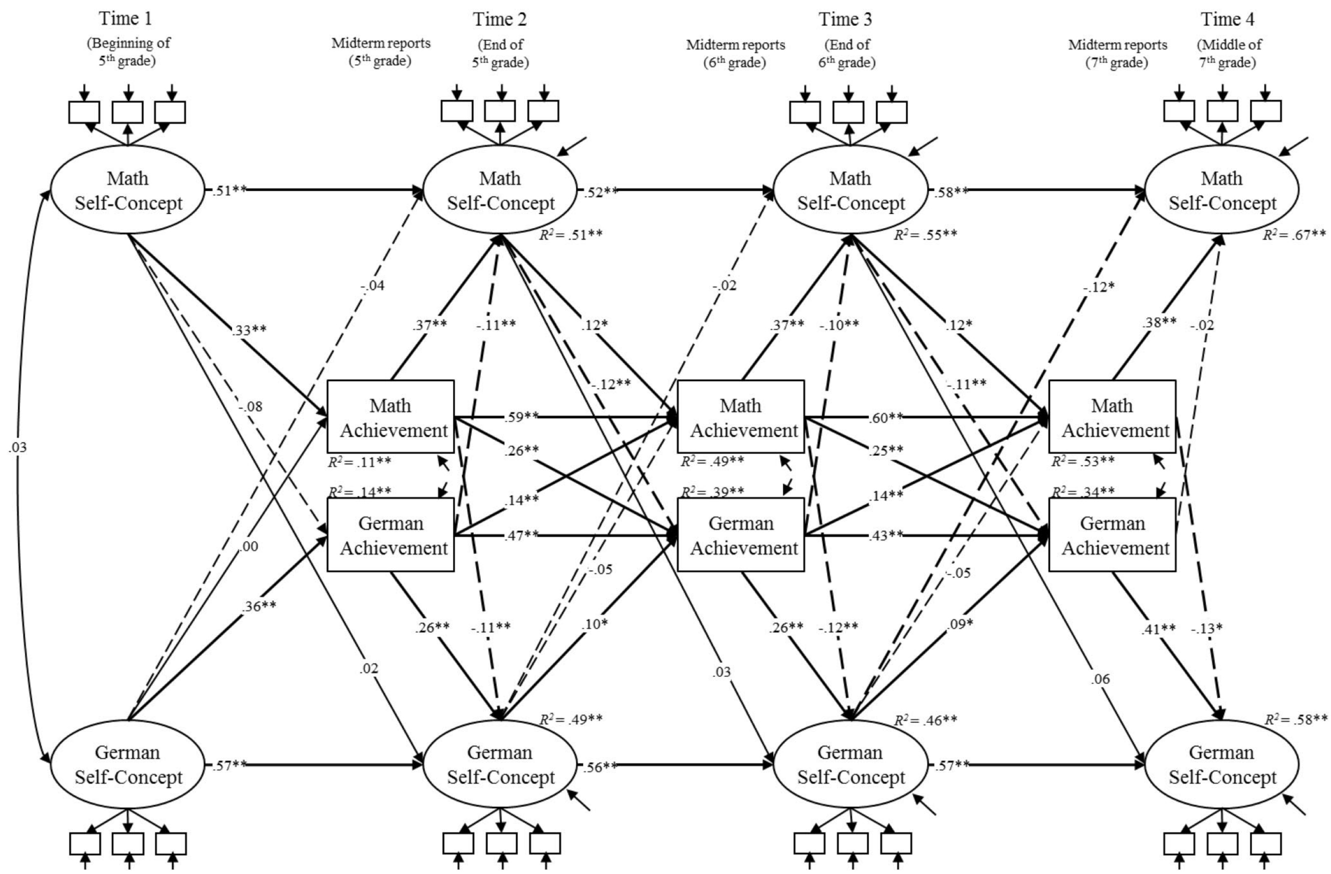

Figure 5. Study 2: Reciprocal internal/external frame of reference model in the German (as verbal subject) and mathematics domains. For the sake of clarity, correlational paths between residuals were omitted from the path diagram. Standardized model parameters are shown. $^{*}p < .05$. $^{**}p < .01$.

porting the prediction of internal frame of reference effects. For the relations between German achievements and subsequent mathematics self-concepts, two of the three tested paths were significant. In sum, the results predominantly supported longitudinal internal frame of reference effects.

Reciprocal cross-domain effects between achievement measures. In line with RI/E model predictions, preceding achievements in one domain were significantly and positively related to subsequent achievements in the other domain.

Cross-domain effects between academic self-concept measures. All paths between preceding German self-concepts and subsequent mathematics self-concepts were negative, but only one of them was significantly different from zero. The path between prior German self-concept at Time 3 (i.e., the end of the sixth grade) and subsequent mathematics self-concept at Time 4 (i.e., in the middle of the seventh grade) featured a small significant negative effect. As in Study 1, none of the paths between preceding mathematics self-concepts and subsequent German self-concepts reached significance.

Cross-domain contrast effects of academic self-concepts on achievements. All resulting coefficients of the cross-domain paths between prior academic self-concepts and subsequent achievements were negative (except for the first path between verbal self-concept and mathematics achievement, which was found to be zero). When controlling for baseline measures of achievement, mathematics self-concepts were negatively and significantly related to subsequent German achievements (i.e., between Time 2 and sixth-grade achievement and between Time 3 and seventh-grade achievement) suggesting cross-domain contrast effects of mathematics self-concept. However, the relations between German self-concepts and subsequent mathematics achievements failed to reach statistical significance.

Study 3

In Study 3, we extended the RI/E model framework. In drawing on a subsample of Study 2, we specified an RI/E model across three domains: one numerical domain (i.e., mathematics) and two verbal (i.e., German and English) domains. Herein, we analyzed students who chose English as their first foreign language in secondary school and were native German speakers. In the following Method section, again, we mainly highlight the ways in which Study 3 differed from Studies 1 and 2.

Method

Sample and procedure. To study RI/E model effects across three domains, we selected those students from Sample 2 who chose

English as their first foreign language in secondary school and were native German speakers. Five hundred thirty-four students chose English as their first foreign language (98 students chose Latin; seven did not provide information); herein, 470 were native German speakers (five of these students were not considered in our analyses because they had both German and English as mother tongues). The resulting subsample comprised data from 465 students (42.37% female) in 22 classes from two German federal states (i.e., Baden-Wuerttemberg and Bavaria). Their mean age was 10.71 years ($SD = 0.44$) at Time 1, 11.28 years ($SD = 0.43$) at Time 2, 12.27 years ($SD = 0.44$) at Time 3, and 13.12 years ($SD = 0.44$) at Time 4.

Variables and measures.

Academic self-concepts. To assess mathematics and German self-concepts, we used the same three items as in Study 2. For the assessment of English academic self-concept, we used the following three SDQ items: "I learn things quickly in English class," "I'm hopeless in English class" (the item was reverse scored in all further analyses), and "I get good marks in English." These three items were already implemented in the PISA study (OECD, 2001). Again, students responded to these items on a 5-point rating scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) with higher scores indicating higher academic self-concept.

Academic achievements. As in Study 2, students' achievements in mathematics, German, and English were assessed by teacher-assigned grades taken directly from the students' midterm records from Grade 5, Grade 6, and Grade 7. We scored grades such that higher numbers indicate better mathematics, German, and English achievement (i.e., 1 = *unsatisfactory* to 6 = *excellent*).

Data analysis. We used the same SEM approach as in Study 2 (as described above). Prior to our main analyses, we (a) standardized our measures within classes, (b) tested for measurement invariance, and (c) tested the assumption of stationarity for 27 corresponding longitudinal paths (i.e., for effects and time spans equivalent to those in Study 2 but extended across three domains). At least 64% to at most 99% of the students (i.e., concerning all academic self-concepts and seventh-grade achievement measures) and at least 32% to at most 56% of the students (i.e., concerning the fifth- and sixth-grade achievement measures) provided data to estimate a certain variance or covariance at any time of measurement.

Results

Preliminary analysis.

Descriptive statistics. Table 5 illustrates the correlations, means, standard deviations, and internal consistencies for all measures at each wave of measurement. Cronbach's α ranged from satisfactory to good (between .71 and .86). The average correlation between the mathematics and German self-concept variables was $r = .02$ (range $r = -.04$ to .11), between mathematics and English self-concept $r = .06$ (range $r = -.02$ to .22), and between German and English self-concept $r = .29$ (range $r = .15$ to .44). The average correlation between the mathematics and German achievement variables was $r = .28$ (range $r = .09$ to .43), between mathematics and English achievement $r = .38$ (range $r = .27$ to .49), and between German and English achievement $r = .44$ (range $r = .26$ to .64).

Measurement invariance. Table 2 shows the results of the tests of measurement invariance. Again, a model with scalar invariance was shown to provide a reasonable approximation to the empirical data. Further analyses imposed scalar invariance.

Testing the RI/E model.

Stationarity of effects. Results for tests of the assumption of stationarity are presented in Table 3. All 27 corresponding paths were successfully restricted to equality without losses in model fit, thus indicating stationarity of all tested effects across time.

RI/E model effects. Results of the specified RI/E model predictions are shown in Table 6 (i.e., standardized path coefficients; for information on model fit, see Table 3, i.e., Study 3, Model B).[3]

Skill development effects and self-enhancement effects. The results clearly supported the predicted skill development and self-enhancement effects in simultaneous examinations of the domain of mathematics and the domains of German and English. We found these predicted effects within each domain and consistently across time.

Autoregressive effects. All autoregressive effects reached statistical significance. The stability of academic self-concepts and achievements was moderate to high.

Internal frame of reference effects. All path coefficients between prior mathematics achievements and subsequent verbal self-concepts (i.e., German *and* English self-concepts) failed to reach statistical significance. Furthermore, English achievements were not significantly associated with subsequent mathematics self-concepts. These results contradict the prediction of internal frame of reference effects. By contrast, two of the three tested paths between prior English achievements and subsequent German self-concepts were negative and significant. The same applied for the longitudinal relations between prior German achievements and subsequent mathematics self-concepts. Again, two of the three tested paths were negative and significant. Concerning the internal frame of reference paths between German achievements and subsequent English self-concepts, the results were not consistent across time: In the first two time slots, the associations were significant and negative; in the third time slot, however, this association was significant and positive. In sum, the results provide some support for longitudinal internal frame of reference effects across one numeric and two verbal domains. All but one of the seven significant paths were found to be negative.

Reciprocal cross-domain effects between achievement measures. Except for the longitudinal association between German achievements (i.e., Grades 5 and 6) and subsequent mathematics achievements (i.e., Grades 6 and 7), all cross-domain paths between achievement measures were positive and significant and had small to moderate effect sizes.

Cross-domain effects between academic self-concept measures. Without considering any baseline measure, the path between prior German self-concept at Time 1 and subsequent English self-concept at Time 2 was significant and positive. Prior mathematics self-concept at Time 3 and subsequent English self-concept at Time 4 were significantly negatively related. None of the others cross-domain

[3] For the sake of clarity, we used a table to depict the resulting standardized coefficients of the more complex extended RI/E model (vs. a figure as was done in Studies 1 and 2).

Table 5
Study 3: Observed Correlations (r), Means (M), Standard Deviations (SD), and Reliabilities (Cronbach's α) for Mathematics, German, and English Self-Concept Scales (MSC/GSC/ESC) and Mathematics, German, and English Achievements (Teacher-Assigned School Grades; MACH/GACH/EACH)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	*M*	*SD*	α
1. MSC Time 1	—																					4.00	0.98	.85
n	459																							
2. MSC Time 2	.57**	—																				3.74	1.05	.86
n	438	444																						
3. MSC Time 3	.49**	.58**	—																			3.46	1.09	.86
n	388	385	394																					
4. MSC Time 4	.51**	.56**	.63**	—																		3.46	1.10	.86
n	323	318	314	328																				
5. GSC Time 1	.00	−.02	−.02	−.04	—																	4.05	0.76	.73
n	459	438	388	323	459																			
6. GSC Time 2	−.03	.07	.02	.04	.53**	—																3.89	0.83	.79
n	438	444	385	318	438	444																		
7. GSC Time 3	−.01	.08	.11*	−.03	.30**	.44**	—															3.76	0.81	.71
n	388	385	394	314	388	385	394																	
8. GSC Time 4	−.03	.04	.08	.08	.29**	.47**	.46**	—														3.88	0.82	.80
n	323	318	314	328	323	318	314	328																
9. ESC Time 1	.16**	.01	−.02	.07	.44**	.25**	.15**	.17**	—													4.24	0.75	.77
n	459	438	388	323	459	438	388	323	459															
10. ESC Time 2	.02	.11*	.07	.02	.27**	.40**	.27**	.23**	.40**	—												4.14	0.83	.77
n	438	444	385	318	438	444	385	318	438	444														
11. ESC Time 3	.00	.10	.22**	−.01	.22**	.27**	.42**	.33**	.28**	.50**	—											4.02	0.87	.78
n	388	385	394	314	388	385	394	314	388	385	394													
12. ESC Time 4	−.01	.03	.01	.13*	.23**	.30**	.28**	.44**	.29**	.49**	.55**	—										3.99	0.78	.78
n	323	318	314	328	323	318	314	328	323	318	314	328												
13. MACH 5	.32**	.51**	.32**	.37**	.08	.09	.02	−.01	.13*	.14*	.03	.01	—									4.78	0.71	
n	256	247	232	199	256	247	232	199	256	247	232	199	261											
14. MACH 6	.28**	.34**	.47**	.41**	−.01	.12	.08	−.01	.08	.25**	.17*	.08	.65**	—								4.35	0.88	
n	232	224	208	168	232	224	208	168	232	224	208	168	148	236										
15. MACH 7	.28**	.32**	.39**	.55**	−.01	.06	−.01	.00	.03	.07	.07	.16**	.48**	.66**	—							4.08	0.90	
n	367	358	339	304	367	358	339	304	367	358	339	304	244	221	373									
16. GACH 5	−.09	−.01	−.18**	−.12	.30**	.33**	.12	.14*	.13*	.17**	.08	.21**	.33**	.13	.09	—						4.62	0.66	
n	256	247	232	199	256	247	232	199	256	247	232	199	261	148	244	261								
17. GACH 6	−.08	−.03	−.03	.01	.27**	.25**	.30**	.26**	.13	.21**	.21**	.24**	.27**	.43**	.29**	.53**	—					4.37	0.73	
n	232	224	208	168	232	224	208	168	232	224	208	168	148	236	221	148	236							
18. GACH 7	−.06	.04	.06	.06	.17**	.23**	.16**	.40**	.11*	.15**	.20**	.32**	.30**	.35**	.36**	.37**	.50**	—				4.38	0.70	
n	367	358	339	304	367	358	339	304	367	358	339	304	244	221	373	244	221	373						
19. EACH 5	−.03	.01	−.11	.02	.10	.11	−.01	−.03	.28**	.44**	.17	.31**	.37**	.29**	.27**	.52**	.36**	.26**	—			4.87	0.71	
n	256	247	232	199	256	247	232	199	256	247	232	199	261	148	244	261	148	244	261					
20. EACH 6	−.13*	.02	−.02	.08	.18**	.24**	.14*	.17*	.22**	.44**	.43**	.46**	.36**	.45**	.38**	.31**	.64**	.51**	.43**	—		4.47	0.82	
n	232	224	208	168	232	224	208	168	232	224	208	168	148	236	221	148	236	221	148	236				
21. EACH 7	.00	.05	.04	.08	.21**	.22**	.18**	.20**	.23**	.29**	.33**	.52**	.35**	.45**	.49**	.31**	.51**	.53**	.40**	.74**	—	4.22	0.87	
n	367	358	339	304	367	358	339	304	367	358	339	304	244	221	373	244	221	373	244	221	373			

Note. All statistics were derived from the original measures (i.e., before standardization within classes); all measures ranged from 1 to 5 except mathematics and verbal achievement: 1 (*worst*) to 6 (*best*).
* $p < .05$. ** $p < .01$.

Table 6
Study 3: Standardized Model Parameters of the Reciprocal Internal/External Frame of Reference Model in the Domains of Mathematics, German (Native Language), and English (Foreign Language)

	Domains		
Paths	Mathematics	German	English
Self-concepts to achievements: Self-enhancement (bold) and cross-domain contrast effects			
Self-concepts (Time 1) to			
MACH (5th grade)	**.38****	.03	−.10
GACH (5th grade)	−.07	**.46****	−.16†
EACH (5th grade)	−.12†	−.06	**.37****
Self-concepts (Time 2) to			
MACH (6th grade)	**.14****	−.05	.00
GACH (6th grade)	−.11*	**.13**†	−.03
EACH (6th grade)	−.11*	.01	**.18****
Self-concepts (Time 3) to			
MACH (7th grade)	**.15****	−.05	.00
GACH (7th grade)	−.10*	**.10**†	−.03
EACH (7th grade)	−.10*	.01	**.17****
Achievements to self-concepts: Skill development (bold) and internal frame of reference effects			
Achievements (5th grade) to			
MSC (Time 2)	**.39****	−.10*	−.04
GSC (Time 2)	−.02	**.29****	−.13*
ESC (Time 2)	.00	−.11*	**.40****
Achievements (6th grade) to			
MSC (Time 3)	**.39****	−.09*	−.04
GSC (Time 3)	−.02	**.30****	−.13*
ESC (Time 3)	.00	−.10*	**.39****
Achievements (7th grade) to			
MSC (Time 4)	**.38****	.01	−.02
GSC (Time 4)	−.10	**.34****	.00
ESC (Time 4)	.03	.14*	**.22****
Self-concepts to self-concepts: Autoregressive (bold) and cross-domain effects			
Self-concepts (Time 1) to			
MSC (Time 2)	**.49****	.03	−.09
GSC (Time 2)	−.01	**.56****	.05
ESC (Time 2)	−.03	.19*	**.29****
Self-concepts (Time 2) to			
MSC (Time 3)	**.50****	−.01	−.01
GSC (Time 3)	.00	**.52****	.08
ESC (Time 3)	.03	.07	**.47****
Self-concepts (Time 3) to			
MSC (Time 4)	**.55****	−.09	−.01
GSC (Time 4)	.09	**.57****	−.02
ESC (Time 4)	−.14**	−.06	**.71****
Achievements to achievements: Autoregressive (bold) and reciprocal cross-domain effects			
Achievements (5th grade) to			
MACH (6th grade)	**.54****	.07	.15*
GACH (6th grade)	.21**	**.35****	.27**
EACH (6th grade)	.23**	.13*	**.47****
Achievements (6th grade) to			
MACH (7th grade)	**.55****	.07	.15*
GACH (7th grade)	.20**	**.30****	.24**
EACH (7th grade)	.23**	.12*	**.44****

Note. MSC = Mathematics Self-Concept; GSC = German Self-Concept; ESC = English Self-Concept; MACH = Mathematics Achievement; GACH = German Achievement; EACH = English Achievement.
† $p < .10$. * $p < .05$. ** $p < .01$.

paths between academic self-concept measures reached statistical significance.

Cross-domain contrast effects of academic self-concepts on achievements. All resulting coefficients of the cross-domain paths between prior mathematics self-concepts and subsequent verbal achievements (i.e., German *and* English achievements) were negative. After controlling for baseline measures, all of them reached statistical significance. The relations between verbal self-concepts (i.e., German *and* English self-concepts) and subsequent cross-domain achievements failed to reach statistical significance except for one marginally significant path: English self-concept was negatively related to subsequent German achievement (i.e., between Time 1 and fifth-grade achievement).

General Discussion

Students develop distinct cognitive and motivational profiles, and these have a large impact on individual academic trajectories and career choices. The development of such profiles comprises the core of the recently developed RI/E model (Möller et al., 2011). The RI/E model focuses on the interplay between students' academic self-concepts and their academic achievements across multiple domains. As students' academic self-concept is strongly related to motivational and affective variables (e.g., interest, academic emotions) and behavioral outcomes (e.g., course selection), academic self-concept is one key variable in educational research. More holistic than conventional self-concept models (e.g., the REM or the I/E model), the RI/E model reflects the joint development of students' motivational (i.e., self-concept) and cognitive (i.e., achievement) profiles by modeling positive reciprocal *and* negative contrast effects across time—thereby considerably advancing self-concept research.

The absence of any replication study that has addressed the entire RI/E model provided the impetus for the present research. Accordingly, our first research aim was the replication of the RI/E model to thereby probe its generalizability. In addition, our second research aim was to significantly extend the RI/E model by examining its predictions across more than two academic domains (i.e., replication-extension study; see Bonett, 2012) and to test whether corresponding empirical relations would be stable across time (stationarity). Overall, our results supported the RI/E model framework. We found evidence for both positive reciprocal effects and negative contrast effects of academic self-concepts and achievements in the long run. Concerning the generalizability of the RI/E model across independent samples, the resulting patterns of longitudinal associations were remarkably similar between Studies 1 and 2. Furthermore, the RI/E model predictions held across the academic domains investigated (i.e., across the domains of mathematics and German, mathematics and English, or German and English; see Study 3). Additionally, for the examined time span in secondary school, we found that the tested RI/E model effects were stable across time. Of note, Möller and colleagues (2011) examined students in secondary school as a whole, whereas we tested the RI/E model framework by focusing on students in the top track of the German secondary school system (i.e., German Gymnasium). Only by controlling for between-secondary-school differences can researchers investigate whether the psychological processes depicted in the RI/E model occur in all subgroups (i.e., certain secondary school types). By focusing on one track, we were therefore able to test the generalizability of the RI/E model framework for this selected group (cf. Shadish et al., 2001, for quasi-experimental designs to generalize causal inferences).

Positive Reciprocal Effects Across Time

According to the RI/E model, due to effects of skill development, students with higher achievements show better self-concept development over time. Additionally, due to self-enhancement effects, students with higher academic self-concepts show better development in their attainment.

In sum (i.e., bringing together Studies 1 and 2), we tested 12 potential longitudinal *skill development effects,* six within each domain (i.e., all of them were controlled for prior academic self-concepts). The results clearly supported these effects; all tested paths were positive and statistically significant. That is, higher achievement is beneficial for academic self-concept development even when controlling for prior academic self-concept. Furthermore, in both studies, we tested 12 potential longitudinal *self-enhancement effects,* six within each domain (i.e., eight of them were controlled for prior achievements, four within each domain). All of them were positive and significant. These findings clearly support beneficial effects of academic self-concepts on the development of achievements within one domain (even when controlling for prior achievement).

Furthermore, we found strong support for positive reciprocal effects between mathematics and German achievements. Thus, performing well in mathematics is beneficial for the development of achievement in German as well—and vice versa.

The positive RI/E model predictions also held across three domains (i.e., mathematics, German, English). For the domain of English, all longitudinal self-enhancement and skill development paths were significant and had moderate effect sizes. Furthermore, English achievements showed positive reciprocal relations with cross-domain achievements. However, different from Studies 1 and 2, the two paths between prior German achievement and subsequent mathematics achievement failed to reach statistical significance, suggesting that—due to shared variance—German achievement did not provide any additional explanatory power over English achievement in the prediction of mathematics achievement. This common variance might reflect that students learning English as a foreign language draw on their verbal ability skills in their native language (i.e., German; see Proctor, August, Carlo, & Snow, 2006).

Negative Contrast Effects Across Time

In Studies 1 and 2, we found strong support for internal frame of reference effects of prior achievement on subsequent academic self-concept in the contrasting domain (11 out of 12 paths that we tested were significant). These findings are in favor of the longitudinal extension of this central I/E model prediction. In both studies, academic self-concepts were always negatively related to subsequent cross-domain achievements. When controlling for baseline measures of achievement, four of eight tested paths differed statistically from zero. The results were even clearer when considering only the paths between mathematics self-concept and subsequent German achievement (three of four tested paths were significant). Furthermore, the observed absolute magnitude of these negative effects was found to be—at least—comparable in size to corresponding positive self-enhancement effects across the same time period (see Figures 3 and 5). Thus, thinking positively about one's abilities in mathematics seems to weaken future achievement in German. Similar dynamics—although weaker—seem to apply for German self-concept and subsequent mathematics achievement.

Furthermore, German self-concept was found to be negatively related to subsequent mathematics self-concept. When controlling for baseline measures in achievement, three of four tested paths achieved statistical significance, indicating small and negative effects. Thus, German self-concepts seem to influence subsequent mathematics self-concepts not only *indirectly* via mathematics achievements (i.e., cross-domain contrast effects on achievement) but also *directly* in the long run. The negative effects of mathematics self-concepts on subsequent German self-concepts, in turn, seem to be completely mediated by German achievements. However, in sum, our results suggest that both mathematics and German self-concepts are negatively related to changes in cross-domain self-concepts—at least indirectly.

Also across three domains, the negative RI/E model effects were predominantly confirmed. Notably, we found support for contrasting internal frame of reference effects between the two verbal domains German and English. Herein, we found five of six tested paths to be significant; all but one were negative. Thus, students seem to contrast their grades in German or English with their self-perceived ability in the other verbal subject. These results are of particular interest because the appearance of longitudinal contrast effects between two verbal domains has been repeatedly doubted in favor of positive assimilation effects (Möller et al., 2006; Xu et al., 2013; see the introduction; see also Möller & Marsh, 2013). We found no evidence for positive long-lasting effects of verbal self-concepts on cross-domain verbal achievements or vice versa. In addition, the internal frame of reference effects between the domains of mathematics and English did not reach statistical significance. After controlling for five concurrent internal frame of reference paths per examined time span, effects might be too weak to reach significance (see limitations below). Between mathematics and German, five out of six internal frame of reference paths were negative; two of them reached statistical significance. Of note, Studies 1 and 2—which could more reliably detect smaller effects (see section on limitations)—revealed internal frame of reference effects between mathematics and German more consistently. However, we found further evidence for negative cross-domain contrast effects of academic self-concepts: Mathematics self-concept was consistently negatively related to subsequent English achievement. Similarly, mathematics self-concept was significantly negatively related to subsequent German achievement (after controlling for baseline measures; cf. Studies 1 and 2). Additionally, we found a negative and significant direct path between mathematics self-concept and subsequent English self-concept (after controlling for baseline measures; without controlling for shared variances, German self-concept was found to be positively related to subsequent English self-concept).

Taken together, the predictions derived from the original RI/E model largely applied to the extended RI/E model as well. Our results suggest that RI/E model effects are generalizable across different domains. Hence, students seem to contrast different academic domains—even if both of them are verbal domains. Thus, rather than subordinating German and English into some kind of general verbal self-concept, students seem to perceive their native and foreign languages in secondary school as academically distinct domains. These findings are in line with recent research on the structure of academic self-concepts, indicating that academic self-concepts are not only highly domain specific (i.e., multidimensional) but also strongly separable across domains and school subjects within domains (i.e., nested Marsh/Shavelson model by Brunner et al., 2010).

Of note, most of the observed negative longitudinal associations involved academic self-concept as a predictor. This finding supports the existence of undesirable effects—that is, the dark side—of academic self-concepts. Even if these effects were rather small, they should not be neglected because they seem to play a central role in students' profile formation (as discussed below).

Stability of RI/E Model Effects Across Time

Overall, our results suggest that the relative impacts of achievements (i.e., skill development effects) and academic self-concepts (i.e., self-enhancement effects) on their corresponding counterparts appear to be fairly stable from fifth grade to seventh grade or from fifth grade to eighth grade. This conclusion holds true except for one tested path in Study 1 (see the Results section of Study 1). However, because we did not replicate this finding in the second sample (i.e., Studies 2 and 3), we have no evidence to generalize it. Thus, we did not interpret this single path any further. Taken together, we found no evidence for variations in the strength of (a) positive skill development effects or (b) negative internal frame of reference effects (i.e., Studies 1, 2, and 3). Furthermore, our findings suggest stability in the relative strengths of these RI/E model effects across time (i.e., Studies 2 and 3): (c) positive self-enhancement effects, (d) negative cross-domain contrast effects of academic self-concepts, (e) positive autoregressive effects (for achievements), and (f) positive reciprocal cross-domain effects between achievements. To conclude, there seem to be no phase-dependent variations in the strengths of the associations between academic self-concepts and achievements (or between adjacent achievements) across time. Thus, RI/E model dynamics seem to apply comparably across Grades 5 to 7 or 8.

Limitations

Even if the temporal ordering of the variables is specified by design, it is impossible to make causal claims on the basis of longitudinal studies alone. In the present study, we borrowed self-concept theory's common terms, which imply that causality stands behind the longitudinal interplay of academic self-concepts and achievements. However, further intervention studies are necessary to test causal RI/E model effects.

To deal with potential clustering effects, we standardized our variables within classes. In the present study, within-class standardization resulted in a fixed-effects design that permitted us to effectively control for (a) grading standards and (b) selection processes between classes (see the methodological advice for making causal inferences in educational settings by Legewie, 2012). Thus, we were able to compare the observed effects between students in different classes. As a consequence of within-class standardization, we slightly restricted our variances, which may have resulted in underestimates of the underlying RI/E model effects. Therefore, the present replication-extension study presents a rather conservative test of the RI/E model (Legewie, 2012).

As Möller and colleagues (2011) did, we analyzed German samples. Thus, further replications outside the German educational system are needed to probe the cross-cultural generalizability of the predictions of the RI/E model in other educational systems.

In all three studies, we implemented teacher-assigned grades as academic achievement measures. Overall, school grades can be regarded as the most relevant achievement indicator for self-concept genesis because regular performance feedback with regard to standardized ability test scores is usually not given to the students (in Germany). Of note, we were able to attenuate potential biases due to different grading standards between classes because we standardized our measures within classes (see above).

The multivariate focus of the current study is reflected by the inclusion of several constructs plus students' achievements in one model. Such integrative models have several advantages for the examination of academic self-concept development compared with models that consider only the mutual relation between (global or one specific facet of) academic self-concept and achievement. Notably, numerous longitudinal relations can be investigated while controlling for shared variances. Besides methodological advantages, more integrative models also feature conceptual advantages: RI/E model research offers genuine and novel insights into the dynamics of students' motivational and cognitive profile formation—also in hinting at academic self-concept's dark side, which has been largely neglected in prior research (see the introduction). Of note, more integrative models are also more complex models. In Study 3, we investigated three domains. Here, statistical power might be an issue for two reasons. First, in Study 3, we used a subsample of Study 2. Thus, as a consequence, we estimated the more complex model in Study 3 with a smaller sample than we did for the less complex model in Study 2. Therefore, Study 3 had less statistical power to estimate the single RI/E model predictions, resulting in a more conservative test than in Studies 1 and 2. Second, the power was further restricted by the fact that we did not have complete data for all students at any time of measurement (power was considerably less of an issue in Studies 1 and 2, where more data were available to estimate the single RI/E model predictions). Thus, Study 3 offered less reliable tests to detect smaller effects. Nevertheless, all RI/E model estimates were properly estimated in Study 3 and were—above all—in line with the RI/E model predictions.

Conclusion

"Individuals with good mathematics skills also tend to have good verbal skills and vice versa, but people think of themselves as either 'math' persons or 'verbal' persons—but not both" (Marsh & Hau, 2004, p. 57). In fact, the RI/E model provides important insights into the underlying mechanisms of becoming a math or verbal person—or even more subject-specifically, for example, a person who has good native language skills or is good at learning foreign languages. In particular, the RI/E model accentuates the propulsive role of academic self-concepts in developing distinct motivational and cognitive profiles. Whereas students' academic attainment for its own sake seems to benefit further attainment even in noncorresponding domains, academic self-concepts seem to strengthen achievement in corresponding domains but simultaneously weaken those in domains that are assumed to contrast with each other. Furthermore, we found some support that academic self-concepts are *directly* and negatively longitudinally related to cross-domain self-concepts. This, in turn, would additionally stimulate students' profile formation. Taken together, thinking of oneself as being either a math or verbal person seems to enhance future corresponding achievements, but—at the same time—it also seems to weaken future noncorresponding achievements *and* self-concepts (at least indirectly).

One could assume that these contrasting RI/E model effects would gain even more importance when students are free to make academic choices (e.g., course selections in more advanced grades in high school or German *Gymnasiale Oberstufe*) or to choose a particular career. Hence, academic self-concepts probably benefit niche-picking processes wherein students actively choose self-congruent environments (i.e., environments that fit their personalities; Caspi, Roberts, & Shiner, 2005). Thus, the RI/E model framework might also be viewed as a promising heuristic for counseling. According to Eccles's (1994) expectancy-value model of educational and occupational choices, peoples' career choices mainly depend on a function of perceptions of task values and expectations for success, wherein the latter are modeled by academic self-concepts. Thus, the RI/E model can substantially refine central facets of Eccles's model by shedding light on underlying developmental dynamics of academic self-concepts in contrasting domains. Other models reflecting academic profiles have also considered cognitive ability (see, e.g., Lubinski & Benbow, 2000). Herein, understanding RI/E model dynamics should help researchers to better understand the gaps between students' actual and perceived abilities.

Hence, an intriguing question that arises is, What would happen if the RI/E model framework was fully integrated with complementary self-regulatory models (e.g., Eccles, 1994; Lubinski & Benbow, 2000)? Thus, an examination of the RI/E model that takes into consideration further cognitive variables (e.g., domain-specific ability), variables of perceived task value (e.g., academic interest), and attribution styles (e.g., performing well in mathematics and German as a result of ability or of effort) would more completely depict the process of self-regulation in educational contexts. Additionally, potential mediators of the longitudinal relations among academic self-concepts and achievements should be investigated. In fact, the underlying modes of action of the influence of academic self-concept on achievements are still unclear to some extent (Möller et al., 2011). A plausible explanation would be that high domain-specific academic self-concepts result in more engagement and time invested in corresponding academic domains (Marsh & Yeung, 1997; Möller et al., 2011). In fact, some research supports these predictions (see Trautwein & Lüdtke, 2007; Trautwein, Lüdtke, Roberts, Schnyder, & Niggli, 2009). Certainly, there is a need to more rigorously examine the potential mediators of the longitudinal relations between academic self-concepts and achievements (Marsh, 2006). At any rate, over and above feedback regarding one's own *actual* achievements, academic self-concepts seem to influence students' perceptions of their learning environment and, thus, their overall engagement in learning opportunities.

Marsh and colleagues (2005) already integrated academic interests into the REM model framework. However, an integration of academic interest (and/or further cognitive/affective variables as mediators or moderators) in the more holistic RI/E model framework still remains to be implemented.

A practical implication of the present study clearly concerns the question of how students can be made to feel good about themselves in several academic domains at the same time. Meta-analytical research on self-concept interventions revealed that interventions focusing on praise and/or feedback yielded the highest effect sizes ($d = 1.13$; O'Mara, Marsh, Craven, & Debus, 2006). Herein, the robustness of RI/E model principles (i.e., across time and different domains) clarifies the importance of providing comprehensible and appreciative performance feedback to students (cf. Marsh & Hau, 2004). Furthermore, the RI/E model framework helps us to understand why even students who have a lot of potential in several domains may develop a below-average academic self-concept in one particular domain (cf. Marsh & Hau, 2004) *and* may even become domain-specific underachievers (cf. Snyder & Linnenbrink-Garcia, 2013). Herein, we shed light on the need to simultaneously consider both students' achievements and academic self-concepts in *different* domains. Maybe more inclusive instructional strategies can help to substantially reduce the negative contrast effects that result in the disadvantage of particular academic domains in the long term, for example, combining elements from different academic domains in instruction (e.g., occasionally teaching mathematics in foreign languages or broaching the issue of radiocarbon age determination in history classes). Such inclusive strategies would emphasize the commonalities between academic fields instead of contrasting supposedly independent academic field skills. Of course, this assumption is speculative in nature and needs to be empirically tested. By presenting our findings, we hope to stimulate further research in this area.

References

Becker, M., Lüdtke, O., Trautwein, U., Köller, O., & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology, 104,* 682–699. doi:10.1037/a0027608

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

Bonett, D. G. (2012). Replication-extension studies. *Current Directions in Psychological Science, 21,* 409–412. doi:10.1177/0963721412459512

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology, 102,* 964–981. doi:10.1037/a0019644

Byrne, B. M. (1996). Academic self-concept: Its structure, measurement, and relation to academic achievement. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 287–316). Oxford, England: Wiley.

Byrne, B. M. (2002). Validating the measurement and structure of self-concept: Snapshots of past, present, and future research. *American Psychologist, 57,* 897–909. doi:10.1037/0003-066X.57.11.897

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

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56,* 453–484. doi:10.1146/annurev.psych.55.090902.141913

Chiu, M.-S. (2012). The internal/external frame of reference model, big-fish-little-pond effect, and combined model for mathematics and science. *Journal of Educational Psychology, 104,* 87–107. doi:10.1037/a0025734

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6,* 330–351. doi:10.1037/1082-989X.6.4.330

Cronbach, L. J. (1976). *Research on classrooms and schools: Formulation of questions, designs, and analysis*. Stanford, CA: Stanford Evaluation Consortium, Stanford University.

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/German Journal of Educational Psychology, 19,* 219–224. doi:10.1024/1010-0652.19.4.219

Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly, 18,* 585–609. doi:10.1111/j.1471-6402.1994.tb01049.x

Goetz, T., Preckel, F., Zeidner, M., & Schleyer, E. (2008). Big fish in big ponds: A multilevel analysis of test anxiety and achievement in special gifted classes. *Anxiety, Stress & Coping, 21,* 185–198. doi:10.1080/10615800701628827

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549–576. doi:10.1146/annurev.psych.58.110405.085530

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

Helmke, A. (1992). *Selbstvertrauen und schulische Leistungen* [Self-confidence and academic attainment]. Göttingen, Germany: Hogrefe.

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

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

Ireson, J., & Hallam, S. (2009). Academic self-concepts in adolescence: Relations with achievement and ability grouping in schools. *Learning and Instruction, 19,* 201–213. doi:10.1016/j.learninstruc.2008.04.001

Kaplan, D. (2009). *Structural equation modeling: Foundations and extensions* (2nd ed.). Los Angeles, CA: Sage.

Kenny, D. A. (1975). Cross-lagged panel correlation: A test for spuriousness. *Psychological Bulletin, 82,* 887–903. doi:10.1037/0033-2909.82.6.887

Köller, O., Klemmert, H., Möller, J., & Baumert, J. (1999). Eine längsschnittliche Überprüfung des Modells des Internal/External Frame of Reference [Achievement feedback and academic self-concepts: A longitudinal test of the internal/external frame of reference model]. *Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology, 13,* 128–134. doi:10.1024//1010-0652.13.3.128

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75,* 63–82. doi:10.3102/00346543075001063

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., . . . Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente* [PISA 2000: Documentation of assessment tools]. Berlin, Germany: Max-Planck-Institut für Bildungsforschung.

Legewie, J. (2012). Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der

Schule [The estimation of causal effects: An introduction to methods of causal inference based on peer effects in education]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 64,* 123–153. doi:10.1007/s11577-012-0158-5

Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association, 90,* 1112–1121. doi:10.1080/01621459.1995.10476615

Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development, 31,* 357–365. doi:10.1177/0165025407077757

Lubinski, D., & Benbow, C. P. (2000). States of excellence. *American Psychologist, 55,* 137–150. doi:10.1037/0003-066X.55.1.137

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science, 7,* 537–542. doi:10.1177/1745691612460688

Marsh, H. W. (1990). *Self-Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept—An interim test manual and a research monograph*. San Antonio, TX: Psychological Corporation.

Marsh, H. W. (1991). Failure of high-ability high schools to deliver academic benefits commensurate with their students' ability levels. *American Educational Research Journal, 28,* 445–480. doi:10.3102/00028312028002445

Marsh, H. W. (2006). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester, England: British Psychological Society.

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

Marsh, H. W., & Hau, K.-T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology, 96,* 56–67. doi:10.1037/0022-0663.96.1.56

Marsh, H. W., & Köller, O. (2004). Unification of theoretical models of academic self-concept/achievement relations: Reunification of East and West German school systems after the fall of the Berlin Wall. *Contemporary Educational Psychology, 29,* 264–282. doi:10.1016/S0361-476X(03)00034-1

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., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K.-T., O'Mara, A. J., & Craven, R. G. (2008). The big-fish–little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review, 20,* 319–350. doi:10.1007/s10648-008-9075-6

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

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of Personality, 74,* 403–456. doi:10.1111/j.1467-6494.2005.00380.x

Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: Relations to academic self-concept and achievement. *American Educational Research Journal, 34,* 691–720. doi:10.3102/00028312034004691

Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal, 35,* 705–738. doi:10.3102/00028312035004705

McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research, 29,* 409–454. doi:10.1207/s15327906mbr2904_5

McArdle, J. J., & Grimm, K. J. (2010). Five steps in latent curve and latent change score modeling with longitudinal data. In K. van Montfort, A. Satorra, & J. H. L. Oud (Eds.), *Longitudinal research with latent variables* (pp. 245–273). doi:10.1007/978-3-642-11760-2_8

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

Meredith, W., & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 203–240). doi:10.1037/10409-007

Möller, J., & Köller, O. (2001a). Dimensional comparisons: An experimental approach to the internal/external frame of reference model. *Journal of Educational Psychology, 93,* 826–835. doi:10.1037/0022-0663.93.4.826

Möller, J., & Köller, O. (2001b). Frame of reference effects following the announcement of exam results. *Contemporary Educational Psychology, 26,* 277–287. doi:10.1006/ceps.2000.1055

Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review, 120,* 544–560. doi:10.1037/a0032459

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

Möller, J., Retelsdorf, J., Köller, O., & Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: An integration of models of relations between academic achievement and self-concept. *American Educational Research Journal, 48,* 1315–1346. doi:10.3102/0002831211419649

Möller, J., Streblow, L., Pohlmann, B., & Köller, O. (2006). An extension to the internal/external frame of reference model to two verbal and numerical domains. *European Journal of Psychology of Education, 21,* 467–487. doi:10.1007/BF03173515

Muthén, L. K., & Muthén, B. O. (1998–2012a). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (1998–2012b). *Mplus Version 7* [statistical software]. Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2006). *Mplus: Statistical analysis with latent variables*. Los Angeles, CA: Muthén & Muthén.

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41,* 181–206. doi:10.1207/s15326985ep4103_4

Organisation for Economic Cooperation and Development. (2001). *Knowledge and skills for life: First results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris, France: Author.

Oud, J. H. L., & Delsing, M. J. M. H. (2010). Continuous time modeling of panel data by means of SEM. In K. van Montfort, A. Satorra, & J. H. L. Oud (Eds.), *Longitudinal research with latent variables* (pp. 201–244). doi:10.1007/978-3-642-11760-2_7

Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science, 7,* 531–536. doi:10.1177/1745691612463401

Pohlmann, B., & Möller, J. (2009). On the benefit of dimensional comparisons. *Journal of Educational Psychology, 101,* 248–258. doi:10.1037/a0013151

Preckel, F., Götz, T., & Frenzel, A. (2010). Ability grouping of gifted students: Effects on academic self-concept and boredom. *British Journal of Educational Psychology, 80,* 451–472. doi:10.1348/000709909X480716

Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology, 98,* 159–169. doi:10.1037/0022-0663.98.1.159

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

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Snyder, K. E., & Linnenbrink-Garcia, L. (2013). A developmental, person-centered approach to exploring multiple motivational pathways in gifted underachievement. *Educational Psychologist, 48,* 209–228. doi:10.1080/00461520.2013.835597

Trautwein, U., & Lüdtke, O. (2007). Students' self-reported effort and time on homework in six school subjects: Between-students differences and within-student variation. *Journal of Educational Psychology, 99,* 432–444. doi:10.1037/0022-0663.99.2.432

Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept. *Journal of Personality and Social Psychology, 90,* 334–349. doi:10.1037/0022-3514.90.2.334

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology, 98,* 788–806. doi:10.1037/0022-0663.98.4.788

Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: Conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology, 97,* 1115–1128. doi:10.1037/a0017048

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

Wagner, J. W. L. (2001). *Leistungsvergleiche in der Schule. Bezugsgruppeneffekte und Fähigkeitskonzepte* [Comparisons of achievement in school: Reference-group effects and academic self-concepts]. Landau, Germany: Verlag Empirische Pädagogik.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). doi:10.1037/10222-009

Xu, M. K., Marsh, H. W., Hau, K.-T., Ho, I. T., Morin, A. J. S., & Abduljabbar, A. S. (2013). The internal/external frame of reference of academic self-concept: Extension to a foreign language and the role of language of instruction. *Journal of Educational Psychology, 105,* 489–503. doi:10.1037/a0031333

Received February 28, 2013
Revision received February 9, 2014
Accepted February 11, 2014 ■