Arens, A. Katrin; Preckel, Franzis

Testing the internal/external frame of reference model with elementary school children. Extension to physical ability and intrinsic value

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
Contemporary educational psychology (2018) 54, S. 199-211

DOI: 10.1016/j.cedpsych.2018.06.003 - *http://dx.doi.org/10.1016/j.cedpsych.2018.06.003*

Testing the internal/external frame of reference model with elementary school children: Extension to physical ability and intrinsic value

A. Katrin Arens
German Institute for International Educational Research, Germany
Franzis Preckel
University of Trier, Germany

Author Note

Correspondence concerning this article should be addressed to A. Katrin Arens, Department of Education and Human Development and Center for Research on Individual Development and Adaptive Education of Children (IDeA), German Institute for International Educational Research, Schloßstraße 29, D-60486 Frankfurt am Main, Germany, Telephone: + 49 69 24708 138, Email: arens@dipf.de.

Acknowledgement

This study is based on data from the THINK Project by Franzis Preckel and Tanja Gabriele Baudson, a cross-sectional study on intelligence and cognitive motivation in German elementary schools. The project was funded by the Hogrefe Publishing Company, Göttingen. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript. Preparation of this article was supported by a Heisenberg fellowship grant from the German Research Foundation to A.K. Arens (AR 877/3-1).

This is the prepublication version of the following manuscript:

Arens, A.K., & Preckel, F. (2018). Testing the internal/external frame of reference model with elementary school children: Extension to physical ability and intrinsic value. *Contemporary Educational Psychology, 54*, 199–211. https://doi.org/10.1016/j.cedpsych.2018.06.003

This paper is not the copy of record and may not exactly replicate the authoritative document published in *Contemporary Educational Psychology*.

Abstract

Using a sample of third grade elementary school students (*N* = 514), we tested a generalized internal/external frame of reference (GI/E) model using math, verbal, and physical education achievements as predictors, and math, verbal, and physical ability self-concepts and intrinsic values as outcomes. Therefore, the present study adds to generalizability of the GI/E model across age and to the investigation of dimensional comparisons. The findings replicated the assumptions of the classic internal/external frame of reference model by showing positive within-domain relations between math (verbal) achievements and self-concepts, but negative cross-domain relations. Positive within-domain and negative cross-domain achievement relations were also found for math and verbal intrinsic values. The findings further showed positive relations between achievement and self-concept as well as between achievement and intrinsic value within the physical ability domain. In addition, the findings demonstrated a negative relation between math achievement and physical ability self-concept and between German achievement and physical ability intrinsic value. Hence, the findings indicated that dimensional comparisons operate across academic and non-academic domains. Boys and girls did not differ with regard to the relations among constructs as depicted in the GI/E model. However, differential mean levels of the constructs were found for boys and girls.

Keywords: academic self-concept; physical ability self-concept; I/E model; dimensional comparisons; elementary school; non-academic self-concept

Academic self-concepts comprise self-perceptions of one’s own academic ability in general and in different academic domains (e.g., math and verbal; e.g., Brunner et al., 2010). They play an important role in educational psychology because they influence students’ achievement, motivation, affect, and well-being from primary to tertiary education (Trautwein & Möller, 2016). Therefore, understanding the formation of academic self-concepts is relevant to students’ education at all levels. Understanding academic self-concept formation in young students seems to be especially relevant because it sets the ground for (early) interventions to support a positive self-concept and academic development (Craven & Marsh, 2008).

The classic internal/external frame of reference (I/E) model (Marsh, 1986; Möller, Pohlmann, Köller, & Marsh, 2009) serves to explain the formation of domain-specific (math and verbal) academic self-concepts by social and dimensional achievement comparisons. The I/E model has entailed a body of research which can be subdivided into two main categories: (1) Examining the generalizability of the I/E model across student characteristics including culture, gender, and age (e.g., Marsh, Abduljabbar et al., 2015; Marsh & Hau, 2004; Tay, Licht, & Tate, 1995), and (2) research focusing on the phenomenon of dimensional comparisons (Möller & Marsh, 2013).

In this study, we combined both strands of research. We tested the I/E model with third grade elementary school students in order to test the generalizability of the I/E model across age. Beyond the classic I/E model, we investigate a generalized internal/external frame of reference (GI/E) model (Möller, Müller-Kalthoff, Helm, Nagy, & Marsh, 2015). The GI/E model serves to investigate the phenomenon of dimensional comparisons as it includes more and other predictor variables (i.e., variables which are subject to dimensional comparisons) and outcome variables (i.e., variables which are affected by dimensional comparisons) from a wide variety of domains. The GI/E model tested in our study included math, verbal, and physical education achievements as predictors and math, verbal, and physical ability self-concepts and values as outcomes. Finally, we examined the generalizability of the model across student gender and tested gender differences in the mean levels of the constructs considered.

The I/E Model

Academic self-concept is highly domain-specific in nature since math and verbal self-concepts are unrelated to each other (Marsh, 1986). Given that math and verbal achievements are, however, substantially related to each other, the clear domain specificity of academic self-concept was first of all surprising. The I/E model explains this finding by the interplay of social and dimensional comparisons (Marsh, 1986; Möller et al., 2009). In social comparisons, students compare their own achievement in one domain with their classmates' achievement in the same domain. In dimensional comparisons, students contrast their own achievement in one domain with their own achievement in another domain. Given that math and verbal achievements are substantially correlated, social comparisons lead to a positive relation between math and verbal self-concepts. Dimensional comparisons lead to a negative correlation between math and verbal self-concepts since the perception of higher math achievement compared to one's own verbal achievement strengthens one's math self-concept and weakens one's verbal self-concept (and vice versa). Thus, the positive relation between math and verbal self-concepts originating from social comparisons and the negative relation due to dimensional comparisons balance out each other, leading to the observed low correlation between math and verbal self-concepts. In regression models, social comparisons result in positive paths between achievement and academic self-concepts within matching domains (e.g., math achievement and math self-concept). Dimensional comparisons, however, result in negative cross-paths between achievement in one domain and academic self-concept in the other domain (e.g., math achievement and verbal self-concept; see Figure 1).

Möller et al. (2009) presented a meta-analysis with strong support for the validity of the I/E model assumptions. Across 69 independent studies, a positive correlation (r = .67) was shown between math and verbal achievements but only a small relation was found (r = .10) between math and verbal self-concepts. In addition, math (verbal) achievement was positively related to math (verbal) self-concept (math: β = .61; verbal: β = .49), but math achievement was negatively related to verbal self-concept (β = -.21), and verbal achievement was negatively related to math self-concept (β = .27). The achievement–self-concept relations within and across domains were higher when using school grades instead of test scores as achievement indicators. Relative to test scores, school grades are salient and directly communicated to the students. They are easy to understand and compare among classmates and entail important implications for students' school careers, leading to the relatively higher relation to achievement compared to test scores (Helmke & van Aken, 1995; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005).

The I/E Model with Elementary School Students

The question whether the I/E model applies to students of all ages has been a matter of debate in research on the generalizability of the I/E model across student characteristics. Hence, it is questioned whether students from all age groups use social and dimensional comparisons for self-concept formation.

Some developmental considerations argue against the validity of the I/E model for younger students including elementary school children. The use of both social and dimensional comparisons requires abstract thinking abilities which might not develop until late childhood and adolescence (Harter, 1999). Social comparisons particularly rely on the ability to observe the achievement of significant others, to interpret others' and one's own achievement in relation to a common metric, to put others' achievement into relation to one's own achievement, and to finally integrate the resulting information into one's own academic self-concept (Nicholls, 1979; Wigfield & Eccles, 2002). Young children might not possess these abilities to a sufficient degree. Accordingly, findings from some studies indicated that young children do not use social comparisons for self-evaluation (e.g., Harter, 1999; Ruble, Boggiano, Feldman, & Loebl, 1980). This might explain why elementary school children are often characterized by an unrealistically high academic self-concept that is not closely related to objective achievement indicators (Chapman & Tunmer, 1995; Marsh, Ellis, & Craven, 2002; Stipek & McIver, 1989). However, other studies indicated that even young children make use of social comparisons for self-concept formation (see for example Butler, 1998). Accordingly, self-concept and achievement were found to be positively related within domains (Arens et al., 2016). In addition, the big-fish-little-pond effect (BFLPE; Marsh et al., 2008) – a phenomenon that particularly reflects the operation of social comparisons – was also demonstrated with elementary school students (Köller, Zeinz, & Trautwein, 2008).

Regarding dimensional comparisons, students need to integrate opposing but coexisting concepts into their self-concepts. Hence, they first have to realize and then to coordinate inconsistencies or even contradictions of elements of their selves across roles, situations, time, or domains (e.g., that one can be friendly to one's peers, but rough to one's parents). This ability is first applied to juxtapositions across separate domains (e.g., across academic and non-academic domains), before later on, students learn to handle contradictory self-evaluations within the same domain [e.g., across the math and verbal domains as two academic domains (Harter, 1988)]. Accordingly, young students' self-concept has been found to be little differentiated across domains, while with increasing age students separate self-concept facets more strongly (Marsh, 1989; Marsh & Ayotte, 2003; Schmidt et al., 2017).

In sum, developmental considerations imply that due to premature cognitive abilities, young students might have difficulties regarding the adequate use of social and particularly dimensional comparisons for academic self-concept formation. Accordingly, the classic I/E model might not apply to young students. Still, respective empirical evidence is scarce. Moderator analyses in the meta-analysis of Möller et al. (2009) demonstrated the generalizability of the I/E model across students of different ages (operationalized by years in school). However, as this meta-analysis incorporated only three (out of 69) studies exclusively considering students below fifth grade, it was not feasible to actually test whether the I/E model also applies to students below fifth grade.

Findings from single studies testing the I/E model with students below fifth grade provided consistent evidence that the I/E model applies to fourth grade students (Pinxten et al. 2015; Schmidt et al., 2017; also see Marsh, Abduljabbar et al., 2015). Yet, support for the I/E model is less clear for children of and below grade level 3. We identified four studies testing the I/E model with children of and below grade level 3, that is, the studies by Ehm, Lindberg, and Hasselhorn (2014), Schmidt et al. (2017), Lohbeck and Möller (2017), and Möller, Kuska and Zaunbauer (2011).

Ehm et al. (2014) tested an I/E model in which verbal achievement and verbal self-concept were represented by separate measures for reading and writing. Math, reading, and writing achievements were assessed by standardized achievement tests. The sample consisted of students attending grade levels 1, 2, and 3. Math, reading, and writing achievements were positively related to the self-concepts of the corresponding domains indicating social comparisons across all grade levels. Regarding the cross-domain relations indicating dimensional comparisons, however, the negative relation as expected according to the I/E model was only found between math and reading achievements and self-concepts in grade 3. Hence, the findings partially supported the applicability of the I/E model to third-grade students, but disconfirmed it for students attending grade levels 1 and 2 and for the relations among math and writing achievements and self-concepts. Yet, in this study, test scores were used as achievement indicators which have been found to share relatively lower relations with self-concept than school grades (Marsh et al., 2005); the effects of dimensional comparisons might have thus been underestimated in this study.

In a recent study, Schmidt et al. (2017) also considered both reading and writing self-concepts as indicators for verbal self-concept and examined the I/E model with third and fourth grade elementary school students. The findings supported the existence of social comparisons given substantial relations between achievements and self-concepts of matching domains (i.e., math achievement and math self-concept, verbal achievement and reading self-concept, verbal achievement and writing self-concept). While evidence of dimensional comparisons was consistently demonstrated by negative cross-domain achievement–self-concept relations in fourth grade students, corresponding results were mixed and varied contingent upon the assumed underlying structure of academic self-concept in third grade students. In essence, negative paths leading from math achievement to reading and writing self-concepts were only found in third grade students when assuming the nested Marsh/Shavelson model (Brunner et al., 2010; a bifactor-like model with a general academic self-concept at the apex and math, writing, and reading self-concepts as specific factors nested under general academic self-concept) and not when assuming a first-order factor model with math, writing, reading, and general academic self-concepts as correlated factors. The findings of this study thus provided partial support for the validity of the I/E model with third-grade students. However, verbal achievement was measured by students' grades in German reflecting students' overall verbal achievement, while verbal self-concept was measured separately with respect to reading and writing. This divergent level of specificity in the operationalization of verbal self-concept and verbal achievement might have undermined the (negative) relations between achievement and self-concept. Relatively higher achievement–self-concept relations are usually found when both achievement and self-concept measures relate to the same domain and are located on the same level of specificity (Arens, Yeung, & Hasselhorn, 2014; Swann, Chang-Schneider, & Larsen McClarty, 2007).

In a study with third grade German elementary school students, Möller et al. (2011) used school grades and self-concepts in math and German to examine the I/E model. The findings supported social comparisons due to positive within-domain achievement–self-concept relations. With regard to the cross-domain relations, only the path leading from German achievement to math self-concept was significantly negative, but not the path leading from math achievement to German self-concept. Again, the match between achievement and self-concept measures was not optimal regarding the verbal domain. While German grades were used as an achievement indicator targeting the verbal domain in general, some items of the German self-concept scale targeted reading only while others addressed both reading and writing. Moreover, the analyses were not based on latent variables and did not correct for measurement error.

In another recent study examining second grade students and considering math and reading math self-concept and achievement measures, Lohbeck and Möller (2017) found significantly positive achievement–self-concept relations within the domains of math and reading, supporting the application of social comparisons. However, the two achievement–self-concept relations across the domains of math and reading were non-significant arguing against the validity of the I/E model in this sample. Yet, test scores were again used as achievement indicators which might have underestimated negative cross-domain self-concept–achievement relations indicating dimensional comparison effects.

Hence, in sum, empirical evidence for the I/E model has been sparse for students attending grade level 3 or lower. While all studies consistently provided empirical support for social comparison processes in the formation of math and verbal self-concepts, empirical evidence for the effects of dimensional comparison processes was ambiguous and inconsistent.

In the present study, we aimed to add to existing research by examining a sample of third grade German students. Limitations of previous studies (Ehm et al., 2014; Lohbeck & Möller, 2017; Möller et al., 2011; Schmidt et al., 2017) were addressed by including measures for verbal achievement and self-concept which are both related to the verbal domain in general and thus of the same level of specificity (i.e., students' school grades in German as a school subject and German self-concept). School grades were considered as achievement indicators. Finally, the within-domain and cross-domain achievement–self-concept relations were analyzed in a latent regression model, controlling for measurement error. These aspects may help detect presumable negative achievement–self-concept relations across the math and verbal domains indicating dimensional comparisons and to gain more insight into the validity of the I/E model for young students.

The GI/E Model

The I/E model has invoked interest in the phenomenon of dimensional comparisons leading to dimensional comparison theory (DCT; Möller & Marsh, 2013). DCT targets the antecedents, circumstances, and consequences of dimensional comparisons. While the classic I/E model focuses on the relation between math and verbal achievements and math and verbal self-concepts, DCT offers a broader definition of dimensional comparisons according to which an individual compares his/her perceptions of aspects of a particular domain A with his/her perceptions of aspects of a particular domain B, and this comparison is assumed to influence any kind of outcomes related to these domains. These assumptions have led to the GI/E model (Möller et al., 2015) which offers an empirical framework for studying dimensional comparisons in a broader way based on DCT. First, the GI/E model allows a broader range of predictor variables beyond math and verbal achievements to be subject to dimensional comparisons, that is, to be compared across domains. Second, the GI/E model allows a broader range of outcome variables beyond math and verbal self-concepts which are affected by dimensional comparisons. In the present study, we tested a GI/E model extending the classic I/E model to the physical ability domain and to intrinsic value.

Extension to the Physical Ability Domain

Studies on the GI/E model predominantly considered students' achievements and self-concepts related to multiple school subjects. They showed that students compare their own achievements, and run dimensional comparisons, across a broad range of school subjects and thus do not only contrast their math and verbal achievements (Arens, Möller, & Watermann, 2016; Jansen, Schroeders, Lüdtke, & Marsh, 2015; Marsh et al., 2014; Marsh, Lüdtke et al., 2015; Möller, Streblow, Pohlmann, & Köller, 2006; Niepel, Brunner, & Preckel, 2014). When including multiple school subjects, cross-paths between achievement and self-concept measures of non-matching domains (depicting dimensional comparisons) can be negative as well as positive. Hence, dimensional comparisons can lead to contrast effects (negative cross-paths) or to assimilation effects (positive cross-paths). Contrast effects are presumed to occur between domains which are perceived as dissimilar (e.g., math and verbal) while assimilation effects are presumed to occur between domains which are perceived as similar (e.g., math and physics) (Helm, Müller-Kalthoff, Nagy, & Möller, 2016).

We expanded the classic I/E model by including students' achievement and self-concept related to physical ability. Although physical education is taught as a school subject and students receive school grades, physical education can be considered to be a non-academic domain. Physical activity and sports are still primarily an extracurricular activity and students might perceive feedback on their physical ability from a variety of sources (peers, parents, and coaches of extracurricular activities) so that students' feedback is not restricted to the school context. Hence, we expanded the classic I/E model to a non-academic domain. In fact, we probed a GI/E model in which we examined the within-domain and cross-domain relations among achievement and self-concept related to math, German, and physical ability.

Findings from previous studies demonstrated that students' achievement and self-concept addressing the physical ability domain are related to each other implicating the use of social comparisons (e.g., Marsh, Gerlach, Trautwein, Lüdtke, & Brettschneider, 2007). In addition, the BFLPE was found for the physical ability domain, providing even stronger evidence for the operation of social comparisons (Chanal, Marsh, Sarrazin, & Bois, 2005; Trautwein, Gerlach, & Lüdtke, 2008). Beyond evidence of social comparisons, so far only the findings reported by Chanal, Sarrazin, Guay, and Boiché (2009) provided evidence for the role of dimensional comparisons in the formation of physical ability self-concept. In this study, students' achievements and self-concepts related to math, French (verbal), and physical ability were integrated in a GI/E model with a sample of French high school students. The findings revealed positive relations between physical education achievement and physical ability self-concept indicating social comparisons. Verbal achievement was negatively related to physical ability self-concept, and physical education achievement was positively related to math self-concept. These findings indicated dimensional comparisons including the physical ability domain: Students seem to compare their achievement in physical education with their math and verbal achievements and these comparisons seem to bear an impact on students' self-concepts in these domains. Yet, from this study, it did not become clear whether dimensional comparisons among math, verbal, and physical education achievements resulted in a contrast or an assimilation effect, since both a positive (assimilation effect) and a negative (contrast effect) cross-domain achievement–self-concept relation were found. Given evidence of positive relations between physical ability or fitness and academic achievement (Castelli, Hillmann, Buck, & Erwin, 2007; Sibley & Etnier, 2003), one can expect positive cross-domain relations, and thus an assimilation effect, between physical education achievement and physical ability self-concept on the one hand and academic (i.e., math, verbal) achievements and self-concepts on the other hand. Alternatively, students might strictly differentiate between academic and non-academic domains, that is, between math and verbal domains on the one hand and physical ability on the other hand. This separation and perceived dissimilarity between academic and non-academic domains might induce a contrast effect leading to negative relations between math and verbal achievements and self-concepts on the one hand and physical education achievement and physical ability self-concept on the other hand. In the present study, we assessed students' math, verbal, and physical education achievements and math, verbal, and physical ability self-concepts with third-grade elementary school students. We were thus able to replicate the findings from Chanal et al. (2009) with younger students and to further explore dimensional comparisons involving the physical ability domain.

Extension to Intrinsic Value

According to the GI/E model, dimensional comparisons may also influence other motivational constructs beyond students' self-concept. In this regard, expectancy-value theory (EVT; Eccles & Wigfield, 1995; Nagengast et al., 2011; Wigfield & Eccles, 2000) can be utilized as a theoretical framework and combined with the GI/E model. EVT assumes that students' motivation encompasses an expectancy component and a value component which jointly impact upon students' aspirations, achievement, effort, and enrolment decisions (Durik, Vida, & Eccles, 2006; Guo, Marsh, Parker, Morin, & Dicke, 2017; Simpkins, Davis-Kean, & Eccles, 2006; Watt et al., 2012). Originally, the expectancy component was presumed to encompass ability beliefs (i.e., individuals' self-evaluations of their competences) and expectancies of success, but both constructs were found to be inseparable and to form a single factor (Eccles, Wigfield, Harold, & Blumenfeld, 1993; Eccles &

Wigfield, 2002). The value component comprises four different facets: intrinsic value (enjoyment, interest), attainment value (subjective importance), utility value (subjective usefulness), and cost (negative consequences of choosing and engaging in a task, and opportunity cost) (Wigfield & Eccles, 2000). Both the expectancy and value components have been found to be domain-specific in nature. Hence, students form separate expectancy and value perceptions in different subject domains (Eccles et al., 1993; Trautwein et al., 2012).

The expectancy component has recently been operationalized as students' academic self-concept (Guo, Marsh, Morin, Parker, & Kaur, 2015; Guo, Parker, Marsh, & Morin, 2015; Nagengast et al., 2011; Trautwein et al., 2012). Given the rich body of empirical evidence for the validity of the I/E model for students' academic self-concept (for meta-analytic findings, see Möller et al., 2009), the I/E model seems well established for the expectancy component. Findings from previous research demonstrated that the GI/E model-like pattern [i.e., positive (negative) achievement-outcome relations within (across) the math and verbal domains] also applies to relations between domain-specific achievements and EVT value constructs (Gaspard et al., 2018; see also Guo et al., 2017). The findings of numerous studies indicated a GI/E model-like pattern when using outcomes resembling the construct of intrinsic value, such as interest (Schurtz, Pfost, Nagengast, & Artelt, 2014), intrinsic motivation (Marsh, Abduljabbar et al., 2015), or enjoyment (Goetz, Frenzel, Hall, & Pekrun, 2008). Yet, these findings need to be replicated with elementary school students. We did this in the present study.

Generalizability across Gender

Boys and girls were found to differ in their mean levels of self-concept and these mean level differences correspond to gender stereotypes (Fredricks & Eccles, 2002; Marsh, 1989; Klomsten, Skaalvik, & Espnes, 2004; Skaalvik & Skaalvik, 2004). Accordingly, boys were found to report higher mean levels of math self-concept and physical ability self-concept, whereas girls were found to display higher levels of verbal self-concept. Parallel gender-stereotypic gender differences were also found for intrinsic value (Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Skaalvik & Skaalvik, 2004; Watt, 2004). Although most of the studies demonstrating gender differences in the mean levels of self-concepts and intrinsic values were based on samples of secondary school students, gender differences in self-concept facets seem to already be in place in elementary school (Fredricks & Eccles, 2002; Herbert & Stipek, 2005).

Despite these mean level differences, boys and girls may differ in their achievement–self-concept relations. Research on whether boys and girls differ in the size of the relations between achievement and self-concept can focus on within-domain relations (e.g., math achievement and math self-concept) and on cross-domain relations (e.g., math achievement and verbal self-concept). Studies considering gender invariance in both within-domain as well as cross-domain self-concept–achievement relations give insight into the generalizability of the I/E model across gender.

Evidence of gender invariance in self-concept–achievement relations within domains has been found for preschool students (Arens et al., 2016), elementary school students (Helmke & van Aken 1995), and secondary school students (Marsh et al., 2005; Marsh & Yeung, 1998; Valentine, DuBois, & Cooper, 2004). In studies focusing on both within-domain and cross-domain relations, the I/E model was found to be similarly applicable to boys and girls (Möller et al., 2009; Tay et al., 1995), but only secondary school students were considered in these studies. Hence, gender invariance of the I/E model has not yet been tested with elementary school students and when the I/E model is expanded to the domain of physical ability and intrinsic value. Yet, given corresponding findings with secondary school students, and since boys and girls are not found to differ in the development and level of cognitive abilities needed to conduct social and dimensional comparison processes (see for example, Ardila, Rosselli, Matute, & Inozemtseva, 2011; Weiss, Kemmlera, Deisenhammerb, Fleischhackera, & Delazer, 2011), gender invariance of the I/E and GI/E models might also be valid for elementary school students.

The Present Study

We aimed to contribute to and to extend contemporary research on the classic I/E model in various ways. By using a sample of third grade elementary school students, we tested the applicability of the I/E model to this age group. Rather than focusing on the classic I/E model only including math and verbal self-concept and achievement measures, we tested a GI/E model (Möller et al., 2015). In essence, we broadened the domains considered by integrating the physical ability domain. In addition, by considering both self-concept and intrinsic value in the three domains (math, German, physical ability), we extended the outcome variables considered. Hence, our study can offer insight into the use of social and dimensional comparisons with third grade elementary school students.

Finally, we examined gender differences in the mean levels of the constructs and tested for gender invariance with respect to the relations among all constructs considered. Our findings thus inform research and theory about whether gender invariance of the classic I/E model and of a GI/E model extended to the physical ability domain and intrinsic value can be postulated for elementary school students.

Method

Sample

The sample analyzed in this study [$N = 514$; boys: $N = 273$ (53.1%), girls: $N = 241$ (46.9%)] is a subsample of the norming sample of the intelligence test THINK 1-4 (Baudson, Wollschläger, & Preckel, 2016) collected in summer 2011. It comprises those students who additionally completed the German short version of the Self Description Questionnaire (SDQ I-GS; Arens, Yeung, Craven, & Hasselhorn, 2013, see below). The students attended the third grade of elementary schools, which were all located in the German federal state of Baden-Wuerttemberg. The mean age of the students was 8.90 years (*SD* = 0.51). The student sample analyzed here came from families in which *N* = 39 (7.6%) students had parents with a leaving certificate from the low-ability (vocational) track of German secondary schools, *N* = 151 (29.4%) students had parents with a certificate from the intermediate track, and *N* = 264 (51.4%) students had parents with a certificate from the high-ability (academic) track as parents' highest educational level. For *N* = 57 (11.1%) students, respective information was missing and *N* = 3 (0.6%) students reported that their parents had not graduated from school at all. *N* = 50 (9.7%) students reported that their mothers worked full time, *N* = 300 (58.4%) that their mothers had a part-time job, and *N* = 102 (19.8%) that their mothers were not employed. Respective information was missing for *N* = 62 (12.1%) students. *N* = 382 (74.3%) students had a father who worked full time, *N* = 19 (3.7%) students had a part-time working father, and *N* = 16 (3.1%) students had a father who was not employed. For *N* = 97 (18.9%) students, information on this measure for their father's occupational status was missing.

Parental consent for participation was obtained for all students. The students were informed about the purpose of the study and the confidential treatment of the data before completing the measures. Data were collected in class by trained research assistants.

Measures

Self-concepts and values. Students' self-concepts and intrinsic values related to math, German, and physical ability were measured by the SDQ I-GS (Arens et al., 2013; see Appendix). With respect to the domains of math, German, and physical ability, the SDQ I encompasses competence-related items asking for students' self-perceptions of competence and affect-related items asking for students' liking, interest, and enjoyment. The competence-related items of the SDQ I reflect students' domain-specific academic self-concepts defined as students' competence self-perceptions (Arens & Hasselhorn, 2015; Skaalvik & Rankin, 1995).

The affect-related items reflect intrinsic value as they address students' interest and enjoyment (Arens & Hasselhorn, 2015; Marsh et al., 2013). Supporting the distinctiveness of these two kinds of item sets, previous studies demonstrated that the competence-related items and the affect-related items define separate factors (Marsh, Craven, & Debus, 1999). They are differentially related to achievement (higher relations for the competence component; Arens, Bodkin-Andrews, Craven, & Yeung, 2014; Arens, Yeung, Craven, & Hasselhorn, 2011) and effort (higher relations for the affect component; Arens & Hasselhorn, 2015; Pinxten, Marsh, De Fraine, Van Den Noortgate, & van Damme, 2014).[1]

The SDQ I-GS scales for math self-concept, German self-concept, math intrinsic value, and German intrinsic value consist of three items. The physical ability self-concept scale comprises two items, while the physical ability intrinsic value scale consists of one single item. The items were worded as statements (e.g., "I am good at sports; I enjoy doing work in German.") and the students were asked to indicate whether these statements were *false*, *mostly false*, *sometimes false and sometimes true*, *mostly true*, or *true* on a 5-point Likert scale. Reliability estimates were sufficient for all multi-item scales: math self-concept: $\alpha = .872$; German self-concept: $\alpha = .865$; physical ability self-concept: $\alpha = .735$; German intrinsic value: $\alpha = .901$; math intrinsic value: $\alpha = .915$).

Achievement. Students' school grades in math, German, and physical education were taken from the latest school reports the students had received before the assessment of self-concept and intrinsic value. In Germany, the grading system ranges from 1 to 6 with 1 ("very good") displaying the highest and 6 ("failed") displaying the lowest grade. For easier interpretation, the grades were reversely coded prior to the analyses so that higher values represent higher school grades, that is, better achievement.

Statistical Analyses

All analyses were conducted with Mplus 7.8 (Muthén & Muthén, 1998-2017) using the maximum likelihood estimator with robust standard errors (MLR). Missing data were handled by the full information maximum likelihood (FIML) approach implemented in Mplus (Enders, 2010). The amount of missing data ranged from 0.2% to 1.8% for the SDQ I-GS items, and between 10.3% and 11.1% for the school grades.[2]

We first ran a confirmatory factor analysis (CFA) model in which we assumed separate factors for the SDQ I-GS scales (math self-concept, German self-concept, physical ability self-concept, math intrinsic value, German intrinsic value, physical ability intrinsic value) and three achievement factors (math achievement, German achievement, physical education achievement). The self-concept and intrinsic value factors (except physical ability intrinsic value) were defined by the corresponding SDQ I-GS items, the achievement factors were defined by students' school grades in math, German, and physical education as single-item indicators, and the physical ability intrinsic value factor was defined by the respective single SDQ I-GS item. This model served to examine the integrity of the SDQ I-GS to assess students' self-concepts and intrinsic values and to investigate the correlations between self-concepts, intrinsic values, and achievements. Based on this CFA model, we stated the GI/E model tested in the present study. To this aim, the six SDQ I-GS factors were simultaneously regressed on the three achievement factors (Figure 2).

In a subsequent set of models (Models I1 to I9 in Table 3), we tested gender invariance of the GI/E model. To this aim, we integrated gender as a grouping variable in the CFA model that assumes separate factors for the SDQ I-GS scales and achievement indicators (see above). We first stated a model of configural invariance in which the same number of factors defined by the same items was assumed for boys and girls but all model parameters were freely estimated across gender (Millsap, 2011; Model I1). Based on this model, step-by-step, we included invariant factor loadings (Model I2), item intercepts (Model I3), and item uniquenesses (Model I4). Factor loading invariance (also known as metric invariance) is most important as it implies that the same constructs with the same meanings are measured across groups. Invariant item intercepts are a prerequisite for testing latent mean invariance. Invariance testing of item uniquenesses is not required in latent analyses correcting for measurement error but is often included to complete the series of measurement invariance testing. The invariance testing procedure continued with testing the invariance of factor variances (Model I5). In the case of invariant factor variances, the invariance of factor correlations can be examined by restricting the factor covariances to invariance (Marsh, 1994; Model I6). The series of invariance tests was concluded by testing the invariance of factor means to explore mean level differences (Model I7).

Finally, we tested gender invariance with respect to the GI/E model. For this purpose, the within-domain and cross-domain relations between the achievement factors (predictors) and the various SDQ I-GS factors (i.e., self-concepts and intrinsic value factors as outcomes) were freely estimated across boys and girls in Model I8. This model was compared to Model I9 in which all relations were set to invariance across gender.

The evaluation of model fit was based on a wide range of descriptive goodness-of-fit indices (e.g., Marsh, Hau, & Wen, 2004). Accordingly, we report the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root mean square error of approximation (RMSEA). For the CFI and TLI, values above .90 and .95 represent an adequate respectively good model fit (Hu & Bentler, 1999). RMSEA values should be below .05 for a close fit, or between .05 and .08 for a reasonable fit (Browne & Cudeck, 1993). Changes in the descriptive goodness-of-fit indices between more and less restrictive models were observed to evaluate the invariance models. Invariance can be assumed as long as the CFI does not drop more than .010 (Cheung & Rensvold, 2002; see also Chen, 2007).

Results

A CFA model stating separate factors for the six SDQ I-GS scales (math self-concept, German self-concept, physical ability self-concept, math intrinsic value, German intrinsic value, and physical ability intrinsic value) and for math, verbal, and physical education achievements resulted in a good fit: $\chi^2(99) = 158.707$, CFI = .986, TLI = .978, RMSEA = .034. The good model fit as well as the resulting factor loadings (see Appendix) supported the integrity of the SDQ I-GS scales[3]. When inspecting the resulting factor correlations (Table 1), math achievement showed the relatively highest relation to math self-concept ($r = .589$, $p < .05$), German achievement had the relatively highest relation to German self-concept ($r = .437$, $p < .05$), and students' physical ability self-concept demonstrated the relatively highest relation to physical education achievement ($r = .449$, $p < .05$). These findings of substantial achievement–self-concept relations within matching domains supported the convergent validity of the SDQ I-GS self-concept scales. The correlations between self-concepts and intrinsic values were high within domains (math: $r = .769$; German: $r = .794$; physical ability: $r = .643$, for all $p < .05$), but they were far from perfect indicating the separation between self-concept and intrinsic value factors. The separation between self-concept and intrinsic value factors was further supported by the finding that within domains, the relations between self-concepts and achievements were higher than the relations between intrinsic values and achievement. For instance, the correlation between math achievement and math self-concept was $r = .589$ ($p < .05$), while the correlation between math achievement and math intrinsic value was $r = .309$ ($p < .05$).

The GI/E Model

Following the logic of the GI/E model, we stated a latent regression model estimating the within-domain and cross-domain relations between math, verbal, and physical education achievements on the one hand and the six SDQ I-GS factors (math self-concept, German self-concept, physical ability self-concept, math intrinsic value, German intrinsic value, physical ability intrinsic value) on the other hand (Figure 2). This model is statistically equivalent to the CFA model assuming separate achievement and SDQ I-GS factors so that both models resulted in the same fit. The resulting path coefficients (Table 2, see also Table S1 of the Online Supplements) replicated the classic I/E model since math achievement was positively related to math self-concept ($\beta = .769$, $p < .05$) and verbal achievement was positively related to German self-concept ($\beta = .615$, $p < .05$), but math achievement was negatively related to German self-concept ($\beta = -.289$, $p < .05$), and verbal achievement was negatively related to math self-concept ($\beta = -.258$, $p < .05$). The I/E model assumptions could also be supported for intrinsic value since math (verbal) achievement was positively related to math (verbal) intrinsic value ($\beta = .505$, resp. $\beta = .491$, for both $p < .05$), while math (verbal) achievement was negatively related to verbal (math) value ($\beta = -.417$, resp. $\beta = -.275$, for both $p < .05$)[4].

Physical education achievement and physical ability self-concept were significantly positively related to each other ($\beta = .508$, $p < .05$). In parallel, physical education achievement showed a significantly positive relation to physical ability intrinsic value ($\beta = .297$, $p < .05$). Physical education achievement was not significantly related to math and verbal self-concepts and intrinsic values. German achievement was not significantly related to physical ability self-concept, but it showed a significant negative relation to physical ability intrinsic value ($\beta = -.156$, $p < .05$). Math achievement displayed a significantly negative relation with physical ability self-concept ($\beta = -.123$, $p < .05$), but was unrelated to physical intrinsic value.

Gender Invariance

Model I1 (Table 3) is a model of configural gender invariance assuming separate achievement and SDQ I-GS factors which are freely estimated across boys and girls. Based on this model, we tested gender invariance by gradually adding invariant factor loadings (Model I2), invariant item intercepts (Model I3), invariant item uniquenesses (Model I4), invariant factor variances (Model I5), and invariant factor covariances (Model I6). Across these models, the CFI value did not decline by more than .01 between less and more restrictive consecutive models. The invariance of factor covariances (Model I6) in case of invariant factor variances (Model I5) indicated that boys and girls did not differ in their relations between the achievement and SDQ I-GS factors. Hence, the GI/E model stated here was found to be similarly valid for boys and girls (see also Models I8 and I9).

The fit declined more substantially (ΔCFI = -.018) when additionally assuming invariant factor means (Model I7). This result indicated that boys and girls displayed differential factor mean levels. For identification purposes, in Model I3 (i.e., the model with invariant factor loadings and item intercepts), we set the factor means to zero in one group (reference group, boys in our case) and freely estimated the factor means of the other group (girls in our case). The results depict the deviations in the mean levels between the reference and comparison groups in standard deviation units. Boys were found to have higher mean levels of math self-concept (-.414), math intrinsic value (-.208), physical ability self-concept (-.249), physical ability intrinsic value (-.284), and math achievement (-.272, for all $p <$. 05). Girls demonstrated higher mean levels of German self-concept (.481), German intrinsic value (.795), and German achievement (.267, for all $p < .05$). Girls and boys were not found to differ in their mean levels of physical education achievement (see also Table 4).

In a final step, we examined gender invariance of the GI/E model. Based on the assumption of invariant factor loadings which is a sufficient precondition for testing invariance of factor relations, Model I8 (Table 3) estimated the GI/E model freely across gender. The CFI value of this model remained robust when restricting the path coefficients to be invariant across gender (Model I9). Hence, for both boys and girls, the within-domain and cross-domain relations between math, verbal, and physical education achievements on the one hand and math self-concept, German self-concept, physical ability self-concept , math intrinsic value, German intrinsic value, and physical ability intrinsic value on the other hand were found to be of similar size for boys and girls.

Discussion

Validity of the I/E model with Third Grade Students

Developmental considerations might question the validity of the I/E model for young students as they might lack the ability to adequately use social and dimensional comparisons for self-concept formation (Harter, 1988, 1999; Ruble et al., 1980). Empirical studies testing the I/E model with elementary school students attending grade level 3 supported social comparisons but provided inconsistent findings with respect to dimensional comparisons (Ehm et al., 2014; Lohbeck & Möller, 2017; Möller et al., 2011; Schmidt et al., 2017). A closer look revealed that these previous studies might suffer from limitations that might have obscured negative cross-domain achievement–self-concept relations indicating dimensional comparisons. The present study aimed to shed further light on the validity of the I/E model for third grade elementary school students by overcoming the limitations of previous studies.

According to our findings, the classic I/E model also applies to third grade elementary school students. In fact, given significantly positive within-domain achievement–self-concept relations, third grade elementary school students seem to use social comparisons for self-concept formation. Further, the results demonstrated negative cross-domain achievement–self-concept relations indicating dimensional comparisons. Hence, third grade elementary school students seem to contrast their math and verbal achievements which affects their math and verbal self-concepts.

The GI/E Model

Another aim of this study was to examine extensions of the classic I/E model based on the GI/E model (Möller et al., 2015) and DCT (Möller & Marsh, 2013). In this context, we extended the I/E model to the physical ability domain and to intrinsic value as a further outcome variable. The findings revealed positive relations between achievement and intrinsic value within all three domains (math, German, physical ability) supporting the application of social comparisons in the formation of domain-specific intrinsic value. The findings further showed negative relations between achievement and intrinsic value across the math and verbal domains. These findings supported the involvement of dimensional comparisons in the formation of intrinsic values. Thereby, we replicated findings from previous studies showing GI/E model-like patterns for constructs conceptually similar to intrinsic value (e.g., interest or enjoyment; Gaspard et al., 2018; Goetz et al., 2008; Guo et al., 2017; Marsh, Abduljabbar et al., 2015; Schurtz et al., 2014), and expanded the respective findings to third grade elementary school students. Moreover, math achievement was negatively related to physical ability self-concept, and German achievement was significantly negatively related to physical ability intrinsic value. That is, students seem to compare their achievements across academic (math, verbal) and non-academic (physical education) domains. This comparison is negatively related to students' physical ability motivation both in terms of students' self-concept and intrinsic value. This finding hints at a partial inconsistency between academic achievement and non-academic (i.e., physical ability) motivation. Higher math achievement is associated with a lower physical ability self-concept, and higher German achievement is related to lower physical ability intrinsic value. Math achievement might be associated with clearer information on one's own abilities than verbal achievement which is less clearly defined. Hence, comparisons involving math achievement might be more relevant for students' self-concept. Still, further research is necessary to explicate this finding.

Developmental theory had originally assumed that students develop the ability to compare elements of their selves between domains earlier (in middle childhood) than within domains (in late childhood and early adolescence; Harter, 1999). In this case, dimensional comparisons might have only been expected between the academic (math, verbal) and non-academic (physical) domain when examining third-grade students. Yet, since the findings hinted at dimensional comparison processes even within the academic domain, the ability to coordinate and compare self-evaluations within a domain might develop earlier. This development might also be enhanced by specific environmental features such as grading practices. In other words, the use of school grades might have stimulated the use of social and dimensional comparisons as school grades are salient, directly communicated to the students, easy to understand, and thus easy to compare. In addition, with students' advancing grade level, the school environment might urge students to focus on social and dimensional comparisons. In fact, environmental characteristics such as salient and frequent performance feedback, evaluation pressure, and the introduction of achievement standards are assumed to trigger the use of social and dimensional comparisons in students (Ames, 1992). Future research is necessary to detect environmental characteristics that augment or weaken students' use of dimensional comparisons.

Gender Invariance and Gender Differences

Boys and girls were comparable regarding the relations among constructs, and thus regarding the GI/E model. The generalizability of the I/E model across gender that was previously demonstrated for secondary school students (Möller et al., 2009; Tay et al., 1995) thus also pertains to (advanced) elementary school students. Gender invariance also applies to our GI/E model extended to intrinsic value and to achievement and self-concept related to the physical ability domain.

Gender differences, however, were found for mean levels. The results match findings from previous studies (Fredricks & Eccles, 2002; Marsh, 1989; Klomsten et al., 2004; Skaalvik & Skaalvik, 2004). Boys reported higher mean levels of math self-concept, physical ability self-concept, and math intrinsic value, and physical ability intrinsic value. Girls displayed higher levels of German self-concept, and German intrinsic value. Hence, the self-concept mean level differences between boys and girls correspond to gender stereotypes indicating that third grade elementary school students have already internalized gender stereotypes in their self-perceptions (Steffens, Jelenec, & Noack, 2010). Finally, boys had better math grades, while girls displayed better grades in German replicating previously found gender differences regarding achievement (De Fraine, van Damme, & Onghena, 2007; Nowell, & Hedges, 1998; Skaalvik & Rankin, 1994).

Practical Implications

In practical terms, the findings of our study indicate that it is important to consider the operation of dimensional comparisons in the formation of students' multidimensional (i.e., academic and non-academic) self-concepts and intrinsic values. Hence, enhancement interventions should not only address students' self-concept and intrinsic values in a single domain. Programs designed to foster students' self-concepts and values (e.g., Gaspard, Dicke, Flunger, Brisson et al., 2015; Hulleman, Godes, Hendricks, & Harackiewicz, 2010; O'Mara, Marsh, Craven, & Debus, 2006) should rather consider the complex interplay of achievements, self-concepts, and intrinsic values within and across various academic and non-academic domains.

Limitations and Future Directions

Our study remains cross-sectional. Longitudinal studies would be required to investigate the within-domain and cross-domain relations among achievements, self-concepts, and intrinsic values across time (see for example Möller, Retelsdorf, Köller, & Marsh, 2011; Niepel et al., 2014). Longitudinal studies would also allow inspecting for mediator variables. The within-domain and cross-domain relations between achievement and value constructs might be mediated through students' academic self-concept (Goetz et al., 2008; Guo et al. 2017; Schurtz et al., 2014). Indeed, social and dimensional comparisons as stated in the I/E model pertain to the comparison of achievements. Thus, they primarily provide students with evaluations of students' competences in specific domains (i.e., domain-specific academic self-concepts) which in turn might affect students' interest, liking, and enjoyment with respect to this subject domain (Harter, 1999; see also Marsh et al., 2005). This mediation assumption needs testing with elementary school students. Indeed, further research on the I/E and GI/E models might benefit from disentangling which variables are directly subject to social and dimensional comparisons and which are only affected indirectly.

We included students' self-concepts and intrinsic values related to three domains (math, German, and physical ability) as outcome variables in our extended I/E model. Thereby, we went beyond the scope of common studies on the I/E model. Nevertheless, as we only considered intrinsic value as one value facet as stated by EVT, further studies should consider the other EVT value facets (i.e., attainment value, utility value, and cost) which have been found to form theoretically and empirically separable constructs (Conley, 2012; Gaspard, Dicke, Flunger, Schreier et al., 2015; Trautwein et al., 2012). A single item was used to measure intrinsic value related to physical ability. Future studies should consider respective scales, that is, multiple items for physical ability intrinsic value to ensure the psychometric property of this measurement. Finally, beyond the math, German, and physical ability domains, future studies should consider a broader range of academic and non-academic domains in GI/E models. Respective studies would allow insight into whether further non-academic domains are involved in dimensional comparisons and are contrasted to academic domains.

The number of participating classes was too small to correct for possibly biased standard errors due to the hierarchical nature of the data. However, students' nesting within classes might only marginally affect students' self-concepts and values as these constructs are located on the individual student rather than on the class level (Marsh et al., 2012). In addition, we considered only third grade elementary school students in our study. While we thereby examined a student sample which has rarely been focused in research on the I/E model and for which existing studies provided inconsistent findings, we cannot draw any conclusions regarding the validity of the classic I/E and GI/E models for students from other grade levels and thus elementary school students in general. Possibly, the ability to adequately engage in social and dimensional comparisons might only evolve during elementary school years so that our findings would not be replicable with even younger students. Thus, future studies are also needed to address the generalizability of the present findings to students attending other grade levels as well as to students from other countries, cultures, and educational systems.

Finally, this study observed the size and direction of within-domain and cross-domain achievement–outcome relations to infer students' use of social and dimensional comparisons. One could assume that the high correlations among achievement measures might invoke the negative cross-domain relations due to a suppression effect (Shieh, 2006). In this case, the negative cross-domain relations between achievement on the one hand and self-concept or other outcome variables on the other hand would only be a statistical artefact and not depict the use of dimensional comparisons. Yet, the vast amount of findings from cross-sectional and longitudinal quantitative studies (e.g., Chiu, 2012; Marsh, Abduljabbar et al., 2015; Marsh & Hau, 2004; Marsh & Köller, 2004; Möller et al., 2011; Niepel et al., 2014), along with the findings of qualitative (Möller & Husemann, 2006) experimental studies (Möller & Köller, 2001; Strickhouser & Zell, 2015), as well as meta-analytic findings (Möller et al., 2009) rather indicated the existence of dimensional comparisons and their use in the formation of self-concepts. Still, respective evidence should be also obtained with elementary school students.

Conclusion

The present study contributes to research on the I/E model and dimensional comparisons in several ways. This study helped clarify the inconsistent findings of previous studies on the I/E model with elementary school students by overcoming shortcomings of previous studies that may have made it difficult to unveil dimensional comparison effects. Beyond testing the classic I/E model, it is one of the first testing a GI/E model with third grade elementary school students. The findings demonstrated that even third grade elementary school students make use of social and dimensional comparisons for the formation of self-concepts and intrinsic values related to the domains of math, German, and physical ability. Hence, the findings of the present study argue that dimensional comparisons are a valid psychological phenomenon that transcends the academic domain and relates to non-academic domains such as physical ability (Chanal et al., 2009; see also Helm, Abele, Müller-Kalthoff, & Möller, 2017) and which also plays a role in the formation of motivational and emotional self-perceptions going beyond academic self-concepts (see also Arens, Becker, & Möller, 2017; Goetz et al., 2008; Marsh, Abduljabbar et al., 2015; Schurtz et al., 2014). The study offers directions for future research on the GI/E model with elementary school students which should consider longitudinal, introspective, and experimental studies, and include even more predictor and outcome variables.

Footnotes

[1] The SDQ I also includes a scale related to students' general school self-concept and general school intrinsic value which reflect students' self-perceived competence respectively enjoyment related to all school subjects. Moreover, the SDQ I encompasses scales for assessing students' self-concepts related to physical appearance, peer-relations, parent-relations, and self-esteem. These scales were not considered in the present study.

[2] A subsample of $N = 48$ (9.3%) students had missing values on all three achievement measures (i.e., math achievement, verbal achievement, physical education achievement).

Compared to the subsample of students with valid achievement measures, this subsample did not differ in the mean levels on all SDQ I-GS scales [physical appearance self-concept: *t* (512) = -0.952; peer self-concept: *t* (512) = -0.443; physical ability self-concept: *t* (512) = -0.234; parent-relations self-concept: *t* (512) = -0.121; self-esteem: *t* (511) = 0.323; general school self-concept: t (512) = 1.406; math self-concept: *t* (512) = 0.269; general school intrinsic value: *t* (512) = 0.677; German intrinsic value: *t* (512) = 1.873; math intrinsic value: *t* (512) = 0.478, physical ability intrinsic value: *t* (509) = 1.441; all *ns*] with the exception of German self-concept [*t* (512) = 2.193; $p < .05$]. Students with missing values on all achievement measures had a lower mean level of German self-concept ($M = 3.60$; $SD = 1.06$) than students with achievement measures ($M = 3.93$; $SD = 1.00$). Moreover, the subsample of students with missing values on all three achievement measures did not differ with respect to gender ratio (without missing values: boys: $N = 247$, girls: $N = 219$; with missing values: boys: $N = 26$, girls: $N = 22$; χ^2 (1) = 0.024, *ns*) or mean age (*t* (512) = 0.475, *ns*).

[3] A CFA model assuming that the self-concept items and intrinsic value items related to math, German, and physical ability load on a common factor along with separate factors for math achievement, verbal achievement, and physical education achievement (i.e., a 6-factor model) had an inadequate fit to the data: $\chi^2(119) = 708.965$, CFI =.862, TLI = 822, RMSEA = .098. This finding supported the validity of the separation of the SDQ I-GS items related to math, German, and physical ability into a self-concept and intrinsic value factor.

[4] Similar findings were found for the within-domain and cross-domain relations among math self-concept, math intrinsic value, German self-concept, German intrinsic value, math achievement, and verbal achievement when estimating a latent regression model without including the physical domain, that is, physical education achievement, physical ability self-concept, and physical ability intrinsic value (see Table S2 of the Online Supplements).

References

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261–271.

Ardila, A., Rosselli, M., Matute, E., & Inozemtseva, O. (2011). Gender differences in cognitive development. *Developmental Psychology, 47*, 984–990. https://doi.org/10.1037/a0023819

Arens, A.K., Becker, M., & Möller, J. (2017). Social and dimensional comparisons in math and verbal test anxiety: Within- and cross-domain relations with achievement and the mediating role of academic self-concept. *Contemporary Educational Psychology, 51*, 240–252. https://doi.org/10.1016/j.cedpsych.2017.08.005

Arens, A.K., Bodkin-Andrews, G., Craven, R.G., & Yeung, A.S. (2014). Self-concept of Indigenous and non-Indigenous Australian students: Competence and affect components and relations to achievement. *Learning and Individual Differences, 32*, 93–103. https://doi.org/10.1016/j.lindif.2014.03.019

Arens, A.K., & Hasselhorn, M. (2015). Differentiation of competence and affect self-perceptions in elementary school students: Extending empirical evidence. *European Journal of Psychology of Education, 30,* 405–419. https://doi.org/10.1007/s10212-015-0247-8

Arens, A.K., Marsh, H.W., Craven, R.G., Yeung, A.S., Randhawa, E., & Hasselhorn, M. (2016). Math self-concept in preschool children: Structure, achievement relations, and generalizability across gender. *Early Childhood Research Quarterly, 36*, 391–403. https://doi.org/10.1016/j.ecresq.2015.12.024

Arens, A.K., Möller, J., & Watermann, R. (2016). Extending the internal/external frame of reference model to social studies: Self-concept and achievement in history and politics. *Learning and Individual Differences, 51*, 91–99. https://doi.org/10.1016/j.lindif.2016.08.044

Arens, A.K., Yeung, A.S., Craven, R.G., & Hasselhorn, M. (2011). The twofold multidimensionality of academic self-concept: Domain specificity and separation between competence and affect components. *Journal of Educational Psychology, 103,* 970–981. http://dx.doi.org/10.1037/a0025047

Arens, A.K., Yeung, A.S., Craven, R.G., & Hasselhorn, M. (2013). A short German version of the Self Description Questionnaire I: Theoretical and empirical comparability. *International Journal of Research & Method in Education, 36,* 415–438. https://doi.org/10.1080/1743727X.2012.710503

Arens, A.K., Yeung, A.S., & Hasselhorn, M. (2014). Native language self-concept and reading self-concept: Same or different? *The Journal of Experimental Education, 82*, 229–252. https://doi.org/10.1080/00220973.2013.813362

Baudson, T.G., Wollschläger, R., & Preckel, F. (2016). *THINK 1–4. Test zur Erfassung der Intelligenz im Kindesalter* (THINK 1-4. Test for assessing intelligence in children). Göttingen: Hogrefe.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen, & J.S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

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. https://doi.org/10.1037/a0019644

Butler, R. (1998). Age trends in the use of social and temporal comparison for self-evaluation: Examination of a novel developmental hypothesis. *Child Development, 69*, 1054–1073.

Castelli, D.M., Hillman, C.H., Buck, S.M., & Erwin, H.E. (2007). Physical fitness and academic achievement in third- and fifth-grade students. *Journal of Sport and Exercise Psychology, 29*, 239–252. https://doi.org/10.1123/jsep.29.2.239

Chanal, J.P., Marsh, H.W., Sarrazin, P.G., & Bois, J.E. (2005). Big-fish-little-pond effects on gymnastics self-concept: Social comparison processes in a physical setting. *Journal of Sport & Exercise Psychology, 27,* 53–70. https://doi.org/10.1123/jsep.27.1.53

Chanal, J.P., Sarrazin, P.G., Guay, F., & Boiché, J. (2009). Verbal, mathematics, and physical education self-concepts and achievements: An extension and test of the internal/external frame of reference model. *Psychology of Sport and Exercise, 10*, 61–66. https://doi.org/10.1123/jsep.27.1.53

Chapman, J.W., & Tunmer, W.E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology, 87*, 154–167. http://dx.doi.org/10.1037/0022-0663.87.1.154

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

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

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. https://doi.org/10.1037/a0025734

Conley, A.M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology, 104*, 32–47.

http://dx.doi.org/10.1037/a0026042

Craven, R.G., & Marsh, H.W. (2008). The centrality of self-concept construct for psychological wellbeing and unlocking human potential: Implications for child and educational psychologists. *Educational and Child Psychology, 25,* 104–118.

De Fraine, B., van Damme, J., & Onghena, P. (2007). A longitudinal analysis for gender differences in academic self-concept and language achievement: A multivariate multilevel growth model. *Contemporary Educational Psychology, 32*, 132–150. https://doi.org/10.1016/j.cedpsych.2006.10.005

Durik, A.M., Vida, M.N., & Eccles, J.S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology, 98*, 382–393. http://dx.doi.org/10.1037/0022-0663.98.2.382

Eccles, J.S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*, 215–225. https://doi.org/10.1177/0146167295213003

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

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

Ehm, J.-H., Lindberg, S., & Hasselhorn, M. (2014). Reading, writing, and math self-concept in elementary school children: influence of dimensional comparison processes. *European Journal of Psychology of Education, 29*, 277–294. https://doi.org/10.1007/s10212-013-0198-x

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

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

Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B.M., Häfner, I., Nagengast, B., & Trautwein, U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology, 51*, 1226–1240. https://doi.org/10.1037/dev0000028

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*, 663–677. http://dx.doi.org/10.1037/edu0000003

Gaspard, H., Wigfield, A., Jiang, Y., Nagengast, B., Trautwein, U., & Marsh, H.W. (2018). Dimensional comparisons: How academic track students' achievements are related to their expectancy and value beliefs across multiple domains. *Contemporary Educational Psychology, 52, 1–14*. https://doi.org/10.1016/j.cedpsych.2017.10.003

Goetz, T., Frenzel, A.C., Hall, N.C., & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology, 33*, 9–33. https://doi.org/10.1016/j.cedpsych.2006.12.002

Guo, J., Marsh, H.W., Morin, A.J.S., Parker, P.D., & Kaur, G. (2015). Directionality of the associations of high school expectancy-value, aspirations, and attainment: A Longitudinal Study. *American Educational Research Journal, 52,* 371–402. https://doi.org/ 10.3102/0002831214565786

Guo, J., Marsh, H.W., Parker, P.D., Morin, A.J.S., & Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in science: Dimensional comparison processes and expectancy-by-value interactions. *Learning and Instruction, 49,* 81–91. https://doi.org/10.1016/j.learninstruc.2016.12.007

Guo, J., Parker, P.D., Marsh, H.W., & Morin, A.J.S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology, 51,* 1163–1176. http://dx.doi.org/10.1037/a0039440

Harter, S. (1988). The construction and conservation of the self: James and Cooley revisited. In D.K. Lapsley & F.C. Power (Eds.), *Self, ego, and identity: Integrative approaches* (pp. 43–70). New York, NY: Springer New York.

Harter, S. (1999). *The construction of the self. A developmental perspective.* New York: The Guilford Press.

Helm, F., Abele, A.E., Müller-Kalthoff, H., & Möller, J. (2017). Applying dimensional comparison theory to the fundamental dimensions of social judgment – Agency and communion. *Learning and Individual Differences, 54,* 116–125. https://doi.org/10.1016/j.lindif.2017.01.011

Helm, F., Müller-Kalthoff, H., Nagy, N., & Möller, J. (2016). Dimensional comparison theory: Perceived subject similarity impacts on students' self-concepts. *AERA Open, 2(2).* https://doi.org/10.1177/2332858416650624

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

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

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

Hulleman, C.S., Godes, O., Hendricks, B.L., & Harackiewicz, J.M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology, 102*, 880–895. https://doi.org/10.1037/a0019506

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

Jansen, M., Schroeders, U., Lüdtke, O., & Marsh, H.W. (2015). Contrast and assimilation effects of dimensional comparisons in five subjects: An extension of the I/E model. *Journal of Educational Psychology, 107*, 1086–1101. http://dx.doi.org/10.1037/edu0000021

Klomsten, A.T., Skaalvik, E.M., & Espnes, G.A. (2004). Physical self-concept and sports: Do gender differences still exist? *Sex Roles, 50*, 119–127. https://doi.org/10.1023/B:SERS.0000011077.10040.9a

Köller, O., Zeinz, H. & Trautwein, U. (2008). Class-average achievement, marks, and academic self-concept in German primary schools. In H.W. Marsh, R.G. Craven & D.M. McInerney (Eds.), *Self-processes, learning, and enabling human potential: Dynamic new approaches* (pp. 331-352). Charlotte, NC: Information Age Publishing Inc.

Lohbeck, A., & Möller, J. (2017). Social and dimensional comparison effects on math and reading self-concepts of elementary school children. *Learning and Individual Differences, 54,* 73–81. https://doi.org/10.1016/j.lindif.2017.01.013

Marsh, H.W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*, 129–149. https://doi.org/10.3102/00028312023001129

Marsh, H.W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81*, 417–430.

Marsh, H.W. (1994). Confirmatory factor analysis models of factorial invariance: A multifaceted approach. *Structural Equation Modeling, 1*, 5–34. https://doi.org/10.1080/10705519409539960

Marsh, H.W., Abduljabbar, A. S., Abu-Hilal, M.M., Morin, A.J.S., Abdelfattah, F.A., Leung, K.C., Xu, M.K., Nagengast, B., & Parker, P.D. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology, 105*, 108–128. http://dx.doi.org/10.1037/a0029907

Marsh, H.W., Abduljabbar, A.S., Parker, P.D., Morin, A.J.S., Abdelfattah, F.A., Nagengast, B., Möller, J., & Abu-Hilal, M.M. (2015). The internal/external frame of reference model of self-concept and achievement relations: Age-cohort and cross-cultural differences. *American Educational Research Journal, 52*, 168–202. https://doi.org/10.3102/0002831214549453

Marsh, H.W., & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology, 95*, 687–706. http://dx.doi.org/10.1037/0022-0663.95.4.687

Marsh, H.W., Craven, R.G., & Debus, R.L. (1999). Separation of competency and affect components of multiple dimensions of academic self-concept: A developmental perspective. *Merrill-Palmer Quarterly, 45*, 567–601.

Marsh, H.W., Ellis, L.A., & Craven, R.G. (2002). How do preschool children feel about themselves? Unraveling measurement and multidimensional self-concept structure. *Developmental Psychology, 38*, 376–393. https://doi.org/10.1037//0012-1649.38.3.376

Marsh, H.W., Gerlach, E., Trautwein, U., Lüdtke, O., & Brettschneider, W.-D. (2007). Longitudinal study of preadolescent sport self-concept and performance: Reciprocal effects and causal ordering. *Child Development, 78*, 1640–1656. . https://doi.org/10.1111/j.1467-8624.2007.01094.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. http://dx.doi.org/10.1037/0022-0663.96.1.56

Marsh, H.W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to cutoff values for fit indexes and dangers in overgeneralizing Hu & Bentler's (1999). *Structural Equation Modeling, 11*, 320–341. https://doi.org/10.1207/s15328007sem1103_2

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. https://doi.org/10.1016/S0361-476X(03)00034-1

Marsh, H.W., Kuyper, H., Seaton, M., Parker, P.D., Morin, A.J.S., Möller, J., & Abduljabbar, A.S. (2014). Dimensional comparison theory: An extension of the internal/external frame of reference effect on academic self-concept formation. *Contemporary Educational Psychology, 39*, 326–341. https://doi.org/10.1016/j.cedpsych.2014.08.003

Marsh, H.W., Lüdtke, O., Nagengast, B., Trautwein, U., Abduljabbar, A.S., Abdelfattah, F., & Jansen, M. (2015). Dimensional comparison theory: Paradoxical relations between self-beliefs and achievements in multiple domains. *Learning and Instruction, 35,* 16–32. https://doi.org/10.1016/j.learninstruc.2014.08.005

Marsh, H.W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A.J.S., Adbuljabbar, A., & Köller, O. (2012). Classroom climate and contextual effects. Methodological issues in the evaluation of group-level effects. *Educational Psychologist, 47*, 106–124. https://doi.org/10.1080/00461520.2012.670488

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. https://doi.org/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. https://doi.org/10.1111/j.1467-8624.2005.00853.x

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. https://doi.org/10.3102/00028312035004705

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

Möller, J., & Husemann, N. (2006). Internal comparisons in everyday life. *Journal of Educational Psychology, 98*, 342–353. https://doi.org/10.1037/0022-0663.98.2.342

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

Möller, J., Kuska, S.K., & Zaunbauer, A.M.C. (2011). Einflussfaktoren auf die Entwicklung des Selbstkonzepts im Grundschulalter. In F. Hellmich (Ed.), *Selbstkonzepte im Schulalter. Modelle – empirische Ergebnisse – pädagogische Konsequenzen* (pp. 161–172). Stuttgarg: Kohlhammer.

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

Möller, J., Müller-Kalthoff, H., Helm, F., Nagy, N., & Marsh, H.W. (2015). The generalized internal/external frame of reference model: An extension to dimensional comparison theory. *Frontline Learning Research, 3*, 37–47. https://doi.org/10.14786/flr.v4i2.16

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. https://doi.org/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. https://doi.org/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, 11*, 467–487. https://doi.org/10.1007/BF03173515

Muthén, L.K., & Muthén, B.O. (1998-2017). *Mplus User's Guide. Eighth Edition.* Los Angeles, CA: Muthén & Muthén.

Nagengast, B., Marsh, H.W., Scalas, L.F., Xu, M.K., Hau, K.-T., & Trautwein, U. (2011). Who took the "x" out of expectancy-value theory?: A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22,* 1058–1066. https://doi.org/10.1177/0956797611415540

Nicholls, J.G. (1979). Development of perception of own attainment and causal attributions for success and failure in reading. *Journal of Educational Psychology, 71*, 94–99. http://dx.doi.org/10.1037/0022-0663.71.1.94

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

Nowell, A., & Hedges, L.V. (1998). Trends in gender differences in academic achievement from 1960 to 1994: An analysis of differences in mean, variance, and extreme scores. *Sex Roles, 39*, 21–43. https://doi.org/10.1023/A:1018873615316

Pinxten, M., Marsh, H.W., De Fraine, B., Van Den Noortgate, W., & van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology, 84*, 152–174. https://doi.org/10.1111/bjep.12028

Pinxten, M., Wouters, S., Preckel, F., Niepel, C., De Fraine, B., & Verschueren, K. (2015). The formation of academic self-concept in elementary education: A unifying model for external and internal comparisons. *Contemporary Educational Psychology, 41,* 124–132. ttps://doi.org/10.1016/j.cedpsych.2014.12.003

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. https://doi.org/10.1207/s15326985ep4103_4

Ruble, D., Boggiano, A. K., Feldman, N. S., & Loebl, J. H. (1980). Developmental analyses of the role of social comparison in self-evaluation. *Developmental Psychology, 16,* 105–115. http://dx.doi.org/10.1037/0012-1649.16.2.105

Schmidt, I., Brunner, M., Keller, L., Scherrer, V., Wollschläger, R., Baudson, T.G., et al. (2017). Profile formation of academic self-concept in elementary school students in grades 1 to 4. *PLoS ONE* 12(5). https://doi.org/10.1371/journal.pone.0177854

Schurtz, I., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction, 34,* 32–41. https://doi.org/10.1016/j.learninstruc.2014.08.001

Shieh, G. (2006). Suppression situations in multiple linear regression. *Educational and Psychological Measurement, 66*, 435–447. https://doi.org/10.1177/0013164405278584

Sibley, B.A., & Etnier, J.L. (2003). The relationship between physical activity and cognition in children: A meta-analysis. *Pediatric Exercise Science, 15,* 243–256. https://doi.org/10.1123/pes.15.3.243

Simpkins, S.D., Davis-Kean, P.E., & Eccles, J.S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*, 70–83. https://doi.org/10.1037/0012-1649.42.1.70

Skaalvik, E.M., & Rankin, R.J. (1994). Gender differences in mathematics and verbal achievement, self-perceptions and motivation. *British Journal of Educational Psychology, 64*, 419–428. https://doi.org/10.1111/j.2044-8279.1994.tb01113.x

Skaalvik, E.M., & Rankin, R.J. (1995). A test of the internal/external frame of reference model at different levels of math and verbal self-perception. *American Educational Research Journal, 32*, 161–184. http://dx.doi.org/10.2307/1163217

Skaalvik, E.M., & Skaalvik, S. (2004). Gender differences in math and verbal self-concept, performance expectations, and motivation. *Sex Roles, 50*, 241–252. https://doi.org/10.1023/B:SERS.0000015555.40976.e6

Steffens, M.C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology, 102*, 947–963. https://doi.org/10.1037/a0019920

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

Strickhouser, J.E., & Zell, E. (2015). Self-evaluative effects of dimensional and social comparison. *Journal of Experimental Social Psychology, 59,* 60–66. https://doi.org/10.1016/j.jesp.2015.03.001

Swann, W.B., Chang-Schneider, C., & Larsen McClarty, K. (2007). Do people's self-views matter? Self-concept and self-esteem in everyday life. *American Psychologist, 62*, 84–94. https://doi.org/10.1037/0003-066X.62.2.84

Tay, M.P., Licht, B., & Tate, R.L. (1995). The internal/external frame of reference in adolescents' math and verbal self-concepts: A generalization study. *Contemporary Educational Psychology, 20,* 393–402. https://doi.org/10.1006/ceps.1995.1026

Trautwein, U., Gerlach, E., & Lüdtke, O. (2008). Athletic classmates, physical self-concept, and free-time physical activity: A longitudinal study of frame of reference effects.

Journal of Educational Psychology, 100, 988–1001. http://dx.doi.org/10.1037/0022-0663.100.4.988

Trautwein, U., Marsh, H.W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*, 763–777. http://dx.doi.org/10.1037/a0027470

Trautwein, U., & Möller, J. (2016). Self-concept: Determinants and consequences of academic self-concept in school contexts. In A.A. Lipnevich, F. Preckel, & R.D. Roberts (Eds.), *Psychosocial skills and school systems in the 21st century. Theory, research, and practice* (pp. 187–214). Berlin, New York: Springer.

Valentine, J.C., DuBois, D.L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist, 39*, 111–131. http://dx.doi.org/10.1207/s15326985ep3902_3

Watt, H.M.G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th- through 11th-grade Australian students. *Child Development, 75,* 1556–1574. https://doi.org/10.1111/j.1467-8624.2004.00757.x

Watt, H.M.G., Shapka, J.D., Morris, Z.A., Durik, A.M., Keating, D.P., & Eccles, J.S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from Australia, Canada, and the United States. *Developmental Psychology, 48*, 1594–1611. https://doi.org/10.1037/a0027838

Weiss, E.M., Kemmlera, G., Deisenhammerb, E.A., Fleischhackera, W.W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences, 35*, 863–875. https://doi.org/10.1016/S0191-8869(02)00288-X

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

Wigfield, A., & Eccles, J.S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J.S. Eccles (Eds.), *A Vol. in the educational psychology series. Development of achievement motivation* (pp. 91-120).

Appendix

Items of the SDQ I-GS Scales and their Standardized Factor Loadings

	λ
Physical Ability Self-Concept	
I am good at sports.	.908
I am a good athlete.	.876
Physical Ability Intrinsic Value	
I enjoy sports and games.	1.000
German Self-Concept	
I am good at German.	.795
Work in German is easy for me.	.841
I learn things quickly in German.	.842
German Value	
I like German.	.830
I enjoy doing work in German.	.862
I look forward to German.	.911
Math Self-Concept	
Work in mathematics is easy for me.	.830
I learn things quickly in mathematics.	.795
I am good at mathematics.	.888
Math Value	
I look forward to mathematics.	.849
I am interested in mathematics.	.892
I like mathematics.	.923

Note. For all $p < .001$.

Table 1

Factor Correlations from a CFA Model

	Physical ability self-concept	Physical ability intrinsic value	German self-concept	German intrinsic value	Math self-concept	Math intrinsic value	Math ACH	German ACH
Physical ability intrinsic value	.643*							
German self-concept	.164*	.049						
German intrinsic value	.039	.027	.794*					
Math self-concept	.094	.092	.181*	-.014				
Math intrinsic value	.106*	.165*	.053	.096	.769*			
Math achievement	-.001	-.078	.137*	-.100*	.589*	.309*		
German achievement	.018	-.108*	.437*	.203*	.247*	.050	.667*	
Physical achievement	.449*	.226*	.151*	-.011	.138*	.037	.321*	.316*

Note. CFA = confirmatory factor analysis. ACH = Achievement. The fit of this model was $\chi^2(99) = 158.707$, CFI = .986, TLI = .978, RMSEA = .034.

* $p \leq .05$.

Table 2

Standardized Path Coefficients of the Extended I/E Model

Outcome	**Predictor**	
Math self-concept		
	Math achievement	.769*
	German achievement	-.258*
	Physical education achievement	-.027
Math intrinsic value		
	Math achievement	.505*
	German achievement	-.275*
	Physical education achievement	-.039
German self-concept		
	Math achievement	-.289*
	German achievement	.615*
	Physical education achievement	.050
German intrinsic value		
	Math achievement	-.417*
	German achievement	.491*
	Physical education achievement	-.032
Physical ability self-concept		
	Math achievement	-.123*
	German achievement	-.060
	Physical education achievement	.508*
Physical ability intrinsic value		
	Math achievement	-.069
	German achievement	-.156*
	Physical education achievement	.297*

Note. The fit of this model was $\chi^2(99) = 158.707$, CFI = .986, TLI = .978, RMSEA = .034.
I/E Model = internal/external frame of reference model.
* $p \leq .05$.

Table 3

Goodness-of-fit Indices of the Gender-Invariance Models

Model		χ^2	df	CFI	TLI	RMSEA
I1	Configural invariance	280.861	198	.981	.971	.040
I2	Factor loading invariance	289.289	207	.981	.972	.039
I3	Factor loading invariance and invariance of item intercepts	301.141	216	.981	.972	.039
I4	Factor loading invariance, invariance of item intercepts, and invariance of item uniquenesses	353.131	230	.972	.962	.046
I5	Factor loading invariance, invariance of item intercepts, invariance of item uniquenesses, and invariance of factor variances	377.761	239	.968	.959	.048
I6	Factor loading invariance, invariance of item intercepts, invariance of item uniquenesses, invariance of factor variances, and invariance of factor covariances	418.441	275	.967	.963	.045
I7	Factor loading invariance, invariance of item intercepts, invariance of item uniquenesses, invariance of factor variances, invariance of factor covariances, and invariance of factor means	504.874	284	.949	.946	.055
I8	GI/E model; free across gender	289.289	207	.981	.972	.039
I9	GI/E model; invariant across gender	306.471	225	.981	.975	.038

Note. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation. All models were conducted with the MLR estimator. All χ^2 values are significant ($p < .001$).

Table 4

Manifest Mean Levels (M), Standard Deviations (SD), and Range of all Variables for Boys and Girls

	Boys			Girls		
	M	*SD*	*Range*	*M*	*SD*	*Range*
Math self-concept	4.24	(0.98)	1-5	3.88	(0.94)	1-5
Math value	4.00	(1.20)	1-5	3.75	(1.17)	1-5
German self-concept	3.72	(1.09)	1-5	4.10	(0.86)	1-5
German value	3.32	(1.26)	1-5	4.06	(0.98)	1-5
Physical ability self-concept	4.46	(0.89)	1-5	4.23	(0.98)	1-5
Physical ability intrinsic value	4.47	(0.96)	1-5	4.14	(1.15)	1-5
Math achievement	5.15	(0.72)	3-6	4.96	(0.76)	2-6
Verbal achievement	4.80	(0.74)	2-6	5.00	(0.71)	2-6
Physical education achievement	5.25	(0.58)	3-6	5.16	(0.64)	3-6

Figure 1

The classic I/E model according to Marsh (1986)

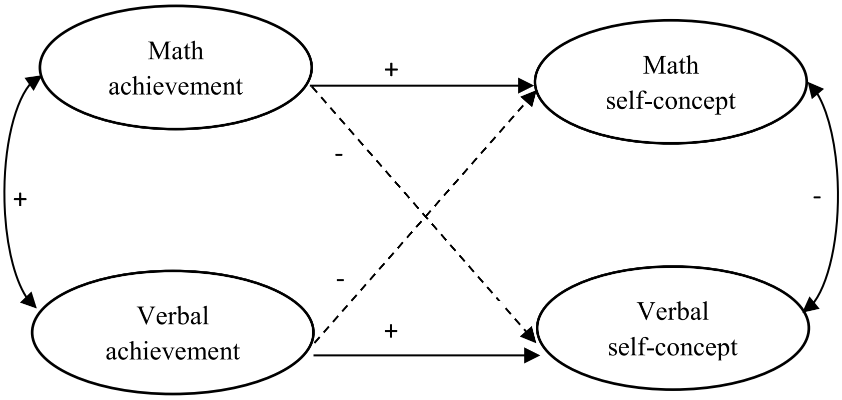
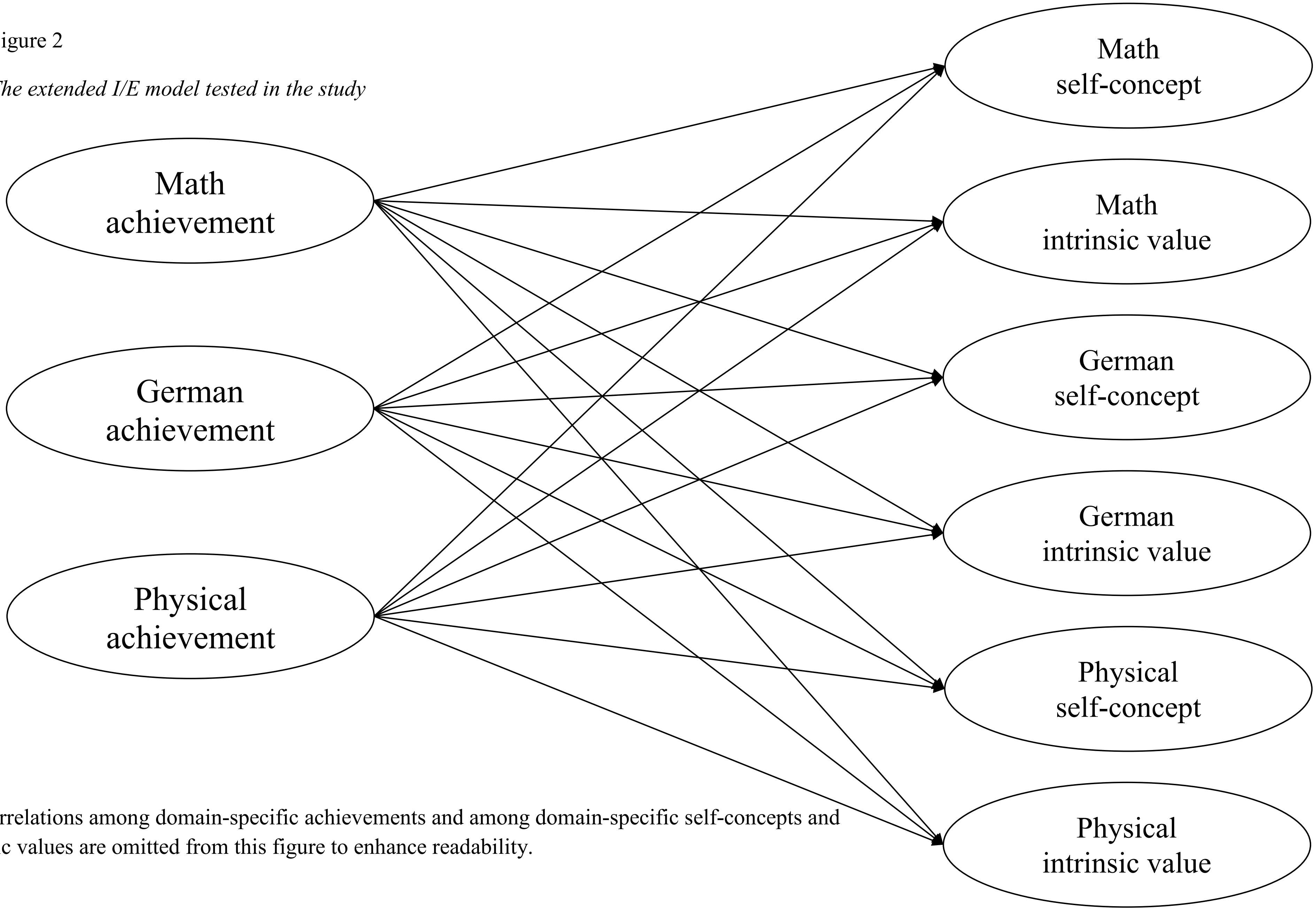


Figure 2

The extended I/E model tested in the study



Note.
The correlations among domain-specific achievements and among domain-specific self-concepts and intrinsic values are omitted from this figure to enhance readability.

Online Supplements to

Testing the internal/external frame of reference model with elementary school children: Extension to physical ability and intrinsic value

Table S1

Factor Correlations of the Extended I/E Model

	Physical ability self-concept	Physical ability intrinsic value	German self-concept	German intrinsic value	Math self-concept	Math achievement	German achievement
Physical ability intrinsic value	.611*						
German self-concept	.170*	.098*					
German intrinsic value	.054	.063	.792*				
Math self-concept	.162*	.174*	.281*	.159*			
Math intrinsic value	.158*	.207*	.136*	.235*	.750*		
Math achievement							
German achievement						.667*	
Physical education achievement						.321*	.316*

Note. The fit of this model was $\chi^2(99) = 158.707$, CFI = .986, TLI = .978, RMSEA = .034.
I/E Model = internal/external frame of reference model.
* $p \leq .05$.

Table S2

Standardized Path Coefficients of the Extended I/E Model without the Measures for the Physical Ability Domain

Outcome	**Predictor**	
Math self-concept		
	Math achievement	.764*
	German achievement	-.261*
Math value		
	Math achievement	.498*
	German achievement	-.279*
German self-concept		
	Math achievement	-.281*
	German achievement	.625*
German value		
	Math achievement	-.422*
	German achievement	.489*

Note. The fit of this model was $\chi^2(74) = 167.245$, CFI = .973, TLI = .962, RMSEA = .050.
I/E Model = internal/external frame of reference model.
* $p \leq .05$.

Table S3

Factor Correlations of the Extended I/E Model without the Physical Ability Domain

	German self-concept	German intrinsic value	Math self-concept	Math achievement
German intrinsic value	.787*			
Math self-concept	.278*	.157*		
Math intrinsic value	.132*	.235*	.749*	
Math achievement				.666*

Note. The fit of this model was $\chi^2(74) = 167.245$, CFI = .973, TLI = .962, RMSEA = .050.

I/E Model = internal/external frame of reference model.

* $p \leq .05$.