Arens, A. Katrin; Becker, Michael; Möller, Jens

The internal/external frame of reference (I/E) model. Extension to five school subjects and invariance across German secondary school ability tracks

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
Learning and individual differences 67 (2018) 2018, S. 143-155

Bitte verwenden Sie in der Quellenangabe folgende URN oder DOI /
Please use the following URN or DOI for reference:
urn:nbn:de:0111-pedocs-167944
10.25656/01:16794

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-167944
https://doi.org/10.25656/01:16794

The internal/external frame of reference (I/E) model: Extension to five school subjects and invariance across German secondary school ability tracks

A. Katrin Arens
German Institute for International Educational Research, Germany

Michael Becker
German Institute for International Educational Research, Germany, &
Leibniz Institute for Science and Mathematics Education, Kiel, Germany

Jens Möller
University of Kiel

Author Note

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

We would like to thank the principal investigators of the BIJU study, Jürgen Baumert, Olaf Köller, and Kai S. Cortina for making the data available for us to conduct the study. 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., Becker, M., & Möller, J. (2018). The internal/external frame of reference (I/E) model: Extension to five school subjects and invariance across German secondary school ability tracks. *Learning and Individual Differences, 67,* 143–155. https://doi.org/10.1016/j.lindif.2018.07.005

This paper is not the copy of record and may not exactly replicate the authoritative document published in *Learning and Individual Differences.*

Abstract

The internal/external frame of reference (I/E) model assumes the interplay of social and dimensional comparisons in the formation of domain-specific academic self-concepts.
The present study tests the generalizability of the I/E model assumptions across students from different ability tracks. While the findings from previous studies implied the similar use of social comparisons with students from different ability tracks, evidence has been missing so far whether students from different ability tracks apply dimensional comparisons to the same extent. Students from lower ability tracks are said to be confronted with negative stereotypes and felt deprivation which might enforce or weaken the use of dimensional comparisons. For the analyses, students from the academic track (N=702) and the vocational track (N=528) of German secondary schools were included as these two groups represent two extreme groups of ability tracks which might thus maximize the power of detecting differences in the use of social and dimensional comparisons. Both the original I/E model only including math and verbal achievement and self-concepts measures and an I/E model extended to five school subjects (math, German, English, physics, and biology) were examined. The results indicated invariance across school tracks for both the original I/E model and the extended I/E model when controlling for students' gender, socioeconomic status, and cognitive ability.

Keywords: I/E model; ability tracking; secondary schooling; invariance tests; social and dimensional comparisons; self-concept

The internal/external frame of reference (I/E) model (Marsh, 1986, 1990a; Möller, Pohlmann, Köller, & Marsh, 2009) offers a theoretical explanation for the separation between math and verbal self-concepts by assuming that two comparison processes are at play in students' academic self-concept formation: social comparisons (comparison of one's own achievement with the achievement of other students in the same subject; Festinger, 1954) and dimensional comparisons (comparison of one's own achievement in one domain with one's own achievement in another domain; Möller & Marsh, 2013). The interplay between both comparison processes leads to positive achievement–self-concept relations within the math and verbal domains, but to negative achievement–self-concept relations across the math and verbal domains. In the context of testing the generalizability of the I/E model across student characteristics, the present study aimed to find out whether the I/E model is similarly applicable to students attending different ability tracks of secondary schooling. While so far, studies have indicated that students from different ability tracks make use of social comparisons for self-concept formation (Liem, Marsh, Martin, McInerney, & Yeung, 2013; Trautwein, Lüdtke, Marsh, & Nagy, 2009; Van Praag, Demanet, Stevens, & Van Houtte, 2017), little research has considered whether students of different ability tracks also make similar use of dimensional comparisons. Students attending lower ability tracks may use dimensional comparisons more often to emphasize individual strengths and areas of success. Yet, students attending lower ability tracks may also apply dimensional comparisons to a lesser extent as part of their anti-school attitudes.

While the original I/E model only involves math and verbal achievement and self-concept measures, recent studies have extended the I/E model to multiple school subjects (Arens, Möller, & Watermann, 2016; Jansen, Schroeders, Lüdtke, & Marsh, 2015; Marsh et al., 2014; Marsh, Lüdtke et al., 2015; Marsh & Yeung, 2001; Möller, Streblow, Pohlmann, & Köller, 2006; Niepel, Brunner, & Preckel, 2014). Here, dimensional comparison processes were not only found to lead to *negative* cross-domain achievement–self-concept relations (i.e., contrast effects), but to also invoke *positive* cross-domain achievement–self-concept relations (i.e., assimilation effects). Therefore, we first tested the generalizability of the original I/E model involving math and verbal achievement and self-concept measures across different ability tracks. In a second step, we tested the generalizability of an I/E model extended to five school subjects (math, German, English, physics, and biology). For doing so, we used a sample of secondary school students in Germany. The German educational system is known for its strict ability tracking. It traditionally leads to two groups of students – students attending the academic track as the highest ability track and students attending the vocational track as the lowest ability track. These groups can be considered as two extreme groups which might help unveil potential differences in the validity of the I/E model between students from different secondary school ability tracks.

1. The I/E Model

Students' academic self-concept – that is, their self-perception of their academic competence – is one of the most important motivational constructs in educational psychology since high levels of academic self-concept have been found to be associated with a variety of desirable educational outcomes such as high levels of achievement, interest, and aspirations (e.g., Marsh & Craven, 2006; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Nagengast & Marsh, 2012). Previous research has stated the domain specificity of academic self-concept as students have been found to display separate math and verbal self-concepts which are only weakly related to each other (Marsh, 1986, 1990a; Möller et al., 2009). Even structural models which assume a global academic self-concept beyond math and verbal self-concepts demonstrate low or negative relations between math and verbal self-concepts (Brunner et al., 2010). The finding of a low correlation between math and verbal self-concepts was surprising in the first place given that math and verbal achievements are substantially correlated. Hence, math and verbal self-concepts were expected to be highly correlated as well. The I/E model (Marsh, 1986, 1990a) offers a theoretical explanation for the low correlation between math and verbal self-concepts despite a substantial correlation between math and verbal achievements. The model assumes that the formation of individual students' math and verbal self-concepts is shaped by an interplay of social and dimensional achievement comparison processes. In *social comparisons*, students compare their own achievement in one domain with their peers' achievement in the same domain. Given that math and verbal achievements are positively correlated, and that math and verbal self-concepts are the subjective representations of these domain-specific achievements, social comparisons are assumed to lead to a high correlation between math and verbal self-concepts. In addition, social comparisons are assumed to invoke positive relations between individual students' achievements and self-concepts in the same (matching) domains (e.g., a positive relation between math achievement and math self-concept) since higher achievement (compared to others) lead to a higher self-concept in the same domain.[1]

In *dimensional comparisons*, students contrast their own achievements across domains. Dimensional comparisons are assumed to invoke a negative correlation between math and verbal self-concepts. If an individual student perceives himself/herself as more able in math, the student's math self-concept increases, but the verbal self-concept decreases. Dimensional comparisons further lead to negative achievement–self-concept relations across domains. For instance, higher math achievement strengthens a student' math self-concept (see social comparisons), but weakens an individual students' verbal self-concept since it enforces an individual's perception of being better in math than in the verbal domain. Hence, with regard to the logic of dimensional comparisons, math (verbal) achievement and verbal (math) self-concept are negatively correlated. The joint operation of social and dimensional comparisons leads to the consistently observed low correlation between math and verbal self-concepts. The I/E model is usually depicted in terms of a path model in which math and verbal self-concepts are regressed on math and verbal achievements. According to the I/E model, math and verbal self-concepts show a lower correlation to each other than math and verbal achievements, and math and verbal achievements and self-concepts share a positive relation within domains, but a negative relation across domains.

2. Extending the I/E Model to Multiple School Subjects

The I/E model has inspired researchers to focus on dimensional comparisons in more detail. As a result, dimensional comparison theory (DCT; Möller & Marsh, 2013) has been proposed which points to the antecedents, psychological processes, and consequences associated with dimensional comparisons as well as to the scope of application. In DCT, dimensional comparisons are defined more broadly as taking place when an individual compares his/her perceptions of aspects of a particular domain A with his/her perceptions of aspects of a particular domain B, this comparison bearing an impact on outcomes related to these domains. In order to empirically test this broader approach to dimensional comparisons, the generalized internal/external frame of reference (GI/E) model (Möller, Müller-Kalthoff, Helm, Nagy, & Marsh, 2016) has been established in the context of DCT. The GI/E model proposed several extensions to the original I/E model. The original I/E model only includes math and verbal achievement and self-concept measures. In this case, the achievement–self-concept relations across domains [i.e., between math (verbal) achievement and verbal (math) self-concept] are negative. A prominent extension of the original I/E model based on the GI/E model refers to its extension to multiple school subjects (Arens et al., 2016; Jansen et al., 2015; Marsh et al., 2014; Marsh, Lüdtke et al., 2015; Marsh & Yeung, 2001; Möller, Streblow, Pohlmann, & Köller, 2006; Niepel et al., 2014). In this case, the original I/E model including math and verbal achievements and self-concept was consistently replicated. Yet, the relations between achievement and self-concept measures across other domains have been found to become negative as well as positive. Hence, dimensional comparisons can invoke contrast effects (negative achievement–self-concept relations across domains) as well assimilation effects (positive achievement–self-concept relations across domains), depending on the domains that are compared.

Contrast effects have been found between math and verbal achievement and self-concept measures as depicted in the original I/E model (e.g., Marsh, 1986; Marsh & Hau, 2004; Möller et al., 2009; Skaalvik & Rankin, 1995). In addition, contrast effects have been observed between math-like and verbal-like achievement and self-concept measures, that is, between achievement and self-concept measures which are conceptually related to the math and verbal domains. For example, German students have been found to show negative relations (i.e., contrast effects) between English (a verbal-like domain related to German students' foreign language) and physics (a math-like domain) achievement and self-concept measures (Arens et al., 2016; Marsh, Lüdtke et al., 2015; Möller, Streblow, Pohlmann, & Köller, 2006). Therefore, contrast effects emerge between conceptually dissimilar subject domains (Helm, Müller-Kalthoff, Nagy, & Möller, 2016; Möller, Streblow, & Pohlmann, 2006). In the Marsh/Shavelson model of academic self-concept (Marsh, 1990b), domain-specific self-concepts are located on a continuum ranging from a verbal pole (represented by the self-concept in the language of instruction) to a math pole (represented by math self-concept). Self-concepts related to other domains are placed along this continuum, somewhere between these two endpoints. Contrast effects are assumed to exist between domains which are located far from each other on this continuum, that is, between math or math-like and verbal or verbal-like domains.

Assimilation effects are assumed to exist between domains which are located close to each other on this continuum (Marsh, Lüdtke et al., 2015), and between domains that are perceived to be conceptually similar to each other (Helm et al., 2016; Möller, Streblow, & Pohlmann, 2006). Correspondingly, assimilation effects have consistently been demonstrated between math and physics achievements and self-concepts as two conceptually related (i.e., math-like) domains closely located near the math endpoint (Arens et al., 2016; Jansen et al., 2015; Marsh, Lüdtke et al., 2015; Möller, Streblow, Pohlmann, & Köller, 2006). Assimilation effects would also be expected between achievements and self-concepts related to different languages (e.g., native and non-native languages) as both domains are verbal-like and placed close to each other in the vicinity of the verbal endpoint of the academic self-concept continuum. However, findings from respective studies were inconsistent and demonstrated only small relations between achievements and self-concepts related to different languages. These relations were reported to be negative indicating a contrast effect in some studies (Marsh, Kong, & Hau, 2001; Marsh, Lüdtke et al., 2015; Marsh & Yeung, 2001; Niepel et al., 2014; Xu et al., 2013). Positive relations have also been reported indicating an assimilation effect in other studies (Marsh et al., 2014; Möller, Streblow, Pohlmann, & Köller, 2006).

Some subject domains such as biology have been placed in the middle of the continuum of academic self-concepts with equidistant relations to the math and verbal endpoints. For these domains, neither contrast nor assimilation effects can be found. The corresponding self-concepts have rather been shown to be independent from achievements in other domains (Jansen et al., 2015; Marsh, Lüdtke et al., 2015).

3. Generalizability of the I/E Model across Ability Tracks

In recent years, efforts have been made to test the generalizability of the I/E model assumptions – both when considering the original and extended I/E models. Respective studies demonstrated the generalizability of the I/E model assumptions across a wide range of student characteristics including gender and different cultural backgrounds (Marsh, Abduljabbar et al., 2015; Marsh & Hau, 2004; Marsh & Köller, 2004; Möller et al., 2009; Pinxten et al., 2015; Skaalvik & Rankin, 1990; Tay, Licht, & Tate, 1995). Hence, students seem to generally apply social and dimensional comparisons for the formation of academic self-concepts. In other words, the use of social and dimensional comparisons can be seen as a universal phenomenon (see also Möller & Husemann, 2006; Möller & Marsh, 2013). This consideration might also suggest the generalizability of the I/E model assumptions across students attending different ability tracks of secondary schooling. Indeed, the findings from quantitative and qualitative studies indicated that students of different ability tracks use social comparisons for self-concept formation (Liem et al., 2013; Trautwein et al., 2009; Van Praag et al., 2017). As outlined above, social comparisons lead to substantial positive achievement–self-concept relations within matching domains. Accordingly, Arens et al. (2017) could demonstrate that math achievement and math self-concept were substantially and similarly related to each other for German students attending the academic, intermediate, and vocational tracks.

Little research has, however, been conducted to examine whether students of different ability tracks use dimensional comparisons to a similar extent which would lead to the generalizability of the I/E model across ability tracks. When considering the situation of lower ability track students in contrast to the situation of higher ability track students, there are arguments for a stronger use of dimensional comparisons as well as arguments for a lesser use of dimensional comparisons with lower ability track students. Lower ability track schools suffer from a poor reputation in society given the low achievement levels of this student population, associated with poorer economic and social prospects and job opportunities (LeTendre, Hofer, & Shimizu, 2003; Knigge & Hannover, 2011; Solga, 2014; see also Ditton, 2013; Dumont, Protsch, Jansen, & Becker, 2017). Students are well aware of the negative stereotype associated with attending lower ability track schools (Hallam & Ireson, 2007; Van Praag et al., 2017). Lower ability track students are thus inclined to avoid the identification with the negative stereotype of low-ability and unsuccessful students. For this purpose, lower ability track students might highlight and emphasize domains of their own success which might become apparent by contrasting different domains, that is, by applying dimensional comparisons: “Students stressed individual intellectual capacities, skills or study motivations to make a clear distinction with their so called “unsuccessful” peers” (Van Praag et al., 2017, p. 615). Hence, lower ability track students might use dimensional comparisons to a greater extent than students from higher ability tracks.

The negative stereotype associated with attending lower ability tracks has served to explain an anti-school culture within lower ability track students. In reaction to the low societal prestige and status of attending lower ability track schools, feelings of inferiority and deprivation, and poorer and more insecure future expectations, lower ability track students might devalue the school system and refrain from the academic domain (Ireson & Hallam, 1999). As such, lower ability track students were found to report higher sense of futility and were characterized by a less study-oriented culture (Van Houtte, 2006; Van Houtte & Stevens, 2009, 2010). Hence, lower ability track students might be less concerned with school matters reducing their engagement in dimensional comparisons.

In addition, the differential learning environments of ability tracks might invoke a differential use of dimensional comparisons in students of different ability tracks. Higher ability tracks explicitly prepare students for university. This characteristic inherently leads to a more pronounced focus on achievement in the respective schools, and teachers' instruction practices might be more focused on students' learning, academic performance, and progress (Boaler, William, & Brown, 2000; Van Houtte, 2004, 2006). As such, for instance, teachers in higher ability track schools have been found to provide higher levels of cognitively activating instruction (Klusmann, Kunter, Trautwein, Lüdtke, & Baumert, 2008; see also Gamoran & Carbonaro, 2002; Gamoran, Nystrand, Berends, & LePore, 1995; Van Houtte, 2004). Given the relatively higher achievement orientation in higher ability track schools, the students might be concerned about their performance and compliance to achievement standards. The students might thus feel a need to understand and evaluate their own relative standing. Therefore, students from higher ability tracks might be particularly inclined to use social comparisons for academic self-concept formation.

Moreover, in higher ability tracks, students are often requested to select courses that are particularly relevant for graduation. Hence, for students attending higher ability tracks, it is important to be aware of one's own strengths and weaknesses and to refine one's profile across various school subjects. To this end, the students might apply social and dimensional comparisons (see studies documenting the role of social and dimensional achievement comparisons in coursework selection; Dickhäuser, Reuter, & Hilling, 2005; Guo, Marsh, Parker, Morin, & Dicke, 2017; Nagy, Trautwein, Köller, Baumert, & Garrett, 2006). Hence, the differential learning environments might lead higher ability track students to use dimensional comparisons more strongly compared to lower ability track students.

In sum, lower ability track students' need to dissociate themselves from the negative stereotype might argue for a stronger use of dimensional comparisons in lower ability track students on the one hand. The anti-school culture and differences in the learning environment between students from higher and lower ability tracks might argue for a weaker use of dimensional comparisons in lower ability track students on the other hand. Still, a finding presented in the supplementary analyses to the study by Jansen et al. (2015) showed that students of different ability tracks from the German educational system use dimensional comparisons to the same extent. In fact, the findings demonstrated the generalizability of an extended I/E model (integrating achievement and self-concept measures for German, math, biology, chemistry, and physics) across a group of German students attending the academic track and a group of German students attending other secondary school ability tracks. However, Jansen et al. only contrasted academic track (i.e., the highest ability track in Germany, see below) students to a group of students attending any other form of lower secondary school tracks, that is, all other forms of secondary school tracks apart from the academic track. However, the German educational system is characterized by a strict ability tracking procedure including multiple secondary school ability tracks which leads to a more differentiated separation of students from different ability tracks.

4. The Secondary School System in Germany

The German secondary school system traditionally bases on a three-tier tracking system in which the students either attend the academic track ("Gymnasium") or two forms of non-academic tracks: the intermediate track ("Realschule") and the vocational track ("Hauptschule"). The different school tracks lead to different secondary school-leaving certificates and occupational options. Successful graduation from the academic track after 12 or 13 school years allows students to enter university and offers admission to all professional careers. The school-leaving certificate from the intermediate track after grade 10 allows the students to begin a vocational training with an emphasis on administrative and commercial professions. After leaving the vocational track after grade nine, students can apply for an apprenticeship in trade and craft with a focus on manual labor (Becker, Neumann, & Dumont, 2016). Students are allocated to different ability tracks for secondary education primarily based on students' school accomplishments during elementary school, and teachers' recommendations, but the allocation to different ability tracks also depends on parents' and students' wishes.

In the context of this three-tier tracking system, German students can be differentiated into two groups of students which represent the extreme groups of the strict German ability tracking procedure, that is, academic track students and vocational track students. Vocational track students have been found to suffer from a negative stereotype and reputation in society (Knigge & Hannover, 2011; Solga, 2014). They might thus be inclined to avoid identification with this negative stereotype either by highlighting their own domains of success (strengthening the use of dimensional comparisons) or by refraining from the academic domain (weakening the use of dimensional comparisons). Hence, one might assume that students attending the vocational track might engage more or alternatively less dimensional comparisons than students attending the academic track.

Yet, so far, no studies have explicitly tested a potential differential use of dimensional comparisons between academic track and vocational track students from Germany. Instead, studies focusing on secondary school students in Germany to validate the I/E model assumptions have not distinguished between students attending different tracks, but merged students from different school tracks [Marsh & Köller, 2004; Marsh, Lüdtke et al., 2015 (Study 2); Möller, Retelsdorf, Köller, & Marsh, 2011; Möller, Zimmermann, & Köller, 2014]. Other studies only considered students from the academic track [Marsh, Lüdtke et al., 2015 (Study 1); Möller & Pohlmann, 2010; Möller, Streblow, Pohlmann, & Köller, 2006; Wolff, Helm, Zimmermann, Nagy, & Möller, 2018]. Hence, previous studies assumed that German students from the academic and vocational tracks do not differ in their use of social and dimensional comparisons and presumed the invariance of the I/E model across secondary school tracks. Finally, Jansen et al. (2015) only distinguished between students attending the academic track and students attending any other kind of lower ability track but did not differentiate further within the group of lower ability track students. This approach might have masked differences between academic track and vocational track students as the two extreme groups of the German ability tracking procedure. The present study helped clarify potential differences between academic track and vocational track students in the use of dimensional comparisons and thus to gain insight into the generalizability of the I/E model across these two groups of students.

5. The Present Study

We aimed to test whether the I/E model assumptions – both related to the original I/E model and an I/E model extended to five school subjects (math, German, English, biology, and physics) – similarly apply to students from different ability tracks of secondary schooling. To this aim, we focused on German students attending the academic track and students attending the vocational track to enhance the probability of unveiling differences in students' use of dimensional comparisons. On the one hand, the negative stereotype vocational track students in Germany are confronted with might induce these students to make stronger use of dimensional comparisons to emphasize and demonstrate individual strength. On the other hand, vocational track students might use dimensional comparisons to a lesser extent to circumvent involvement with academic matters. In addition, the academic focus of the learning environment of academic tracks might foster the use of dimensional comparisons in academic track students, also leading to a differential use of dimensional comparisons between academic track and vocational track students. Still, existing research has documented the generalizability of the I/E model assumption across different groups of students including students from different secondary school ability tracks (Marsh, Abduljabbar et al., 2015; Jansen et al., 2015; Marsh & Hau, 2004; Marsh & Köller, 2004; Möller et al., 2009; Pinxten et al., 2015; Skaalvik & Rankin, 1990; Tay et al., 1995). Hence, there are some empirical and theoretical arguments supporting the generalizability of the I/E model across academic and vocational track students as examined in our study, while the assumption of generalizability is rejected by other lines of argument. It was thus difficult and unwarranted to state clear a-priori assumptions so that our study remains exploratory.

The analyses for testing the generalizability of the I/E model across the German academic track and the German vocational track were conducted while controlling for student' gender, socioeconomic status (SES), and general cognitive ability (intelligence). These student characteristics were chosen as control variables as they have been found to be associated with students' academic self-concept, achievement, and tracking. Girls and boys have been found to differ in their mean levels of academic self-concept facets with these mean levels following gender stereotypes. Girls have been found to display higher levels of verbal self-concept, while boys have been found to report higher levels of math self-concept (e.g., Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Skaalvik & Skaalvik, 2004). Moreover, girls display higher achievement in verbal subjects than boys (e.g., De Fraine, Van Damme, & Onghena, 2007; Van de gaer, Pustjens, Van Damme, & De Munter, 2006). Boys have not been consistently found to demonstrate higher levels of math achievement. Instead, findings regarding gender differences on math achievement are more ambiguous across studies and seem to vary contingent upon the achievement indicator (standardized test scores or school grades; e.g., Brunner, Krauss, & Kunter, 2008; Hyde, Fennema, & Lamon, 1990; Matteucci &

Mignani, 2011). Finally, gender is associated with tracking since more boys than girls attend lower ability tracks (Hallam & Ireson, 2007; Jones, Vanfossen, & Ensminger, 1995).

Students from lower SES families demonstrate lower levels of achievement (Bradley & Corwyn, 2002; Sirin, 2005) and more likely attend the vocational track (Baumert, Watermann, & Schümer, 2003; Jones et al., 1995; Maaz, Trautwein, Lüdtke, & Baumert, 2008). High levels of cognitive abilities have often been found to be linked to high levels of student achievement (Frey & Detterman, 2004; Rhode & Thompson, 2007). Students with higher levels of cognitive ability are moreover more likely to attend the academic track of German secondary schooling (Arens et al., 2017).

In Germany, federal states are responsible for education. In recent years, the 16 German federal states have started to modify their secondary school system and thus to diverge from each other regarding the range of secondary school tracks. These differences among German federal states largely affect the number and structure of non-academic, lower ability tracks (i.e., the intermediate and vocational tracks), whereas the different federal states have consistently adhered to the academic track (Becker et al., 2016). In fact, some federal states have merged different forms of non-academic, lower ability tracks. In addition, comprehensive tracks have been introduced which open the opportunity to reach different school certificates. These modifications on the non-academic, lower ability school tracks can be seen as a reaction to the increasing reluctance of students and their parents to select the vocational track for secondary schooling. This reluctance might itself originate from the bad reputation and inferior job prospects of attending the vocational track (LeTendre et al., 2003; Knigge & Hannover, 2011; Solga, 2014; see also Dumont et al., 2017). The boundaries between German secondary school tracks have thus become blurred and school certification has become more flexible and less dependent on the specific school track attended (Becker et al., 2006).

To still be able to compare German students attending the traditional academic and vocational tracks regarding the use of social and dimensional comparisons and the validity of the I/E model assumption, we applied data from the large longitudinal project “Learning Processes, Educational Careers, and Psychosocial Development in Adolescence and Young Adulthood (BIJU)” conducted under the aegis of the Max Planck Institute for Human Development, Berlin, Germany (for more information on this data set, see for example Baumert et al., 1996). The BIJU study investigated the educational and psychosocial development from early adolescence up to young adulthood. Students from four German federal states (North Rhine-Westphalia, Mecklenburg-Western Pomerania, Saxony-Anhalt, and Berlin) participated in the overall BIJU study, but only students from the federal state of North Rhine-Westphalia could be grouped into academic track and vocational track students at the time. In the other German federal states participating in the BIJU study, vocational tracks were not part of the secondary school system at the time.

6. Method

6.1 Sample

In this study, we focused on the second measurement point of the BIJU study which took place in the midst of the 1991/1992 school year when the students attended grade level 7. We included students with valid information on their secondary school track and students who learned English as the first foreign language. The total study sample consisted of $N = 4001$ students [$N = 1884$ (47.1%) boys, $N = 2057$ (51.4%) girls, and $N = 60$ (1.5%) students without gender indicated] from 298 different classes of 169 different schools in four federal states (North Rhine-Westphalia: N=2441; Mecklenburg-Western Pomerania: $N = 166$; Saxony-Anhalt: $N = 235$; Berlin: $N = 1159$). In this study, we focused on the subsample of students from the federal state of North Rhine-Westphalia [$N = 2441$; $N = 1116$ (45.7%) boys, $N = 1293$ (53.0%) girls, $N = 32$ (1.3%) no gender indicated] as these students could be separated into a subsample of students attending the academic track [$N = 702$; $N = 460$ (65.5%) girls; N

= 237 (33.8%) boys; N = 5 (0.7%) students without indicated gender] and a subsample of students attending the vocational track [N = 528; N = 250 (47.3%) girls, N = 272 (51.5%) boys, N = 6 (1.1%) students without indicated gender]. Hence, we used a sample of N = 1230 students of academic and vocational track students from 68 classes in 34 schools. Data were assessed in entire classes by trained research assistants. Approval to realize the study was granted by the responsible ministries of education. Human subjects standards were approved by the ethics committee of the respective research institution. Informed consent was obtained from the participating students' parents, and the students were informed about the purpose of the study, the voluntary nature of participation, and the confidential treatment of the collected data.

6.2 Measures

6.2.1 Academic self-concept. Five items asking for students' self-perceptions of competence were used to measure students' academic self-concept related to the five school subjects of math, German, English, biology, and physics. The items were adapted from Jopt (1978) and Jerusalem (1984). They were formulated in parallel across the five domains as they had the same wordings and only differed in their targeted domain (e.g., "Nobody is perfect but I am just not good at math/German/English/biology/physics."; "I am not particularly good at math/German/ English/biology/physics."; "Math/German/ English/biology/physics just isn't my thing."). The items were rated on a 4-point Likert-type scale (1="does not apply at all" to 4="fully applies") in order that higher values indicated higher levels of self-concept. The different academic self-concept scales demonstrated good reliability estimates in terms of Cronbach's alpha reliability coefficient in the merged subsample of academic and vocational track students from North Rhine-Westphalia examined here: math self-concept: α = .904; German self-concept: α = .865; English self-concept: α = .858; biology self-concept: α = .858; physics self-concept: α = .884.

6.2.2 Achievement. The school grades the students had obtained in their last school report in math, German, English, biology, and physics as school subjects served as achievement indicators. In Germany, school grades range from 1 to 6, with 1 representing the best, and 6 the poorest grade. To facilitate interpretation of the results, the grades were reversely coded before all analyses, thus higher values indicated higher levels of achievement.[2]

6.2.3 Control variables. Students' gender, intelligence, and SES were used as control variables. Gender was a dichotomous variable (0 = girls; 1 = boys). For intelligence, two measures were used which were assessed along with the self-concept and achievement measures in the midst of the school year when students attended grade level 7. As a first measure of verbal intelligence, we used the verbal analogies subscale of the KFT 4–13+ (Heller, Schoen-Gaedike, & Weinlaeder, 1985), an adapted German version of Thorndike's Cognitive Abilities Test (Thorndike & Hagen, 1971). As a second measure of numerical intelligence, we used the numerical intelligence subtests of Amthauer's (1955) Intelligence Structure Test (IST). The reliability estimated by the Kuder-Richardson Formula 20 was adequate for both measures: verbal analogies subscale: $\alpha = .680$; numerical intelligence subtest: $\alpha = .949$. For obtaining a measure of students' SES, we considered the reports on mothers' and fathers' occupation. The information was coded using the International Standard Classification of Occupations (Ganzeboom & Treiman, 1996). The highest value of either mother's or father's occupational status was selected for each individual student and used as an indicator for this individual student's SES.

6.3 Statistical Analyses

All analyses were conducted within the framework of structural equation modeling (SEM; e.g., Kline, 2005) using the statistical package *Mplus 8.0* (Muthén & Muthén, 1998-2017). All models were estimated by applying the robust maximum likelihood estimator (MLR), which has been shown to be robust against violations of normality assumptions and accounts for the treatment of items responded on a Likert-type scale as continuous variables (Beauducel & Herzberg, 2006). Missing values on the achievement and self-concept variables were estimated by the Full Information Maximum Likelihood (FIML) implemented in Mplus. The FIML approach is known to be reliable in handling missing data, making less restrictive assumptions than, for example, listwise deletion (Enders, 2010; Graham, 2009).

The data set has a multilevel structure since the participating students were nested into classes (Raudenbush & Bryk, 2002). Students attending the same class can be assumed to be more similar to each other than students attending different classes so that the student ratings cannot be treated as independent observations. All analyses were therefore conducted using the *Mplus* option “type = complex” with students’ classes treated as clustering variables to correct for possible biased standard errors resulting from the hierarchical nature of the data. Furthermore, all models considered correlated uniquenesses between parallel-worded items for measuring domain-specific academic self-concepts to account for potential shared method variance attributed to the wordings of the items (Marsh et al., 2013).

The first set of models addressed the original I/E model and thus only considered math and verbal (i.e., German) achievement and self-concept measures. The second set of models addressed an extended I/E model integrating achievement and self-concept measures related to the five school subjects of math, German, English, biology, and physics. Within both sets, a confirmatory factor analytic (CFA) model (Brown, 2006) was first estimated in which separate factors for the domain-specific achievement and self-concept measures were assumed. The achievement factors were single-item factors defined by students’ school grades, while the self-concept factors were defined by the respective domain-specific set of items. Afterwards, the I/E model assumptions were tested in a latent regression model in which the domain-specific self-concepts served as outcome variables and were regressed on the domain-specific achievement factors as predictor variables.

For the purpose of comparing students from different ability tracks, measurement invariance has to be established as a precondition (Meredith, 1993; Millsap, 2011). Hence, students' school tracks were introduced as a grouping variable in the CFA model assuming separate achievement and self-concept factors. This grouping variable consisted of two groups, that is, students attending the academic track and students attending the vocational track. A first model of configural invariance stated the same number of factors defined by the same set of items with the parameter estimates freely estimated across groups. This model was expanded by assuming invariant factor loadings in a second step (weak measurement invariance; Meredith, 1993). The invariance of factor loadings ensures that the measures assess the same constructs with the same underlying meanings across groups, and it is essential when comparing the relations among constructs (i.e., achievement–self-concept relations as stated in the I/E model) across groups. Based on this model of invariant factor loadings, the I/E model in terms of regressions of domain-specific self-concepts on domain-specific achievements was freely estimated in each group. Based on the approach realized by Xu et al. (2013), we increasingly introduced invariance constraints on the paths coefficients. Here, we first restricted the paths of within-domain achievement–self-concept relations depicting social comparisons to be of equal size across groups. Second, we restricted the paths of between-domain achievement–self-concept relations depicting dimensional comparisons to be of equal size across groups. Finally, we restricted all paths (i.e., within-domain and between-domain achievement–self-concept relations) to be of equal size across groups. In all multi-group models using students' ability track as a grouping variable, students' gender, the two measures of students' intelligence, and the SES measure were introduced as manifest control variables.[3]

For model fit evaluation, we followed the advice to consider a wide range of descriptive goodness-of-fit indices (e.g., Marsh, Hau, & Wen, 2004; Marsh, Hau, & Grayson, 2005). Accordingly, we reported the comparative fit index (CFI), the Tucker-Lewis index

(TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). For the CFI and TLI, values above .90 and .95 represent an adequate respectively good model fit (Hu & Bentler, 1999). For the RMSEA, values should be below .05 for a close fit, or between .05 and .08 for a reasonable fit (Browne & Cudeck, 1993). Regarding the SRMR, Hu and Bentler (1999) propose values below .08 as indicative of a good model fit. These descriptive goodness-of-fit indices were also used to evaluate the invariance models. According to the guidelines proposed by Cheung and Rensvold (2002) and Chen (2007), invariance can be accepted as long as the CFI does not drop more than .01 between more and less restrictive models. These cut-off criteria for the different descriptive goodness-of-fit indices for the purpose of model fit and invariance evaluation should be treated as guidelines instead of "golden rules". Along with considering a range of resulting fit indices, researchers are advised to base their model evaluation on different types of information including the resulting parameter estimates, statistical conformity, and theoretical adequacy of the models (Marsh et al., 2004).

7. Results

7.1 The Original I/E Model

The CFA model for the student sample including academic track and vocational track students assuming separate factors for math and German achievements and self-concepts fitted the data well (Model 1 in Table 1)[4]. Based on this model, we stated the original I/E model (Model 2)[5] regressing math and German self-concepts on math and German achievements. The original I/E model assumptions could be replicated since the paths between math (German) self-concept and math (German) achievement were significantly positive (math: $\beta = .489$; German: $\beta = .326$; for both $p < .05$), while the paths leading from achievement to self-concept across domains were significantly negative: math achievement → German self-concept: $\beta = -.103$; German achievement → math self-concept: $\beta = -.129$; for both $p < .05$ (Table 2). The correlation between math and German self-concepts ($r = .333$, $p <$

.05) was lower than the correlation between math and German achievements ($r = .446$, $p < .05$) (Table S2 of the Online Supplements).

In the next step, the two groups of students attending the academic track and the vocational track were included as a grouping variable in the baseline CFA model. The model fit did not decline substantially between a model of configural invariance (Model 3) and a model of factor loading (i.e., weak measurement) invariance (Model 4). This finding of weak measurement invariance indicated that the same constructs were measured in both groups, allowing the comparisons of relations among constructs across groups. We thus freely estimated the I/E model across school track groups (Model 5)[6]. Compared to this model, the CFI value did not decline by more than $\Delta = .01$ when adding invariance constraints to the paths irrespective of whether these invariance constraints addressed the within-domain achievement–self-concept relations only (Model 6), the between-domain achievement–self-concept relations only (Model 7), or all paths of achievement–self-concept relations (Model 8, Table 2).

Regarding the covariates, boys demonstrated higher levels of math self-concept within both subsamples of academic track and vocational track students. In the academic track subsample, students with higher levels of numerical intelligence additionally showed higher levels of math self-concept. All control variables were related to math and German achievements in the subsample of academic track students in order that higher levels of verbal and numerical intelligence as well as higher levels of SES were positively associated with math and verbal achievements. Boys had higher levels of math achievement, and girls had higher levels of German achievement. In the vocational track student subsample, girls were also found to demonstrate higher levels of German achievement, and boys were also found to display higher levels of math achievement. Moreover, higher levels of verbal intelligence were positively associated with German achievement and higher levels of SES were positively associated with math achievement (Table 2).

7.2 The Extended I/E Model

The CFA model assuming separate achievement and self-concept factors related to five school subjects (math, German, English, biology, physics; Model 9 in Table 1) fitted the data well.[7] The corresponding extended I/E model (Model 10)[8] replicated the original I/E model given positive relations between math (German) achievement and math (German) self-concept (β = .486, resp. β = .294; both $p < .05$), and negative cross-domain relations between math achievement and German self-concept (β = -.137, $p < .05$) and between German achievement and math self-concept (β = -.107, $p < .05$; Table 3, see also Table S4 in the Online Supplements). When considering the other domains (i.e., English, biology, and physics), significantly positive within-domain achievement–self-concept relations were consistently found (English: β = .549; biology: β = .319; physics: β = .307; for all $p < .05$). Negative cross-domain contrast effects were found between English and math [see the paths leading from English achievement to math self-concept (β = -.094, $p < .05$), and the path leading from math achievement to English self-concept (β = -.218, $p < .05$)] indicating contrast effects between verbal-like (English) and math domains. Similarly, there was a negative path leading from German achievement to physics self-concept (β = -.119, $p < .05$) indicating a contrast effect between math-like and verbal domains. Surprisingly, the path between physics achievement and English self-concept was significantly positive (β = .118, $p < .05$). Regarding biology self-concept, corresponding to a contrast effect, the path leading from English achievement to biology self-concept was significantly negative (β = -.101, $p < .05$). Neither a contrast nor an assimilation effect was found for the relations between German and English achievements and self-concepts, but English (German) achievement was unrelated to German (English) self-concept (β = .023, resp. β = .040; both *ns*).

After testing these models based on a sample including both academic and vocational track students, we conducted models considering the two groups of ability tracks as a grouping variable and included gender, the two measures of intelligence, and SES as control variables. Weak measurement invariance across ability tracks could be demonstrated in Model 12 which revealed only a small decline in the CFI value (Δ = -.002) relative to the more relaxed model of configural invariance (Model 11). We thus freely estimated the extended I/E model across school tracks (Model 13)[9]. Relative to this model, the CFI value did not decline by more than Δ = .01 when restricting the path coefficients of the within-domain achievement–self-concept relations (Model 14), the path coefficients of the cross-domain achievement–self-concept relations (Model 15), and all path coefficients (Model 16) to invariance across academic and vocational track students.

Considering the relations between the covariates and the academic self-concept facets assessed here, boys were found to display higher levels of math and physics self-concepts regardless of the secondary school ability track attended. In the academic track, higher levels of numerical intelligence were additionally found to be related to higher levels of math self-concept. Regarding the relations between the covariates and achievement, girls showed higher levels of German achievement in both academic and vocational tracks, while boys had higher levels of math and physics achievements in both tracks. Verbal intelligence was positively associated with German and biology achievements in both the academic and vocational tracks, and with math, English, and physics achievements in the academic track. Numerical intelligence was positively related to German, math, English, and biology achievements in the academic track, but unrelated to the achievement measures in the vocational track. Higher SES students had higher levels of German and biology achievements in the academic track. In both the academic and vocational tracks, higher SES students showed higher levels of math and physics achievements (Table 3).

8. Discussion

8.1 The Generalizability of the I/E Model across School Tracks

The I/E model assumes that the formation of students' domain-specific academic self-concepts relies on an interplay between social and dimensional comparisons. Recently, many studies have tested the generalizability of the I/E model across student characteristics (e.g., Marsh, Abduljabbar et al., 2015; Marsh & Hau, 2004). Another prominent line of research targets the extension of the I/E model to multiple school subjects (Arens et al., 2016; Jansen et al., Marsh et al., 2014; Marsh, Lüdtke et al., 2015; Marsh & Yeung, 2001; Möller, Streblow, Pohlmann, & Köller, 2006; Niepel et al., 2014). Combining these two contemporary lines of research, the present study examined whether the original I/E model and its extension to five school subjects are applicable to students attending different secondary school ability tracks.

Previous studies implied the similar use of social comparisons across students from different ability tracks. It has yet remained unclear whether students from different ability tracks apply dimensional comparisons in similar ways. Given that lower ability track students suffer from a negative stereotype in society and felt deprivation (LeTendre et al., 2003; Knigge & Hannover, 2011; Solga, 2014), they might use dimensional comparisons to a stronger extent in order to distance themselves from the negative stereotype by highlighting their own strengths. Alternatively, the negative stereotype and the consequential anti-school culture may lead lower ability track students to dissociate from the academic and school domains, weakening their engagement in dimensional comparisons. Still, invariance might also be assumed given the so far found generalizability of the I/E model assumptions across various student groups, and thus universality of social and dimensional comparisons (e.g., Jansen et al., 2015; Marsh, Abduljabbar et al., 2015; Marsh & Hau, 2004).

In order to investigate the generalizability of the use of dimensional comparisons and the I/E model assumptions across students from different ability tracks, we compared students attending the academic track (“Gymnasium”) and students attending the vocational track (“Hauptschule”) in Germany. The comparison of academic and vocational track students might maximize the power of detecting any differences in the use of dimensional comparisons and thus in the validity of the I/E model between students of different ability tracks since these two groups can be considered as two extreme groups of students attending different secondary school ability tracks.

Our findings indicated invariance of the paths coefficients across students from the academic and vocational tracks both when considering the within-domain achievement–self-concept relations indicating social comparisons and when considering the between-domain achievement–self-concept relations indicating dimensional comparisons. Hence, students attending the academic and vocational tracks do not seem to differ in their use of social and dimensional comparisons, but the findings argue for the generalizability of the I/E model assumptions across students from different ability tracks. Generalizability of the I/E model across students from different ability tracks was found for the original I/E model as well as for an I/E model extended to five school subjects.

These findings are in line with other empirical studies showing that the I/E model is robust against various student characteristics (Marsh, Abduljabbar et al., 2015; Marsh & Hau, 2004; Marsh & Köller, 2004; Möller et al., 2009; Pinxten et al., 2015; Skaalvik & Rankin, 1990; Tay et al., 1995). Moreover, our findings match findings from previous studies demonstrating the validity of the I/E model for different ability groups [i.e., for high-ability students (Mueller & Winsor, 2016; Plucker & Stocking, 2001) as well as for students with learning disabilities (Möller, Streblow, & Pohlmann, 2009)] – however, the respective studies failed to compare different ability groups with each other. Finally, we replicated the findings of Jansen et al. (2015). Our study yet constituted a stricter test of invariance of the I/E model across students from different ability tracks since we contrasted two extreme groups (i.e., academic track and vocational track students). However, supplementary analyses (Tables S5 to S11 in the Online Supplements) demonstrated that invariance of the achievement–self-concept relations and thus the original and extended I/E models even holds when only comparing academic track students and students attending any other form of secondary school tracks, that is, when applying the approach by Jansen et al.. For these supplementary analyses we could rely on the total sample, that is, students from all German federal states participating in the BIJU study. Social and dimensional achievement comparisons can thus be seen as a valid psychological phenomenon that applies to different student samples (see also Möller & Husemann, 2006). Researchers might thus be allowed to combine different samples including students from different ability tracks in studies on the I/E model. This might be helpful for research practice since many educational systems, in particular secondary school systems, implement at least some kind of tracking or ability grouping (Chmielewski, Dumont, & Trautwein, 2013; LeTendre et al., 2003).

For DCT, the study offers new insight into the broad application of dimensional, along with social, comparisons for academic self-concept formation. In addition, students attending different ability tracks do not only seem to similarly apply social and dimensional comparisons in the formation of math and verbal self-concepts (see the original I/E model) but also in the formation of a wide range of domain-specific academic self-concepts. For applied practice, the findings indicate that for all students, high achievement levels in one domain might weaken the self-concept in another domain. Self-enhancement programs should thus be domain-specific in nature and should try to avoid potential adversary side effects such as negative effects of a math intervention on verbal self-concept (Gaspard et al., 2016; O'Mara, Marsh, Craven, & Debus, 2006).

8.2 Findings on the Original and Extended I/E Models

With regard to the total sample, the present study replicated well established findings on the original I/E model given positive achievement–self-concept relations within the math and verbal (German) domains and negative achievement–self-concept relations across the math and verbal domains. When extending the I/E model to five school subjects, contrast effects were found for relations between math or math-like (i.e., physics) achievement and self-concept measures on the one hand and verbal or verbal-like (i.e., English) achievement and self-concept measures on the other hand. As such, the findings showed negative relations between math achievement and English self-concept, between English achievement and math self-concept, and between German achievement and physics self-concept. English achievement was found to be negatively related to biology self-concept. Some studies found no contrast or assimilation effects involving biology self-concept, indicating that biology self-concept is neither verbal-like nor math-like (Jansen et al., 2015; Marsh, Lüdtke et al., 2015). Our study, instead, demonstrated a contrast effect for biology self-concept originating from English achievement as a verbal-like achievement measure, indicating that biology self-concept is a more math-like domain. However, given a found contrast effect between physics and biology, Guo et al. (2017) proposed a more verbal-like nature of biology self-concept. Hence, future research on biology self-concept is necessary. The specific characterization of biology self-concept as more math-like or verbal-like might also depend on students' school curriculum, that is, whether science is taught as an interdisciplinary subject (Jansen, Schroeders, Lüdtke, & Pand, 2014) and on students' perceived similarities of domains (Helm et al., 2016; Möller, Streblow, & Pohlmann, 2006).

English and German achievement and the non-corresponding self-concept measures were found to be unrelated to each other, so that our findings again reject an assimilation effect between verbal-like achievements and self-concepts (see also Marsh et al., 2001, 2014; Marsh, Lüdtke et al., 2015; Marsh & Yeung, 2001; Möller, Streblow, Pohlmann, & Köller, 2006; Niepel et al., 2014; Xu et al., 2013). Our findings could not support a significant assimilation effect between math and physics, although positive, but non-significant, relations were found between math and physics achievements and self-concepts. Another surprising finding relates to the significant positive relation between physics achievement and English self-concept. A negative contrast effect would have been assumed given the math-like nature of physics and the verbal-like nature of English. This significant finding disappeared when considering the total sample consisting of students from different German federal states of the BIJU study (Tables S5 to S11 in the Online Supplements). Hence, it might originate from specific characteristics of academic and vocational track students from North Rhine-Westphalia.

8.3 The Inclusion of Covariates

In all analyses including students' ability track, we controlled for students' gender, (verbal and numerical) intelligence, and SES as these background variables have been found to be associated with students' self-concept, achievement, and the allocation to different ability tracks. Our findings are in accord with findings from previous studies showing that boys have higher levels of math and physics self-concepts (Fredricks & Eccles, 2002; Jacobs et al., 2002; Schilling, Sparfeldt, & Rost, 2006). These gender differences related to students' self-concept were similarly found in the academic and vocational tracks.

Considering achievement, girls were found to have higher levels of verbal achievement in both tracks – a result that corresponds to previous findings (De Fraine et al., 2007; Van de gaer et al., 2006). Boys were found to have higher levels of math and physics achievements in both tracks. This finding on math achievement counters findings from previous studies showing no gender differences on math school grades as the math achievement indicator applied here (Arens et al, 2017; Marsh, Trautwein et al., 2005). Yet, findings on gender differences on math achievement seem to be inconsistent, to depend on the indicator of math achievement and the student sample considered, and to be generally small in size (Hyde et al., 1990; Nowell & Hedges, 1998). Verbal and numerical intelligence as well as SES were also found to be associated with students' domain-specific achievements, whereby more relations were found within the academic than in the vocational track. Hence, the effects of covariates partially differed between academic and vocational track students. Further research should address the varying effects of covariates on student achievement contingent upon students' ability track and learning environment.

8.4 Limitations and Directions for Future Research

In this study, we restricted our analyses to students from the German federal state of North Rhine-Westphalia. Only these students experienced the traditional allocation to academic, intermediate, comprehensive, and vocational tracks of secondary school allowing us to compare academic and vocational track students. Still, the sample of students was not representative across German federal states.

Our findings implied the invariance of the original and extended I/E models across students attending different school tracks. Hence, the assumption that academic and vocational track students might differ in their use of dimensional comparisons for self-concept formation has to be rejected. In other words, the negative stereotype associated with attending the vocational track does not seem to induce a stronger or lesser use of dimensional comparisons. Yet, the students were not asked for their felt deprivation and perception of the negative stereotype due to attendance of the vocational track.[10] Nor were the students directly asked for their use of social and dimensional comparisons for self-concept formation; the use of these comparison processes was only derived from the observation of achievement–self-concept relations. Hence, our study might motivate researchers to further detect variables which might impact on students' use of achievement comparison processes. Qualitative studies (Möller & Husemann, 2006) and experimental studies (Möller & Köller, 2001; Möller & Savyon, 2003; Pohlmann & Möller, 2009; Strickhouser & Zell, 2015) might help find out whether, when, to which extent, and why students apply social and dimensional comparisons for academic self-concept formation. Respective studies further aim to get insight into factors which might influence the use of social and dimensional comparisons for academic self-concept formation, and whether the factors differ for students attending different ability tracks.

The present study only realized a cross-sectional design and thus did not examine achievement–self-concept relations across time (Möller et al., 2011; Niepel et al., 2014). Differences in the application of social and dimensional comparisons between students from different ability tracks might however vary across students' grade levels. For instance, respective differences might be more pronounced at the time when students have just been allocated to different ability tracks, that is, when the vocational track students are fully aware of their allocation to a low-status and negatively stereotyped group rather than when they have become habituated to their learning environment (see also Arens & Watermann, 2015). Although the extended I/E model already takes five school subjects into account, the range of school subjects could be further broadened by, for instance, including multiple foreign languages or art-related school subjects (e.g. music and art; Vispoel, 1995). Moreover, the present findings have to be replicated when considering students from other educational systems and tracking types (Chmielewski et al., 2013).[11] The data analysed here were quite dated, but these data allowed us to contrast academic and vocational track students. Our findings thus do not automatically imply the invariance of the I/E model assumptions across more recent forms of tracking in the German educational system. For example, a recently implemented form of tracking is to merge the vocational and intermediate tracks. By this, the German school system aims to reduce environmental differences between the traditional school tracks (i.e., between the academic and non-academic tracks) and to overcome the traditionally strict boundaries between school tracks (Becker et al., 2016). Given that our findings showed that invariance regarding the I/E model assumptions even held when contrasting academic and vocational track students, invariance might also apply when comparing students who probably differ less regarding their experienced school environment. This conjecture is supported by the results of our supplementary analyses (Tables S5 to S11 in the Online Supplements) where we demonstrated invariance of the original and extended I/E models for academic track students and students attending any other form of secondary school tracks, that is, when pursuing a more coarse-grained approach.

Finally, only school grades were used as achievement indicators. Given the differential characteristics including achievement–self-concept relations of school grades and achievement test scores (Marsh et al., 2014; Marsh, Trautwein et al., 2005), future studies should replicate the present findings when using test scores as achievement indicators and when combining both types of achievement indicators (Möller et al., 2014). In sum, our study offers interesting findings on the generalizability of the original I/E model and its extension to multiple school subjects across students from different ability tracks. It might also present an incentive to further investigations on students' differences in the application of social and dimensional achievement comparison processes for academic self-concept formation.

9. Footnotes

[1] The I/E model only considers the individual student level where social comparisons lead to positive relations between achievements and self-concepts in matching domains. On the between-level, class-average or school-average achievement is negatively related to individual students' academic self-concepts as depicted in the big-fish-little-pond effect (BFLPE; Marsh et al., 2008). For integrating the BFLPE and the I/E model, see Parker, Marsh, Lüdtke, and Trautwein (2013).

[2] In preliminary analyses, we tested whether the variances of achievement (i.e., school grades) were similar across ability tracks in order to preclude that differences in the strength of achievement–self-concept relations (i.e., the I/E model assumptions) originate from different achievement variability. The Levene's tests of homogeneity of variances did not indicate differences in achievement (i.e., school grade) variances contingent upon students' ability tracks.

[3] Since Mplus excludes cases with missing data on covariates defined as exogenous variables, we allowed covariances among the continuous covariates, that is, verbal intelligence, numerical intelligence, and SES. In this case, FIML is also applied to handle missing data on these covariates. However, with regard to gender as a dichotomous covariate, students with missing information on gender ($N = 11$ in the subsample of students from North Rhine-

Westphalia) were deleted from the multi-group analyses where gender was included as a covariate.

[4] The resulting factor correlations are depicted in Table S1 of the Online Supplements.

[5] This model is statistically equivalent to the CFA model (Model 1 in Table 1) assuming separate factors for math and German achievements and self-concepts since the factor correlations were replaced by path coefficients.

[6] This model is statistically equivalent to Model 4 (i.e., the CFA model with math and German achievement and self-concept factors stating factor loading invariance across academic and vocational track students), since the factor correlations were replaced by path coefficients.

[7] The resulting factor correlations are depicted in Table S3 of the Online Supplements.

[8] This model is statistically equivalent to the CFA model (Model 9 in Table 1) assuming separate factors for math, German, English, biology, and physics achievements and self-concepts, since the factor correlations were replaced by path coefficients.

[9] This model is statistically equivalent to Model 12 (i.e., the CFA model with achievement and self-concept factors for math, German, English, biology, and physics stating factor loading invariance across academic and vocational track students), since the factor correlations were replaced by path coefficients.

[10] Yet, academic and vocational track students were found to differ in their achievement levels. Vocational track students demonstrated lower mean levels of achievement when considering standardized achievement tests in math [t (1206) = -27.729, $p < .05$], English [t (1106) = -40.359, $p < .05$]), biology [t (1211) = -25,043, $p < .05$], and physics [t (1219 = -15.604, $p < .05$].

[11] In additional analyses, we further tested whether the patterns of achievement–self-concept relations were invariant across gender. For this purpose, we conducted models with a grouping variable consisting of four groups: girls attending the academic track, boys attending the academic track, girls attending the vocational track, and girls attending the vocational

track. As can be seen from Table S12 in the Online Supplements, the results supported invariance both when considering the original I/E model and when considering the I/E model extended to five school subjects.

10. References

Amthauer, R. (1955). I-S-T. *Intelligenz-Struktur-Test: Handanweisung für die Durchführung und Auswertung [Intelligence Structure Test: Manual for administration and analysis] (2nd ed.).* Göttingen, Germany: Hogrefe.

Arens, A.K., Marsh, H.W., Pekrun, R., Lichtenfeld, S., Murayama, K., & vom Hofe, R. (2017). Math self-concept, grades, and achievement test scores: Long-term reciprocal effects across five waves and three achievement tracks. *Journal of Educational Psychology, 109*, 621–634. http://doi.org/10.1037/edu0000163

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. http://dx.doi: 10.1016/j.lindif.2016.08.044

Arens, A.K., & Watermann, R. (2015). How an early transition to high-ability secondary schools affects students' academic self-concept: Contrast effects, assimilation effects, and differential stability. *Learning and Individual Differences, 37,* 64–71. https://doi.org/10.1016/j.lindif.2014.11.007

Baumert, J., Roeder, P. M., Gruehn, S., Heyn, S., Köller, O., ... Rimmele, R. (1996). Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU) [Educational careers and psychosocial development in adolescence and young adulthood]. In K.-P. Treumann, G. Neubauer, R. Möller, & J. Abel (Eds.). *Methoden und Anwendungen empirischer pädagogischer Forschung* (pp. 170–180). Münster: Waxmann.

Baumert, J., Watermann, R., & Schümer, G. (2003). Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs. Ein institutionelles und individuelles Mediationsmodell

[Disparities in educational participation and attainment: An institutional and individual mediation model]. Zeitschrift für Erziehungswissenschaft, 6, 46–71. https://doi.org/10.1007/s11618-003-0004-7

Beauducel, A., & Herzberg P.Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling, 1,* 186–203. https://doi.org/10.1207/s15328007sem1302_2

Becker, M., Neumann, M., & Dumont, H. (2016). Recent developments in school tracking practices in Germany: An overview and outlook on future trends. *Orbis Scholae,10*, 1–17. https://doi.org/10.14712/23363177.2017.8

Boaler, J., William, D., & Brown, M. (2000). Students' experiences of ability grouping. Disaffection, polarisation and the construction of failure. *British Educational Research Journal, 26*, 631–646. https://doi.org/10.1080/713651583

Bradley, R.H., & Corwyn, R.F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*, 371–399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.

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

Brunner, M., Krauss, S., & Kunter, M. (2008). Gender differences in mathematics: Does the story need to be rewritten? *Intelligence, 36*, 403–421. https://doi.org/10.1016/j.intell.2007.11.002

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

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

Chmielewski, A.K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal, 50*, 925–957. https://doi.org/10.3102/0002831213489843

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

Dickhäuser, O., Reuter, M., & Hilling, C. (2005). Coursework selection: A frame of reference approach using structural equation modelling. *British Journal of Educational Psychology, 75*, 673–688. https://doi.org/10.1348/000709905X37181

Ditton, H. (2013). Wer geht auf die Hauptschule? Primäre und sekundäre Effekte der sozialen Herkunft beim Übergang nach der Grundschule [Who goes to the German „Hauptschule" (the lower ability band secondary school)? Primary and secondary effects of social origin in the transition to secondary school]. *Zeitschrift für Erziehungswissenschaft, 16*, 731–749. https://doi.org/10.1007/s11618-013-0440-y

Dumont, H., Protsch, P., Jansen, M., & Becker, M. (2017). Fish swimming into the ocean: How tracking relates to students' self-beliefs and school disengagement at the end of schooling. *Journal of Educational Psychology, 109*, 855–870. http://dx.doi.org/10.1037/edu0000175

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

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140. https://doi.org/10.1177/001872675400700202

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

Frey, M.C., & Detterman, D.K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. P*sychological Science, 15*, 373–378.

Gamoran, A., & Carbonaro, W. (2002). High school English: A national portrait. *High School Journal, 86*, 1–13. https://doi.org/10.1353/hsj.2002.0021

Gamoran, A., Nystrand, M., Berends, M., & LePore, P.C. (1995). An organizational analysis of the effects of ability grouping. *American Educational Research Journal, 32*, 687–715. https://doi.org/10.3102/00028312032004687

Ganzeboom, H.B.G, & Treiman, D.J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research, 25,* 201–239. https://doi.org/10.1006/ssre.1996.0010

Gaspard, H., Dicke, A.-L., Flunger, B., Häfner, I., Brisson, B.M., Trautwein, U., & Nagengast, B. (2016). Side effects of motivational interventions? Effects of an intervention in math classrooms on motivation in verbal domains. *AERA Open, 2*. https://doi.org/10.1177/2332858416649168

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

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

Hallam, S., & Ireson, J. (2007). Secondary school pupils' satisfaction with their ability grouping placements. *British Educational Research Journal, 33*, 27–45. http://www.jstor.org/stable/30032722

Heller, K.A., Schoen-Gaedike, A.-K., & Weinlaeder, H. (1985). *Kognitiver Fähigkeitstest: KFT 4-13+ [Cognitive Ability Test: KFT 4−13+] (2nd ed.).* Weinheim, Germany: Beltz.

Helm, F., Müller-Kalthoff, H., Nagy, N., & Möller, J. (2016). Dimensional comparison: theory Perceived subject similarity impacts on students' self-concepts. *AERA Open, 2,* 1–9. https://doi.org/10.1177/2332858416650624

Hu, L.T., & Bentler, P.M. (1999). Cutoff 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

Hyde, J.S., Fennema, E., & Lamon, S.L. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139–155. http://dx.doi.org/10.1037/0033-2909.107.2.139

Ireson, J., & Hallam, S. (1999). Raising standards: Is ability grouping the answer? *Oxford Review of Education, 25,* 343–358. https://doi.org/10.1080/030549899104026

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., & Pant, H.A. (2014). Interdisziplinäre Beschulung und die Struktur des akademischen Selbstkonzepts in den naturwissenschaftlichen Fächern [Interdisciplinary science classes and the dimensional structure of academic self-concept in the sciences]. *Zeitschrift für Pädagogische Psychologie, 28*, 43–49. https://doi.org/10.1024/1010-0652/a000120

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. https://doi.org/10.1037/edu0000021

Jerusalem, M. (1984). *Selbstbezogene Kognitionen in schulischen Bezugsgruppen. Eine Längsschnittstudie* [Self-related cognitions in school reference groups: A longitudinal study]. Berlin, Germany: Free University Press.

Jones, J.D., Vanfossen, B.E, & Ensminger, M.E. (1995). Individual and organizational predictors of high school track placement. *Sociology of Education, 68,* 287–300. Retrieved from http://www.jstor.org/stable/2112743

Jopt, U.-J. (1978). *Selbstkonzept und Ursachenerklärung in der Schule. Zur Attribuierung von Schulleistungen* [Self-concept and causal attributions in schools. On attributions of school achievement]. Bochum, Germany.

Kline, R.B. (2005). *Principles and practice of structural equation modeling.* New York: Guildford.

Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Teachers' occupational well-being and quality of instruction: The important role of self-regulatory patterns. *Journal of Educational Psychology, 100,* 702–715. https://doi.org/10.1037/0022-0663.100.3.702

Knigge, M., & Hannover, B. (2011). Collective school-type identity: Predicting students' motivation beyond academic self-concept. *International Journal of Psychology, 46*, 191–205. https://doi.org/10.1080/00207594.2010.529907.

LeTendre, G.K., Hofer, B.K., & Shimizu, H. (2003). What is tracking? Cultural expectations in the United States, Germany, and Japan. *American Educational Research Journal, 40*, 43–89. https://doi.org/10.3102/00028312040001043

Liem, G.A.D., Marsh, H.W., Martin, A.J., McInerney, D.M., & Yeung, A.S. (2013). The big-fish-little-pond effect and a national policy of within-school ability streaming: Alternative frames of reference. *American Educational Research Journal, 50*, 326–370. https://doi.org/10.3102/0002831212464511

Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives, 2*, 99–106. https://doi.org/10.1111/j.1750-8606.2008.00048.x

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. (1990a). Influences of internal and external frames of reference on the formation of math and English self-concepts. *Journal of Educational Psychology, 82*, 107–116. http://dx.doi.org/10.1037/0022-0663.82.1.107

Marsh, H.W. (1990b). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology, 82*, 623–636.

Marsh, H.W., Abduljabbar, A.S., Abu-Hilal, M.M., Morin, A.J.S., Abdelfattah, F., Leung, K. C., . . . Parker, P. (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. https://doi.org/10.1037/a0029907

Marsh, H.W., Abduljabbar, A.S., Parker, P., Morin, A.J.S., Abdelfattah, F., Nagengast, B., . . . Abu-Hillal, M.M. (2015). The internal/external frame of reference model: Age-cohort and cross-national differences in paradoxical relations between TIMSS mathematics and science achievement, self-concept and intrinsic motivation. *American Educational Research Journal, 52,* 168–202. https://doi.org/10.3102/0002831214549453

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

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

Marsh, H.W., Hau, K.T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares, & J.J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 275–340). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

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

Marsh, H.W., Kong, C.K., & Hau, K. (2001). Extension of the internal/external frame of reference model of self-concept formation: Importance of native and nonnative languages for Chinese students. *Journal of Educational Psychology, 93*, 543–553. 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., 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

Marsh, H.W., & Yeung, A.S. (2001). An extension of the internal/external frame of reference model: A response to Bong (1998). *Multivariate Behavioral Research, 36*, 389–420. . https://doi.org/10.1207/S15327906389-420.

Matteucci, M., & Mignani, S. (2011). Gender differences in performance in mathematics at the end of lower secondary school in Italy. *Learning and Individual Differences, 21*, 543–548. http://dx.doi.org/10.1016/j.lindif.2011.03.001

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

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., & 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.v4i2.169

Möller, J., & Pohlmann, B. (2010). Achievement differences and self-concept differences: Stronger associations for above or below average students? *British Journal of Educational Psychology, 80*, 435–450. https://doi.org/10.1348/000709909X485234

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., & Savyon, K. (2003). Not very smart thus moral: Dimensional comparisons between academic self-concept and honesty. *Social Psychology of Education, 6,* 95–106. https://doi.org/10.1023/A:1023247910033

Möller, J., Streblow, L., & Pohlmann, B. (2006).The belief in a negative interdependence of math and verbal abilities as determinant of academic self-concepts. *British Journal of Educational Psychology, 76,* 57–70. https://doi.org/10.1348/000709905X37451

Möller, J., Streblow, L., & Pohlmann, B. (2009). Achievement and self-concept of students with learning disabilities. *Social Psychology of Education, 12,* 113–122. https://doi.org/10.1007/s11218-008-9065-z

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

Möller, J., Zimmermann, F., & Köller, O. (2014). The reciprocal internal/external frame of reference model using grades and test scores. *British Journal of Educational Psychology, 84,* 591–611. https://doi.org/10.1111/bjep.12047

Mueller, C.E., & Winsor, D.L. (2016). Math and verbal academic self-concept: Subject-specificity across four distinctive groups of high ability adolescents. *Learning and Individual Differences, 50*, 240–245. https://doi.org/10.1016/j.lindif.2016.08.002

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

Nagengast, B., & Marsh, H.W. (2012). Big fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology, 104*, 1033–1053. https://doi.org/10.1037/a0027697

Nagy, G., Trautwein, U., Köller, O., Baumert, J., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation, 12,* 323–345. https://doi.org/10.1080/13803610600765687

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

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

Parker, P. D., Marsh, H. W., Lüdtke, O., & Trautwein, U. (2013). Differential school contextual effects for math and English: Integrating the big-fish-little-pond effect and the internal/external frame of reference. *Learning and Instruction*, *23*, 78–89. https://doi.org/10.1016/j.learninstruc.2012.07.001

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. https://doi.org/10.1016/j.cedpsych.2014.12.003

Plucker, J.A., & Stocking, V.B. (2001). Looking outside and inside: Self-concept development of gifted adolescents. *Exceptional Children, 67*, 535–548. https://doi.org/10.1177/001440290106700407

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

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd Ed.).* Thousand Oaks, CA: Sage.

Rhode, T.E., & Thompson, L.A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*, 83–92. https://doi.org/10.1016/j.intell.2006.05.004

Schilling, S.R., Sparfeldt, J.R., & Rost, D.H. (2006). Facetten schulischen Selbstkonzepts: Welchen Unterschied macht das Geschlecht? [Sex differences in subject-specific academic self-concepts]. *Zeitschrift für Pädagogische Psychologie, 20*, 9–18. http://dx.doi.org/10.1024/1010-0652.20.12.9.

Sirin, S.R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*, 417–453. https://doi.org/10.3102/00346543075003417

Skaalvik, E.M., & Rankin, R.J. (1990). Math, verbal, and general academic self-concept: The internal/external frame of reference model and gender differences in self-concept structure. *Journal of Educational Psychology, 82*, 546–554. http://dx.doi.org/10.1037/0022-0663.82.3.546

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. https://doi.org/10.3102/00028312032001161

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

Solga, H. (2004). Increasing risks of stigmatization: Changes in school-to-work transitions of less-educated West Germans. *Yale Journal of Sociology, 4*, 99–129. Retrieved from http://www.yale.edu/socdept/yjs/YJSFALL2004.pdf.

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

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

Thorndike, R.L., & Hagen, E. (1971). *Cognitive abilities test.* Boston: Houghton-Mifflin.

Trautwein, U., Lüdtke, O., Marsh, H.W., & Nagy, G. (2009). Within-school social comparison: How students perceive the standing of their class predicts academic self-concept. *Journal of Educational Psychology, 101,* 853–866. https://doi.org/10.1037/a0016306

Vispoel, W.P. (1995). Self-concept in artistic domains: An extension of the Shavelson, Hubner, and Stanton (1976) model. *Journal of Educational Psychology, 87,* 134–153. http://dx.doi.org/10.1037/0022-0663.87.1.134

Van de gaer, E., Pustjens, H., van Damme, J., & De Munter, A. (2006). The gender gap in language achievement: The role of school-related attitudes of class groups. *Sex Roles, 55,* 397–408. https://doi.org/10.1007/s11199-006-9092-1

Van Houtte, M. (2004). Tracking effects on school achievement: A quantitative explanation in terms of the academic culture of school staff. *American Journal of Education, 110*, 354–388. Retrieved from http://www.jstor.org/stable/10.1086/422790

Van Houtte, M. (2006). School type and academic culture: Quantitative evidence for the differentiation-polarisation theory. *Journal of Curriculum Studies, 38*, 273–292. https://doi.org/10.1080/00220270500363661

Van Houtte, M., & Stevens, P.A.J. (2009). Study involvement of academic and vocational students: Does between-school tracking sharpen the difference? *American Educational Research Journal, 46*, 943–973. https://doi.org/10.3102/0002831209348789

Van Houtte, M., & Stevens, P.A.J (2010). The culture of futility and its impact on study culture in technical/vocational schools in Belgium. *Oxford Review of Education, 36*, 23–43. https://doi.org/10.1080/03054980903481564

Van Praag, L., Demanet, J., Stevens, P.A.J., & Van Houtte, M. (2017). ‘Everyone has their own qualities’: Tracking and academic self-appraisal in Flemish secondary education. *Social Psychology of Education, 20*, 601–618. https://doi.org/10.1007/s11218-017-9371-4

Wolff, F., Helm, F., Zimmermann, F., Nagy, G., & Möller, J. (2018). On the effects of social, temporal, and dimensional comparisons on academic self-concept. *Journal of Educational Psychology.* Advance online publication. https://doi.org/10.1037/edu0000248

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

Table 1

Goodness-of-fit Indices with Students from the Federal State of North Rhine-Westphalia

	χ^2	df	CFI	TLI	RMSEA	SRMR	
Original I/E Model							
	Total Sample Analyses						
1	153.007	45	.976	.965	.044	.023	CFA model with separate factors for math and German self-concepts and achievements
2	153.007	45	.976	.965	.044	.023	I/E model with math and German self-concepts and achievements
	Analyses with Academic and Vocational Track Students as a Grouping Variable						
3	271.030	160	.981	.973	.034	.025	CFA model with separate factors for math and German self-concepts and achievements; configural invariance across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
4	290.255	168	.979	.972	.035	.033	CFA model with separate factors for math and German self-concepts and achievements; factor loading invariance across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
5	290.254	168	.979	.972	.035	.033	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; free estimation of the I/E model across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
6	312.158	170	.976	.968	.037	.043	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; invariance of the within-domain achievement–self-concept relations across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
7	303.931	170	.977	.970	.036	.038	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; invariance of the between-domain achievement–self-concept relations across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
8	323.386	172	.974	.966	.038	.043	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; invariant estimation of the I/E model across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates

(continued)

Table 1 (continued)

	χ^2	df	CFI	TLI	RMSEA	SRMR	
Extended I/E Model							
	Total Sample Analyses						
9	660.955	315	.973	.963	.030	.031	CFA model with separate factors for math, German, English, biology, and physics self-concepts and achievements
10	660.955	315	.973	.963	.030	.031	I/E model with math, German, English, biology, and physics self-concepts and achievements
	Analyses with Academic and Vocational Track Students as a Grouping Variable						
11	1286.871	796	.968	.955	.032	.036	CFA model with separate factors for math, German, English, biology, and physics self-concepts and achievements; configural invariance across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
12	1328.431	816	.966	.954	.032	.039	CFA model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
13	1328.432	816	.966	.954	.032	.039	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; free estimation of the extended I/E model across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
14	1354.408	821	.965	.952	.033	.041	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; invariance of the within-domain achievement–self-concept relations across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
15	1369.825	836	.965	.953	.032	.042	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; invariance of the between-domain achievement–self-concept relations across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
16	1395.067	841	.963	.952	.033	.044	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; invariant estimation of the extended I/E model across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates

Note. All models are estimated with the Robust Maximum Likelihood (MLR) estimator; all χ^2 are significant ($p < .05$). CFA = confirmatory factor analyses; CFI = comparative fit index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

Table 2

Standardized Path Coefficients of the Original I/E Models with Students from the Federal State of North Rhine-Westphalia

	Model 2	Model 8	
Original I/E Model	**β**	**β AT**	**β VT**
German achievement → German self-concept	.326*	.348*	.304*
Math achievement → German self-concept	-.103*	-.122*	-.111*
German achievement → Math self-concept	-.129*	-.117*	-.101*
Math achievement → Math self-concept	.489*	.457*	.412*
Relations of Covariates to Self-concept			
Verbal intelligence → German self-concept	-	.018	-.006
Numerical intelligence → German self-concept	-	.052	.027
SES → German self-concept	-	.011	.027
Gender → German self-concept	-	.016	.023
Verbal intelligence → Math self-concept	-	-.042	.024
Numerical intelligence → Math self-concept	-	.162*	.078
SES → Math self-concept	-	.054	.040
Gender → Math self-concept	-	.192*	.162*
Relations of Covariates to Achievement			
Verbal intelligence → German achievement	-	.250*	.103*
Numerical intelligence → German achievement	-	.137*	.056
SES → German achievement	-	.131*	.079
Gender → German achievement	-	-.151*	-.117*
Verbal intelligence → Math achievement	-	.266*	.066
Numerical intelligence → Math achievement	-	.148*	.083
SES → Math achievement	-	.147*	.097*
Gender → Math achievement	-	.110*	.119*

Note. AT = academic track; VT = vocational track.
* $p < .05$.

Table 3

Standardized Path Coefficients of the Extended I/E Models with Students from the Federal State of North Rhine-Westphalia

	Model 10	Model 16	
Extended I/E Model	**β**	**β AT**	**β VT**
German achievement → German self-concept	.294*	.316*	.273*
Math achievement → German self-concept	-.137*	-.157*	-.141*
English achievement → German self-concept	.023	.004	.004
Biology achievement → German self-concept	.022	.029	.027
Physics achievement → German self-concept	.074	.099	.097
German achievement → Math self-concept	-.107*	-.085*	-.073*
Math achievement → Math self-concept	.486*	.473*	.426*
English achievement → Math self-concept	-.094*	-.113*	-.098*
Biology achievement → Math self-concept	-.012	.017	.016
Physics achievement → Math self-concept	.099	.038	.037
German achievement → English self-concept	.040	.042	.036
Math achievement → English self-concept	-.218*	-.230*	-.210 *
English achievement → English self-concept	.549*	.582*	.511*
Biology achievement → English self-concept	-.047	-.053	-.050
Physics achievement → English self-concept	.118*	.100	.099
German achievement → Biology self-concept	-.011	-.025	-.020
Math achievement → Biology self-concept	-.058	-.068	-.058
English achievement → Biology self-concept	-.101*	-.118*	-.097*
Biology achievement → Biology self-concept	.319*	.346*	.308*
Physics achievement → Biology self-concept	.022	-.023	-.021
German achievement → Physics self-concept	-.119*	-.076	-.071
Math achievement → Physics self-concept	.024	-.005	-.005
English achievement → Physics self-concept	-.004	-.037	-.035
Biology achievement → Physics self-concept	-.030	.001	.001
Physics achievement → Physics self-concept	.307*	.242*	.256*
Relations of Covariates to Self-concept			
Verbal intelligence → German self-concept	-	.007	-.009
Numerical intelligence → German self-concept	-	.053	.032
SES → German self-concept	-	.028	.017
Gender → German self-concept	-	-.013	.005
Verbal intelligence → Math self-concept	-	-.043	.029
Numerical intelligence → Math self-concept	-	.169*	.071
SES → Math self-concept	-	.037	.016
Gender → Math self-concept	-	.191*	.150*
Verbal intelligence → English self-concept	-	.014	-.078
Numerical intelligence → English self-concept	-	-.010	.035
SES → English self-concept	-	.065	.056
Gender → English self-concept	-	.078	.043
Verbal intelligence → Biology self-concept	-	.063	.072
Numerical intelligence → Biology self-concept	-	-.054	.076

(continued)

Table 3 (continued)

	Model 10	Model 16	
	β	**β AT**	**β VT**
SES → Biology self-concept	-	.025	-.056
Gender → Biology self-concept	-	.015	-.003
Verbal intelligence → Physics self-concept	-	.050	.007
Numerical intelligence → Physics self-concept	-	.014	-.002
SES → Physics self-concept	-	.077	-.068
Gender → Physics self-concept	-	.174*	.149*
Relations of Covariates to Achievement			
Verbal intelligence → German achievement	-	.251*	.104*
Numerical intelligence → German achievement	-	.136*	.056
SES → German achievement	-	.131*	.074
Gender → German achievement	-	-.153*	-.118*
Verbal intelligence → Math achievement	-	.265*	.068
Numerical intelligence → Math achievement	-	.145*	.083
SES → Math achievement	-	.152*	.100*
Gender → Math achievement	-	.110*	.118*
Verbal intelligence → English achievement	-	.189*	.086
Numerical intelligence → English achievement	-	.145*	-.022
SES → English achievement	-	.006	-.030
Gender → English achievement	-	.020	-.053
Verbal intelligence → Biology achievement	-	.178*	.105*
Numerical intelligence → Biology achievement	-	.092*	.003
SES → Biology achievement	-	.122*	.079
Gender → Biology achievement	-	-.065	.004
Verbal intelligence → Physics achievement	-	.243*	.054
Numerical intelligence → Physics achievement	-	.023	-.027
SES → Physics achievement	-	-.087	.142*
Gender → Physics achievement	-	.244*	.183*

Note. AT = academic track; VT = vocational track.
* $p < .05$.

Online Supplements to "The internal/external frame of reference (I/E) model: Extension to five school subjects and invariance across German secondary school ability tracks"

Table S1

Factor Correlations of Model 1 (Table 1 in the Main Manuscript)

	Math self-concept	German self-concept	Math achievement
German self-concept	.269*		
Math achievement	.432*	.042	
German achievement	.089*	.280*	.446*

Note. * $p < .05$.

Table S2

Factor Correlations of the Original I/E Models with Students from the Federal State of North Rhine-Westphalia

	Math self-concept	Math achievement
German self-concept	.333*/.268 */.403*	
German achievement		.446*/.457*/.350*

Note. The first coefficient refers to Model 2 of Table 1 in the main manuscript (the original I/E model with the total sample), the second and third coefficients refer to Model 8 of Table 1 in the main manuscript (the original I/E model with a grouping variable consisting of academic track students and vocational track students). Thereby, the second coefficient depicts the factor correlations for academic track students; the third coefficient depicts the factor correlations for vocational track students of Model 8.
* $p < .05$.

Table S3

Factor Correlations of Model 9 (Table 1 in the Main Manuscript)

	Math sc	German sc	English sc	Physics sc	Biology sc	Math ach	German ach	English ach	Physics ach
German sc	.267*								
English sc	.160*	.348*							
Physics sc	.338*	.315*	.213*						
Biology sc	.288*	.400*	.191*	.424*					
Math ach	.434*	.042	.057	.078*	.038				
German ac	.093*	.282*	.250*	-.009	.059*	.446*			
English ach	.077*	.145*	.495*	.018	-.019	.418*	.517*		
Physics ach	.221*	.146*	.172*	.257*	.122*	.396*	.371*	.270*	
Biology ach	.163*	.137*	.120*	.076*	.266*	.421*	.441*	.334*	.486*

Note. sc = self-concept. ach = achievement.
* $p < .05$.

Table S4

Factor Correlations of the Extended I/E Models with Students from the Federal State of North Rhine-Westphalia

	Math sc	German sc	English sc	Physics sc	Math ach	German ach	English ach	Physics ach
German sc	.329*/.264*/.412*							
English sc	.254*/.111*/.406*	.317*/.301*/.370*						
Physics sc	.324*/.313*/.308*	.337*/.344*/.342*	.254*/.133*/.338*					
Biology sc	.313*/.259*/.405*	.410*/.404*/.463*	.254*/.156*/.322*	.436*/.477*/.413*				
German ach					.446*/.456*/.350*			
English ach					.418*/.384*/.404*	.517*/.536*/.435*		
Physics ach					.396*/.296*/.364*	.371*/.333*/.379*	.270*/.325*/.232*	
Biology ach					.421*/.371*/.382*	.441*/.370*/.419*	.334*/.249*/.392*	.486*/.442*/.479*

Note. The first coefficient refers to Model 10 of Table 1 in the main manuscript (the extended I/E model with the total sample of students from the federal state of North Rhine-Westphalia), the second and third coefficients refer to Model 16 of Table 1 in the main manuscript (the extended I/E model with a grouping variable consisting of academic track students and vocational track students). Thereby, the second coefficient depicts the factor correlations for academic track students; the third coefficient depicts the factor correlations for vocational track students of Model 16.
sc = self-concept; ach = achievement.
* $p < .05$.

Table S5

Goodness-of-fit Indices for the Models with Students from all Federal States

	χ^2	df	CFI	TLI	RMSEA	SRMR	
Original I/E Model							
	Total Sample Analyses						
S1	311.069	45	.983	.975	.038	.025	CFA model with separate factors for math and German self-concepts and achievements
S2	311.069	45	.983	.975	.038	.025	I/E model with math and German self-concepts and achievements
	Analyses with Students' Ability Tracks as a Grouping Variable						
S3	502.383	160	.981	.974	.033	.024	CFA model with separate factors for math and German self-concepts and achievements; configural invariance across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
S4	553.934	168	.979	.972	.034	.032	CFA model with separate factors for math and German self-concepts and achievements; factor loading invariance across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
S5	553.934	168	.979	.972	.034	.032	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; free estimation of I/E Model across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
S6	576.347	170	.978	.970	.035	.036	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; invariance of the within-domain achievement–self-concept relations across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
S7	566.156	170	.978	.971	.034	.034	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; invariance of the between-domain achievement–self-concept relations across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
S8	586.617	172	0.977	0.970	0.035	0.037	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; invariant estimation of I/E Model across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates

(continued)

Table S5 (continued)

	χ^2	df	CFI	TLI	RMSEA	SRMR	
Extended I/E Model							
	Total Sample Analyses						
S9	1263.903	315	.976	.967	.027	.030	CFA model with separate factors for math, German, English, biology, and physics self-concepts and achievements
S10	1263.903	315	.976	.967	.027	.030	I/E model with for math, German, English, biology, and physics self-concepts and achievements
	Analyses with Students' Ability Tracks as a Grouping Variable						
S11	1887.957	796	.975	.965	.026	.030	CFA model with separate factors for math, German, English, biology, and physics self-concepts and achievements; configural invariance across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
S12	1970.237	816	.973	.964	.027	.032	CFA model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
S13	1970.237	816	.973	.964	.027	.032	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; free estimation of the extended I/E Model across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
S14	1983.285	821	.973	.964	.027	.033	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; invariance of the within-domain achievement–self-concept relations across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates
S15	2007.697	836	.973	.964	.027	.035	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; invariance of the between-domain achievement–self-concept relations across ability track groups, SES, verbal intelligence, and numerical intelligence as covariates
S16	2034.413	841	.972	.963	.027	.035	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; invariant estimation of the extended I/E Model across ability track groups; gender, SES, verbal intelligence, and numerical intelligence as covariates

Note. All models are estimated with the Robust Maximum Likelihood (MLR) estimator; all χ^2 are significant ($p < .05$). Since Mplus excludes cases with missing values on any covariates defined as exogenous variable only, covariances among the continuous covariates (i.e., verbal intelligence, numerical intelligence, and SES) were allowed. FIML was then used to handle missing data on these variables. Still, for $N = 60$ students, no information was available regarding gender as a dichotomous covariate. These students were dropped from the multi-group analyses using ability tracks as a grouping variable so that the sample size analyzed was $N = 3941$ in this case. CFA = confirmatory factor analyses; CFI = comparative fit index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

Table S6

Factor Correlations of Model S1 in Table S5

	Math self-concept	German self-concept	Math achievement
German self-concept	.172*		
Math achievement	.401*	.040*	
German achievement	.018	.358*	.414*

Note. * $p < .05$.

Table S7

Standardized Path Coefficients of the Original I/E Models with Students from all Federal States

	Model S2	**Model S8**	
Original I/E Model	**β**	**β AT**	**Β NAT**
German achievement → German self-concept	.412*	.445*	.398*
Math achievement → German self-concept	-.131*	-.149*	-.139*
German achievement → Math self-concept	-.179*	-.153*	-.140*
Math achievement → Math self-concept	.475*	.442*	.420*
Relations of Covariates to Self-concept			
Verbal intelligence → German self-concept	-	.004	-.006
Numerical intelligence → German self-concept	-	.061*	-.024
SES → German self-concept	-	-.007	.008
Gender → German self-concept	-	.027	.008
Verbal intelligence → Math self-concept	-	-.011	.002
Numerical intelligence → Math self-concept	-	.162*	.104*
SES → Math self-concept	-	.004	-.001
Gender → Math self-concept	-	.173*	.164*
Relations of Covariates to Achievement			
Verbal intelligence → German achievement	-	.207*	.115*
Numerical intelligence → German achievement	-	.115*	.010
SES → German achievement	-	.164*	.114*
Gender → German achievement	-	-.153*	-.201*
Verbal intelligence → Math achievement	-	.236*	.069*
Numerical intelligence → Math achievement	-	.134*	.108*
SES → Math achievement	-	.085*	.077*
Gender → Math achievement	-	.114 *	.086*

Note. AT = academic track; NT = non-academic tracks.
* $p < .05$.

Table S8

Factor Correlations of the Original I/E Models with Students from all Federal States

	Math self-concept	Math achievement
German self-concept	.260*/.273*/.260*	
German achievement		.414*/.431*/.391*

Note. The first coefficient refers to Model S2 in Table S5 (the original I/E model with the total sample of students from all federal states), the second and third coefficients refer to Model S8 in Table S5 (the original I/E model with a grouping variable consisting of academic track students and non-academic track students). Thereby, the second coefficient depicts the factor correlations for academic track students; the third coefficient depicts the factor correlations for non-academic track students of Model S8.

* $p < .05$.

Table S9

Factor Correlations of Model S9 in Table S5

	Math sc	German sc	English sc	Physics sc	Biology sc	Math ach	German ach	English ach	Physics ach
German sc	.171*								
English sc	.124*	.327*							
Physics sc	.375 *	.239*	.151*						
Biology sc	.262*	.339*	.140*	.377*					
Math ach	.403*	.041*	.048*	.100*	.017				
German ach	.021	.359*	.253*	-.004	.056*	.416*			
English ach	.040*	.174*	.471*	.015	-.030	.430*	.550*		
Physics ach	.224*	.110*	.100*	.274*	.141*	.466*	.370*	.344*	
Biology ach	.131*	.126*	.118*	.100*	.290*	.402*	.400*	.363*	.502*

Note. sc = self-concept; ach = achievement.
* $p < .05$.

Table S10

Standardized Path Coefficients of the Extended I/E Model with Students from all Federal States

	Model S2	**Model S18**	
Extended I/E model		**β AT**	**β NAT**
German achievement → German self-concept	.406*	.439*	.392*
Math achievement → German self-concept	-.141*	-.159*	-.147*
English achievement → German self-concept	.001	-.004	-.003
Biology achievement → German self-concept	.010	.014	.014
Physics achievement → German self-concept	.020	.020	.020
German achievement → Math self-concept	-.145*	-.120*	-.109*
Math achievement → Math self-concept	.468*	.440*	.417*
English achievement → Math self-concept	-.113*	-.116*	-.106*
Biology achievement → Math self-concept	-.010	-.004	-.004
Physics achievement → Math self-concept	.104*	.093*	.092*
German achievement → English self-concept	.046	.078*	.069 *
Math achievement → English self-concept	-.193*	-.216*	-.197*
English achievement → English self-concept	.536*	.581*	.512*
Biology achievement → English self-concept	-.016	-.028	-.027
Physics achievement → English self-concept	-.003	-.015	-.014
German achievement → Biology self-concept	.014	.015	.013
Math achievement → Biology self-concept	-.092*	-.114 *	-.100*
English achievement → Biology self-concept	-.143*	-.170 *	-.144*
Biology achievement → Biology self-concept	.346*	.355*	.335*
Physics achievement → Biology self-concept	.054	.034	.031
German achievement → Physics self-concept	-.102*	-.057	-.053
Math achievement → Physics self-concept	.014	-.032	-.031
English achievement → Physics self-concept	-.042	-.055	-.051
Biology achievement → Physics self-concept	-.014	.009	.009
Physics achievement → Physics self-concept	.327*	.282*	.282*

(continued)

Table S10 (continued)

Relations of covariates to self-concept			
Verbal intelligence → German self-concept	-	.004	-.007
Numerical intelligence → German self-concept	-	.058	-.024
SES → German self-concept	-	-.005	.006
Gender → German self-concept	-	.023	.008
Verbal intelligence → Math self-concept	-	-.009	-.001
Numerical intelligence → Math self-concept	-	.161*	.100*
SES → Math self-concept	-	-.004	.000
Gender → Math self-concept	-	.166*	.152*
Verbal intelligence → English self-concept	-	.009	.010
Numerical intelligence → English self-concept	-	.007	.043*
SES → English self-concept	-	-.013	-.007
Gender → English self-concept	-	.118*	.087*
Verbal intelligence → Biology self-concept	-	.062*	.029
Numerical intelligence → Biology self-concept	-	.019	.020
SES → Biology self-concept	-	-.031	.047
Gender → Biology self-concept	-	-.011	.044
Verbal intelligence → Physics self-concept	-	.089*	.048
Numerical intelligence → Physics self-concept	-	.029	.000
SES → Physics self-concept	-	.039	-.015
Gender → Physics self-concept	-	.266*	.188*
Relations of covariates to achievement			
Verbal intelligence → German achievement	-	.206*	.115*
Numerical intelligence → German achievement	-	.112*	.011
SES → German achievement	-	.165*	.117*
Gender → German achievement	-	-.152*	-.202*
Verbal intelligence → Math achievement	-	.236*	.071*
Numerical intelligence → Math achievement	-	.133*	.108*
SES → Math achievement	-	.114*	.076*
Gender → Math achievement	-	.085*	.084*
Verbal intelligence → English achievement	-	.201*	.069*
Numerical intelligence → English achievement	-	.108*	-.004

(continued)

Tables S10 (continued)

SES → English achievement	-	.045	.098*
Gender → English achievement	-	-.014	-.087*
Verbal intelligence → Biology achievement	-	.168*	.081*
Numerical intelligence → Biology achievement	-	.072*	.022
SES → Biology achievement	-	.091*	.086*
Gender → Biology achievement	-	-.031	-.071*
Verbal intelligence → Physics achievement	-	.181*	.090*
Numerical intelligence → Physics achievement	-	.098	.025
SES → Physics achievement	-	.063	.090*
Gender → Physics achievement	-	.139*	.081*

Note. AT = academic track; NAT = non-academic tracks.

* $p < .05$.

Table S11

Factor Correlations of the Extended I/E Models with Students from all Federal States

	Mah sc	German sc	English sc	Physics sc	Math ach	German ach	English ach	Physics ach
German sc	.259*/.265*/.261*							
English sc	.236*/.127*/.264*	.285*/.268*/.305*						
Physics sc	.355*/.331*/.328*	.282*/.306*/.282*	.200*/.129*/.195*					
Biology sc	.294*/.239*/.319*	.358*/.357*/.363*	.199*/.143*/.203*	.377*/.371*/.385*				
German ach					.416*/.433*/.392*			
English ach					.430*/.425*/.400*	.550*/.559*/.511*		
Physics ach					.466*/.420*/.425*	.370*/.301*/.362*	.344*/.326*/.303*	
Biology ach					.402*/.359*/.384*	.400*/.396*/.350*	.363*/.302*/.344*	.502*/.381*/.505*

Note. The first coefficient refers to Model S10 of Table S5 (the extended I/E model with the total sample of students from all federal states), the second and third coefficients refer to Model S16 of Table S5 (the extended I/E model with a grouping variable consisting of academic track students and non-academic track students). Thereby, the second coefficient depicts the factor correlations for academic track students; the third coefficient depicts the factor correlations for non-academic track students in Model S16.
sc = self-concept; ach = achievement.
* $p < .05$.

Table S12

Goodness-of-fit Indices for Models with Gender x Track as a Grouping Variable

χ^2	df	CFI	TLI	RMSEA	SRMR	
Original I/E Model						
386.898	276	.981	.972	.036	.031	CFA model with separate factors for math and German self-concepts and achievements; configural invariance across school track groups; SES, verbal intelligence, and numerical intelligences as covariates
427.183	300	.978	.970	.037	.047	CFA model with separate factors for math and German self-concepts and achievements; factor loading invariance across school track groups; SES, verbal intelligence, and numerical intelligences as covariates
427.183	300	.978	.970	.037	.047	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; free estimation of I/E Model across groups; SES, verbal intelligence, and numerical intelligences as covariates
466.973	306	.973	.963	.042	.062	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; invariance of the within-domain achievement–self-concept relations across groups; SES, verbal intelligence, and numerical intelligences as covariates
449.996	306	.975	.967	.039	.055	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; invariance of the between-domain achievement–self-concept relations across groups; SES, verbal intelligence, and numerical intelligences as covariates
483.109	312	.971	.962	.042	.061	Path model with separate factors for math and German self-concepts and achievements; factor loading invariance; invariant estimation of I/E Model across groups; SES, verbal intelligence, and numerical intelligences as covariates

(continued)

Table S12 (continued)

χ^2	df	CFI	TLI	RMSEA	SRMR	
Extended I/E Model						
2295.997	1515	.951	.932	.041	.055	CFA model with separate factors for math, German, English, biology, and physics self-concepts and achievements; configural invariance across groups; SES, verbal intelligence, and numerical intelligences as covariates
2396.582	1575	.948	.931	.041	.061	CFA model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance across groups; SES, verbal intelligence, and numerical intelligences as covariates
2396.582	1575	.948	.931	.041	.061	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; free estimation of the extended I/E Model across groups; SES, verbal intelligence, and numerical intelligences as covariates
2451.132	1590	.946	.928	.042	.064	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements: factor loading invariance; invariance of the within-domain achievement–self-concept relations across groups; SES, verbal intelligence, and numerical intelligences as covariates
2518.381	1635	.944	.928	.042	.068	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; invariance of the between-domain achievement–self-concept relations across groups; SES, verbal intelligence, and numerical intelligences as covariates
2583.906	1650	.941	.925	.043	.070	Path model with separate factors for math, German, English, biology, and physics self-concepts and achievements; factor loading invariance; invariant estimation of the extended I/E Model across groups; SES, verbal intelligence, and numerical intelligences as covariates

Note. The grouping variable considered in these multi-group models consisted of four groups of students attending the academic track or vocational track in the German federal state of North Rhine-Westphalia, i.e., girls in the academic track (N = 460), boys in the academic track (N = 237), girls in the vocational track (N = 250), and boys in the vocational track (N = 272). N = 11 students had to be discarded from the analyses due to missing information on gender. All models are estimated with the Robust Maximum Likelihood (MLR) estimator; all χ^2 are significant ($p < .05$). CFA = confirmatory factor analyses; CFI = comparative fit index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.