Arens, A. Katrin; Helm, Friederike; Wolff, Fabian; Möller, Jens

Social and dimensional comparisons in the formation of German students' language self-concepts

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
The journal of experimental education 88 (2020) 2, S. 245-264

Running head: I/E MODEL WITH LANGUAGES

Social and Dimensional Comparisons in the Formation of German Students' Language Self-concepts

A. Katrin Arens

DIPF | Leibniz Institute for Research and Information in Education

Friederike Helm, Fabian Wolff, & Jens Möller,

Kiel University

Author Note

Correspondence concerning this article should be addressed to A. Katrin Arens, DIPF | Leibniz Institute for Research and Information in Education, Department on Research on Education and Human Development and Centre for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Rostocker Str. 6, D-60323 Frankfurt am Main, Germany, Email: arens@dipf.de, Phone: +49 69 24708 138.
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., Helm, F., Wolff, F., & Möller, J. (2020). Social and dimensional comparisons in the formation of German students' language self-concepts. *The Journal of Experimental Education, 88,* 245–264. https://doi.org/10.1080/00220973.2019.1635069

This paper is not the copy of record and may not exactly replicate the authoritative document published in *The Journal of Experimental Education*.

Abstract

This study extended the internal/external reference model to multiple languages including students' language of instruction, first foreign language, and second foreign language. We examined whether social and dimensional comparisons play similar roles in the formation of students' self-concepts related to different languages, and whether dimensional comparisons result in contrast or assimilation effects. All students had German as the language of instruction and English as the first foreign language, and were divided into a subsample (*N*=487) learning French and a subsample (*N*=481) learning Latin as a second foreign language. Invariance tests demonstrated that the achievement–self-concept relations were similar across the subsamples, but interesting group differences became apparent when analyzing the French and Latin subsamples separately.

Keywords: I/E model; academic self-concept; social comparisons; dimensional comparisons; languages; contrast and assimilation effects

Student motivation has been found to be important in language learning (e.g., Dörnyei, 2003; Gardner, 2010). Among various facets of students' motivation in education, academic self-concept is one of the most important and widely researched constructs (Marsh, 2007; Marsh & O'Mara, 2008). Students' academic self-concept is defined as students' competence self-perceptions related to the academic domain (Marsh & Craven, 2006; Shavelson, Hubner, & Stanton, 1976). Academic self-concept is a domain specific construct since students establish separate self-concepts for different domains and school subjects. Hence, students were found to depict separate self-concepts for different languages including their language of instruction (LOI), first foreign language (FFL), and second foreign language (SFL) (Arens & Jansen, 2006; Marsh et al. 2015; Möller, Streblow, Pohlmann, & Köller, 2006).

A high level of academic self-concept is associated with desirable outcomes since it is related to academic achievement (Marsh & Craven, 2006), but also to adaptive learning behavior such as effort, persistence, coursework selection (Marsh & O'Mara, 2008; Trautwein & Möller, 2016). Hence, high levels of language self-concepts are also desirable for students' language learning. This insight leads to the task for educational practice to enhance students' language self-concepts in order to contribute to students' successful language learning. Intervention approaches benefit from knowledge about the sources of language self-concepts as these sources should be specifically targeted in intervention programs. Hence, comprehending the sources and formation of language self-concepts is directly associated with practical implications.

Social and dimensional comparisons have been proposed to be major sources of domain-specific academic self-concepts. This is the core assumption of the internal/external frame of reference (I/E) model (Marsh, 1986; Möller, Pohlmann, Köller, & Marsh, 2009). In our study, we examined whether social and dimensional comparisons are also involved in the formation of language self-concepts. We therefore extended the original I/E model to students' LOI, FFL and SFL. We investigated a sample of German secondary school students,

all having German as the LOI, English as the FFL, but either French or Latin as the SFL. Therefore, we could also examine whether social and dimensional comparisons operate similarly or differently in the formation of SFL self-concepts when considering French and Latin as two different SFLs.

The I/E Model

The I/E model assumes that the formation of math and verbal self-concepts relies on an interplay between social comparisons (comparing one's own achievement in one domain with the achievement of others in the same domain) and dimensional comparisons (comparing one's own achievement across different domains) (Marsh, 1986; Möller et al., 2009). The I/E model is commonly depicted in a regression model where domain-specific academic self-concepts are regressed on domain-specific achievement indicators. Social comparisons evoke positive relations (i.e., regression paths) between achievements and self-concepts within the same domains (e.g., math achievement and math self-concept) and a positive correlation between math and verbal self-concepts. Dimensional comparisons evoke negative relations (i.e., regression paths) between achievements and self-concepts across different domains (e.g., math achievement and verbal self-concept) and a negative correlation between math and verbal self-concept. The positive correlation between math and verbal self-concepts due to social comparisons and the negative correlation due to dimensional comparisons balance out leading to a near-zero correlation. The I/E model thus offers a theoretical explanation for the consistently observed low correlation between math and verbal self-concepts despite substantial correlations between math and verbal achievements (Marsh, 1986).

Extending the I/E Model to Multiple School Subjects

Recently, the original I/E model has been generalized and extended by including other predictor and outcome variables beyond math and verbal achievements and self-concepts (see the generalized I/E (GI/E) model; Möller, Müller-Kalthoff, Helm, Nagy, & Marsh, 2016). One such extension is the inclusion of a wide range of school subjects since the original I/E

model only considers one math and one verbal domain (Arens, Möller, & Watermann, 2016; Jansen, Schroeders, Lüdtke, & Marsh, 2015; Marsh et al., 2014, 2015; Möller et al., 2006). Respective studies showed substantial positive relations between achievements and self-concepts within the same domains. Hence, social comparisons seem to be consistently involved in the formation of domain-specific self-concepts. Across different domains, the relations between achievements and self-concepts were not always negative as is the case in the original I/E model, but they were also found to be positive in some cases. Hence, dimensional comparisons can invoke negative achievement–self-concept relations across domains (contrast effects). In this case, good performance in one domain entails lower levels of self-concept in the compared domain. In addition, dimensional comparisons can invoke positive achievement–self-concept relations across domains (assimilation effects). Here, good performance in one domain entails higher levels of self-concept in the compared domain.

The finding that dimensional comparisons can result in both contrast and assimilation effects has been linked to the Marsh/Shavelson model of academic self-concept (Marsh, 1990). Accordingly, domain-specific self-concepts are placed on a continuum ranging from a pure math endpoint to a pure verbal endpoint. Domain-specific self-concepts are thus categorized based on the similarity between domains since self-concepts addressing similar domains and sharing conceptual overlap (e.g., math and physics) are placed next to each other on the math-verbal continuum, while self-concepts of dissimilar domains (e.g., math and languages) are more distant from each other. Dimensional comparisons are assumed to result in contrast effects when two domains are considered that are located far from each other on the continuum, thus between dissimilar domains such as math and languages. In turn, assimilation effects are assumed to occur between closely related domains, thus between domains sharing some overlap or similarity.

Dimensional comparison theory (DCT; Möller & Marsh, 2013) has been explicitly formulated in response to the vast amount of research on the I/E model and considers the

scope of application and consequences of dimensional comparisons. DCT also notes that dimensional comparisons can result in both contrast and assimilation effects. According to DCT, dimensional comparisons lead to contrast effects when considering achievements and self-concepts related to domains for which the respective abilities are believed to be negatively correlated, see Hypothesis VI in Möller, Helm, Müller-Kalthoff, Nagy, & Marsh, 2015). Assimilation effects in turn are assumed to occur between domains for which the respective abilities are believed to be positively correlated, see Hypothesis VII in Möller et al., 2015). Hence, the assumptions of DCT regarding contrast and assimilation effects are closely linked to the assumptions derived from the Marsh/Shavelson model of academic self-concept since domains with (dis)similar underlying abilities might be closer to (far away from) each other on the math-verbal continuum leading to assimilation (contrast) effects.

Findings from previous studies rendered support for the assumptions regarding the occurrence of contrast and assimilation effects derived from the Marsh/Shavelson model of academic self-concept and DCT. In fact, the findings consistently demonstrated negative relations, and thus contrast effects, between achievements and self-concepts related to math and verbal domains (Möller et al., 2009). Math and verbal domains are located at the opposite ends of the math-verbal continuum of academic self-concepts in the Marsh/Shavelson model (Marsh, 1990). Moreover, students might suppose that math and verbal achievements are influenced by different underlying abilities. Positive relations and thus assimilation effects have been found between achievements and self-concepts related to math and physics (Arens et al., 2016; Jansen et al., 2015; Marsh et al., 2014, 2015; Möller et al., 2006). Math and physics might be conceptualized as math-like domains, are thus located close to each other on the math-verbal continuum of academic self-concepts. Moreover, physics achievements might be linked to math abilities.

Contrast and Assimilation Effects between Languages

In the original I/E model, verbal achievement and self-concept address students' LOI. In line with the GI/E model, some recent studies have expanded the verbal domain by including two languages, that is, students' LOI and FFL. One might assume a positive relation between achievements and self-concepts related to different languages, thus an assimilation effect. Different languages can be allocated to the verbal endpoint of the academic self-concept continuum, they share conceptual similarity due to the verbal nature, and students might suppose underlying verbal abilities to be responsible for language achievement in general. Moreover, transfer effects were demonstrated for language learning since students' proficiency in one language has been found to facilitate learning in another language (Chen, Xu, Nguyen, Hong, & Wang, 2010; Cunningham & Graham, 2000; Gebauer, Zaunbauer, & Möller, 2013; Gottardo, Yan, Siegel, & Wade-Woolley, 2001; Kellerman, 1995).

However, surprising findings have been reported regarding the relations between achievements and self-concepts related to different languages. In fact, these relations were negative or non-significant indicating a contrast effect rather than an assimilation effect between languages. For instance, in the study by Xu et al. (2000) with a sample of secondary school students from Hong Kong, non-significant paths between Chinese (English) achievement and English (Chinese) self-concept were found. Also examining a sample of students from Hong-Kong, Marsh, Kong, and Hau (2001) demonstrated negative paths – some of which were statistically significant – between Chinese (English) achievement and English (Chinese) self-concept. Moreover, Marsh and Yeung (2001) demonstrated negative relations between Spanish achievement and a higher-order factor of verbal self-concept (encompassing English, history, and general verbal self-concepts) and between verbal achievement and Spanish self-concept. The occurrence of contrast effects between languages was also corroborated by numerous studies with German student samples learning English as the FFL (Arens et al., 2016; Marsh et al., 2015; Niepel, Brunner, & Preckel, 2014). In particular, the studies documented weak, but primarily negative relations between German (English)

achievement and English (German) self-concept. Hence, across student samples from different countries (Germany, Hong- Kong, US), dimensional comparisons between languages were found to rather lead to contrast than to assimilation effects.

So far, most of the studies on I/E models including more than one language have only considered students' LOI and FFL, that is, two languages. There has been a shortage of studies investigating the pattern of achievement–self-concept relations to disclose the operation of social and dimensional comparisons when including three languages, that is, students' LOI, FFL, and SFL. Using a sample of German secondary school students, Marsh et al. (2015; Study 1) examined an extended I/E model inter alia including German as students' LOI, English as students' FFL, and students' SFL. Supporting the operation of social comparisons in the formation of language self-concepts, there were substantial positive relations between achievements and self-concepts when the achievements and self-concept measures addressed the same domains. The relations across achievements and self-concepts addressing the three different languages (German, English, and SFL) were either significantly negative or not statistically significant. Accordingly, from this study, one can conclude that, if there are any dimensional comparisons affecting the formation of language self-concepts related to German, English, and a SFL, they invoke contrast effects rather than assimilation effects across these languages.

So far, the study by Marsh et al. (2015) has been the only study investigating an I/E model with three languages and thus needs to be replicated. Moreover, in the above mentioned study the SFL was not specified; hence, it was not clear which specific SFL the students learned and whether all students learned the same SFL. Hence, further studies are necessary which specify the SFL and preferably include different SFLs. Such studies would render it possible to examine whether the pattern of achievement–self-concept relations hinting at social and dimensional comparisons in the formation of language self-concepts similarly applies to different SFLs. The study by Marsh et al. (2015) mainly indicated contrast

effects in achievement–self-concept relations across languages when including a SFL. Still, the pattern of achievement–self-concept relations across languages and thus the occurrence of contrast and assimilation effects might vary contingent upon the specific SFL considered. Therefore, we conducted a study with German students with German as the LOI, learning English as the FFL, and learning French or Latin as their SFL. We thus examined whether social and dimensional comparisons apply to the formation of self-concepts related to different languages including SFLs (replicating the results found by Marsh et al., 2015), and whether dimensional comparisons similarly lead to contrast or assimilation effects when considering French and Latin as different SFLs (extending the study by Marsh et al., 2015).

Framework of the Present Study

Language Learning in German Secondary Schools

The German secondary school system is characterized by a relatively strict tracking procedure. After four years of elementary school, students are commonly allocated to one ability track of secondary schooling (Becker, Neumann, & Dumont, 2017). The academic track is the highest track, and graduation from this track allows entrance to university. Students in the academic track have to learn two foreign languages at least until upper secondary levels when they have more options to choose courses. English is most often learned as the FFL. For SFLs, students are given a choice. The final selection depends on personal and parental preferences but is also determined by the attended school, since schools differ in the range of SFLs students can select. French and Latin constitute the most commonly learned SFLs and, accordingly, most of the academic track secondary schools offer French and Latin as SFLs.

In Germany, federal states are responsible for education. This leads to differences in the secondary school systems across federal states, although the formal characteristics of the academic track are very similar across states (Becker et al., 2016). Still, while FFL learning starts in elementary school and continues in secondary school in all federal states

(Fleckenstein, Möller, & Baumert, in press), the onset of SFL learning varies across states with some starting in year 6, and others starting in year 7. We therefore restricted our sample to students from one German federal state (Schleswig-Holstein) in order to eliminate possible confounding effects due to variations in the secondary school systems and differing onsets of SFL learning. In our sample, all participating students had German as their LOI and learned English as their FFL. All students started learning the SFL in year 7. One student subsample learned French and one subsample learned Latin as the SFL.

Aims and Hypotheses

In sum, the aim of the present study was to test whether social and dimensional comparisons are involved in the formation of language self-concepts including SFL self-concept and whether the findings are similar or different when considering French and Latin as two different SFLs. To this aim, we examined an I/E model extended to three languages besides math (Figure 1). Thus, four subject domains, including three languages, were considered in the present study (i.e., math, German as the LOI, English as the FFL, and French or Latin as the SFL). We explicate our assumptions regarding the role and effects of social and dimensional comparisons in the formation of language self-concepts in the following.

Social comparisons. We expected that social comparisons are involved in the formation of all domain-specific academic self-concepts. Hence, we consistently presumed positive achievement–self-concept relations within the same domains irrespective of the specific domain considered.

Dimensional comparisons. We differentiate between dimensional comparisons occurring across math and languages, and across different languages. As indicated above, the occurrence of contrast versus assimilation effects due to dimensional comparisons is assumed to depend upon the similarity and conceptual overlap of domains (Marsh et al., 2014, 2015). Here, we point out the following categories of (dis)similarity based on which we framed our

considerations regarding contrast and assimilation effects across domains: (a) origin of the language, (b) active/spoken versus passive/non-spoken nature of the language, and (c) logical approach to the subject.

(1) Across math and languages: The French and Latin subsamples were not expected to differ with regard to the effects of dimensional comparisons involving math and German (LOI), and math and English (FFL). Based on the findings from previous studies and given the different and distant locations of math and languages on the math-verbal continuum of academic self-concept, we assumed contrast effects between math and German, and between math and English (Arens et al., 2016; Jansen et al., 2015; Marsh et al., 2015; Möller et al., 2006). With respect to dimensional comparisons involving math and SFL, differences might occur between French and Latin. Given the verbal nature of French, parallel to the presumed contrast effect between math and German and between math and English, a contrast effect was expected between math and French. The expectations regarding Latin were less clear. On the one hand, the verbal character of Latin as a language might also lead to a contrast effect between math and Latin. On the other hand, Latin has a logical and deductive nature requiring reasoning abilities (Ortner, Asanger, Kubinger, & Proyer, 2008). Given this similarity to math, one could also expect an assimilation effect between math and Latin.

(2) Across languages: We expected small negative relations (contrast effects) or no relations between German and English given the findings from previous studies (Arens et al., 2016; Marsh et al., 2015; Niepel et al., 2014).

When considering the relations involving German and students' SFL, it has to be noted that German is a Germanic language, while the two SFLs (French and Latin) are Romance languages. Hence, contrast effects might occur between German and French as well as between German and Latin due to the different origins of languages. A particularly strong contrast effect might be expected between German and Latin. Latin differs from German

not only in its origin, but Latin is a passive (non-spoken) language while German is an active (spoken) language.

Still, it is plausible to assume an assimilation effect between German and Latin. Students might gain a higher level of competence and understanding of the German grammatical structure by learning Latin. Haag and Stern (2000) showed that German students learning Latin as their FFL were superior on some indicators of German competencies compared to students who did not learn Latin or only learned Latin as the SFL. Accordingly, other studies demonstrated a positive effect of learning Latin on students' proficiency in other languages (Barber, 1985; Masciantonio, 1977).

Regarding the relations between English as students' FFL and the SFL, an assimilation effect might occur, irrespective of whether French or Latin is considered. Students might perceive their FFL and SFL as belonging to an overarching category of foreign languages. Yet, such an assimilation effect was not documented by Marsh et al. (2015) although this finding needs to be replicated with other student samples. Still, the existence of contrast versus assimilation effects between FFL and SFL might vary contingent upon the specific SFL considered. French and Latin are Romance languages, while English is a Germanic language. Thus, a contrast effect might indeed be assumed for the relations between English and French and between English and Latin due to the different language origins. This contrast effect might be even enhanced for the relations between English and Latin given that English is an active (spoken) and Latin is a passive (nonspoken) language. However, regarding the relation between English and Latin, one might also presume an assimilation effect given the facilitating aspect of Latin learning for learning other languages including English (Barber, 1985; Masciantonio, 1977).

Summary

In sum, our research builds on previous studies extending the I/E model to multiple school subjects, but particularly considers multiple languages, that is students' LOI (German),

FFL (English), and SFL (French or Latin). Therefore, we tested the effects of social and dimensional comparisons in the formation of German students' self-concepts for different languages that are central in the German secondary school system. Furthermore, we compared the pattern of findings across the two SFLs French and Latin. Based on the various characteristics of the different languages along with their (dis)similarities, we referred to several considerations regarding the achievement–self-concept relations across languages to figure out contrast and assimilation effects due to dimensional comparisons. Given the lack of research on I/E models involving more than two languages, we had to treat diverse, albeit similarly plausible, deliberations regarding the pattern of achievement–self-concept relations in parallel.

Method

Sample

The data analyzed in this study were part of a larger data set examining the formation of students' academic self-concepts related to multiple school subjects. Only students attending the academic track of secondary schools ("Gymnasium") were selected. All students (N=970) attended grade levels 9 to 11. Two subsamples were identified: students learning French (N=489), and students learning Latin (N=481) as the SFL. The two subsamples were similar regarding student characteristics including age, school grades, grade point average, and years of learning the SFL (Table S1 of the Online Supplements)[1]. Parental consent was obtained for all participants and students were informed about the purpose of the study and the confidential treatment of their data. Data collection took place in students' classrooms during regular lessons and was administered by trained research assistants. The study was conducted from the beginning of May to the beginning of July 2017.

Measures

Self-concepts. A questionnaire which contained scales regarding the self-concepts in various school subjects was administered to the students. The students were instructed to only

complete the scales which referred to school subjects they were being taught. In this context, the students completed questionnaires in which they were asked for their self-concepts in German (LOI), English (FFL), math, and their SFL. All students completed self-concept measures with regard to German, English, and math, the student subsamples analyzed here differed with regard to their completed measure of SFL self-concept. All students completed self-concept measures with regard to German, English, and math; students taking French as the SFL (the French subsample) responded to the French self-concept items but not to the Latin self-concept items; students taking Latin as the SFL (the Latin subsample) responded to the Latin self-concept items but not to the French self-concept items.

The scales for measuring the different domain-specific self-concepts all consisted of four items, which were adapted from Jopt (1978) and Jerusalem (1984). They have successfully been used in contemporary self-concept research (Helm & Möller, 2017; Möller, Zimmermann, & Köller, 2014; Zimmermann, Möller, & Köller, 2018). The items were worded in parallel across the domains: "I can achieve at most things in [subject]", "Nobody's perfect but I'm just not good at [subject].", "With some of the topics in [subject], I know from the start that I just won't get them", "I am good at [subject]." The students responded to the items on a 6-point Likert scale ranging from 1=*not true at all* to 6=*fully true*. Prior to analyses, the items were consistently coded so that higher ratings depicted higher levels of self-concept. All scales demonstrated good reliability (Table S1 of the Online Supplements).

Achievement. The students reported the school grades they had obtained in their latest school report (i.e., mid-term report received in January 2017) in German, English, math, and their SFL. Student reported grades have been found to be of sufficient validity (Dickhäuser & Plenter, 2005). 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, grades were reversely coded before all analyses, thus higher values indicated higher levels of achievement.

Statistical Analyses

The analyses were conducted within the framework of structural equation modeling (SEM; e.g., Kline, 2005) using the statistical package Mplus 8.2 (Muthén & Muthén, 1998-2018). All models were estimated by applying the robust maximum likelihood estimator which has been shown to be robust against violations of normality assumptions. Missing values on all variables were estimated by the Full Information Maximum Likelihood (FIML) implemented in Mplus. The FIML approach is known to be reliable, to lead to unbiased parameter estimates, and to be as trustworthy as multiple imputation when handling data that are missing at random or missing completely at random (Enders, 2010; Graham, 2009). The amount of missing data ranged between 0.0% and 2.8% for the self-concept items, and between 2.7% and 23.5% for the school grades. Little's Missing Completely at Random Test was non-significant for the self-concept items ($\chi^2(465)=468.46$, $p=.45$), and it was not significant for the school grades ($\chi^2(14)=23.06$, $p=.06$). To account for the multilevel structure of the data (students nested in classes), the Mplus option "type=complex" was used, with students' classes treated as the clustering variables[2]. This option corrects for possible bias in standard errors resulting from the hierarchical nature of the data. Finally, all models included correlated uniquenesses between parallel-worded self-concept items across domains to account for potentially shared method variance (Marsh et al., 2013).

To evaluate the fit of the latent models, we relied on several commonly accepted goodness-of-fit indices as there is no conclusive index for evaluating the goodness of fit of latent models (Marsh, Hau, & Wen, 2004). We thus considered 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 between .90 and .95 are commonly accepted as indication of a good model fit, although some authors (e.g., Hu & Bentler, 1998) suggest a stricter criterion of .95. Concerning the RMSEA, Browne and Cudeck (1993) proposed values below .05 as indicative of a close fit, values between .05 and .08 as indicative of a reasonable fit, and values greater than .10 as indicative of a poor fit.

However, Hu and Bentler (1998) also conceived a value close to .06 as a good fit. Regarding the SRMR, Hu and Bentler (1998) recommend a cut-off value of .08, while others (e.g., Kline, 2005) accept the less conservative value of .10. Despite these cut-off criteria for the different descriptive goodness-of-fit indices for the purpose of model fit, it has to be noted that they should rather be treated as guidelines instead of "golden rules". Along with 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).

Separate analyses for the French and Latin subsamples. We first ran separate analyses for the French and Latin samples. Here, we stated confirmatory factor analytic (CFA; Brown, 2006) models assuming separate achievement and self-concept factors for each domain. Hence, for instance, for the French subsample, separate factors were stated for achievements and self-concepts in German, English, math and French, that is, eight factors in total. The achievement factors were single-item factors defined by students' school grades with the measurement errors fixed to zero. The self-concept factors were defined by the respective domain-specific set of items.

We then tested an extended I/E model in which the different self-concept factors were regressed on the different achievement factors (Figure 1). This model provided information about the paths leading from domain-specific achievements to domain-specific self-concepts while controlling for the other relations. Significantly positive relations within the same domains indicated the operation of social comparisons in the formation of domain-specific self-concept facets. Significant cross-domain relations hinted at the operation of dimensional comparisons with positive cross-domain relations indicating assimilation effects and negative cross-domain relations indicating contrast effects.

Invariance tests. Given that the French and Latin subsamples completed different self-concept measures (French versus Latin self-concept scales) and reported on their French

versus Latin achievement, they were independent subsamples. This precluded comparisons of the path coefficients for the achievement–self-concept relations across groups. To circumvent this limitation, we created new variables in each sample, that is, SFL achievement and SFL self-concept. These variables were defined by the French achievement and self-concept measures in the French subsample, and by the Latin achievement and self-concept measures in the Latin subsample. The French and Latin subsamples were merged, and multi-group models were conducted to compare the pattern of achievement–self-concept relations across groups. To this aim, in a CFA model including domain-specific achievement and self-concept factors, all model parameters were first freely estimated across groups (configural invariance). We then tested for metric invariance and thus constrained the factor loadings to be invariant across the French and Latin subsamples. Metric invariance ensures that the same constructs with the same underlying meanings are measured across groups. Metric invariance is a necessary, yet sufficient precondition for the inspection of invariance in the relations among constructs, that is, to test group invariance in the pattern of relations among domain-specific achievements and self-concepts within the same and across different domains (i.e., the extended I/E model including multiple languages; Meredith, 1994; Millsap, 2011). For examining the invariance of the extended I/E model including multiple languages, we followed the approach by Xu et al. (2013). Hence, we examined an I/E model with all path coefficients stated to be equal across groups, an I/E model in which only the relations within the same domains depicting social comparisons were stated to be equal across groups, and an I/E model in which only the relations across different domains depicting dimensional comparisons were stated to be equal across groups. These I/E models with invariance constraints were compared to an extended I/E model without any invariance constraints, that is, a model in which the pattern of achievement–self-concept relations within and across domains was freely estimated across groups. In order to evaluate the invariance models, we relied on the guidelines proposed by Cheung and Rensvold (2002). Accordingly, invariance

can be accepted as long as the CFI does not drop more than .01 between more and less restrictive models.

Results

Separate Analyses for the French and Latin Subsamples

The CFA models stating separate factors for German, English, math, and SFL self-concepts and achievements fitted the data well in both the French and Latin subsamples (Table 1). The self-concept items had positive loadings on their respective self-concept factors indicating the integrity of the used measures (Table S2 of the Online Supplements). While the domain-specific achievements were positively correlated with each other (French subsample: *r*s=.280 to .488; Latin subsample: *r*s=.267 to .545), the domain-specific academic self-concepts showed lower intercorrelations (French subsample: *r*s=-.029 to .330; Latin subsample: *r*s=-.029 to .310; Table S3 of the Online Supplements).

The latent regression models (Table 2 and Table S4 of the Online Supplements) in which the self-concept factors were regressed on the achievement factors (i.e., an I/E model including math, LOI, FFL, and SFL, Figure 1) are statistically equivalent to the CFA models as the factor correlations were only replaced by path coefficients. Thus, both types of models resulted in the same fit. Supporting social comparisons in the formation of domain-specific self-concepts, the findings showed substantial positive paths between achievements and self-concepts within the same domains in the French (German: β=.651; English: β=.731; math: β=.860; French: β=.753, for all $p<.001$) and in the Latin subsample (German: β=.680; English: β=.806; math: β=.854; Latin: β=.832, for all $p<.001$).

Regarding cross-domain relations involving math, the findings documented contrast effects between math and German achievements and self-concepts in both the French and Latin subsamples. Math achievement was negatively related to German self-concept (French subsample: β=-.156; Latin subsample: β=-.201; both $p<.001$), and German achievement was negatively related to math self-concept in both subsamples (French subsample: β=-.115; Latin

subsample: β=-.157; both $p<.001$). With regard to English, the findings also indicated contrast effects to math in both subsamples. Math achievement displayed negative paths to English self-concept (French subsample: β=-.128; Latin subsample: β=-.149; both $p<.01$), and English achievement was negatively related to math self-concept (French subsample: β=-.101; Latin subsample: β=-.163; both $p<.001$). We then considered the relations between math and SFL. In the French subsample, math achievement was significantly negatively related to French self-concept (β=-.116, $p<.01$), but French achievement was not significantly related to math self-concept (β=-.023, *ns*). In the Latin subsample, the path between math achievement and Latin self-concept was not significant (β=-.050, *ns*), and Latin achievement was not related to math self-concept (β=-.020, *ns*).

Regarding cross-domain relations among languages, we first considered the relations between German and English. Indicating a contrast effect, German achievement was negatively related to English self-concept in the French subsample (β=-.180, $p<.001$) as well as in the Latin subsample (β=-.197, $p<.001$). The paths leading from English achievement to German self-concept were not significant in both subsamples (French subsample: β=-.102; Latin subsample: β=.024; both *ns*). We then inspected the relations between German on the one hand and French or Latin as students' SFL on the other hand. In both subsamples, the relation between German achievement and SFL self-concept was not significant (French subsample: β=.006; Latin subsample: β=-.030; both *ns*). The relation between French achievement and German self-concept was significantly positive (β=.121, $p<.05$), but the relation between Latin achievement and German self-concept was not significant (β=.015, *ns*). Finally, we examined the relations involving English (FFL) and French or Latin (SFL). The relations involving English and French were all not significant (English achievement and French self-concept: β=-.064; French achievement and English self-concept: β=.066; both *ns*). In the Latin subsample, the findings showed a significantly negative path from English

achievement to Latin self-concept (β=-.109, p<.05), while the relation between Latin achievement and English self-concept was not significant (β=-.060, *ns*).

Invariance Tests

Invariance tests were applied to compare the French and Latin subsamples. They based on the newly created SFL achievement and self-concept factors. Compared to a model of configural invariance in which all model parameters were freely estimated across the French and Latin subsamples, the CFI value only dropped by Δ=-.001 when assuming invariant factor loadings (Table 1). This allowed us to compare factor relations across groups, that is, to test whether the French and Latin subsamples differed in their achievement–self-concept relations within the same and across different domains. Based on the model with invariant factor loadings, we estimated the extended I/E model (Figure 1) freely across the French and Latin subsamples. This model is statistically equivalent to the CFA model including invariant factor loadings and thus resulted in the same fit. Relative to the freely estimated extended I/E model, the CFI value only dropped by Δ=-.001 when assuming invariance of the path coefficients for all achievement–self-concept relations, when assuming invariance of the path coefficients for the achievement–self-concept relations within the same domains only, and when assuming invariance of the path coefficients for the achievement–self-concept relations across different domains only. Hence, the sizes of the achievement–self-concept relations were invariant across the French and Latin subsamples irrespective of whether only the within-domain, only the cross-domain, or all relations were considered.

The findings revealed significantly positive relations between achievements and self-concepts within the same domains (Table 3; see also Table S5 of the Online Supplements). Math and German achievements and self-concepts displayed significantly negative relations. Significantly negative relations were also found between math and English achievements and self-concepts. German achievement showed a significantly negative relation to English self-concept, but the path between English achievement and German self-concept was not

significant. SFL self-concept displayed negative relations to English and math achievements, but was unrelated to German achievement. SFL achievement did not demonstrate any significant relations to self-concepts of other domains, that is, to German, English, and math self-concepts.

Discussion

Drawing on the I/E model framework, this study was the first one investigating whether social and dimensional comparisons are involved in the formation of students' language self-concepts when considering three languages simultaneously (LOI, FFL, and SFL). The I/E model originally only included students' LOI (e.g., Marsh, 1986; Möller et al., 2009). In line with the GI/E model (Möller et al., 2016), the original I/E model was then extended to students' LOI and FFL (e.g., Arens et al., 2016; Möller et al., 2006; Niepel et al., 2014), but has so far rarely also integrated students' SFL (but see Marsh et al., 2015). Since one subsample of the participating students learned French as the SFL, while another subsample learned Latin as the SFL, we could find out whether the pattern of achievement–self-concept relations varied across French and Latin as SFLs. Our double-staged analytic approach firstly included separate analyses for the French and Latin subsamples and secondly invariance tests for which the French and Latin subsamples were merged using newly generated SFL achievement and self-concept variables.

Summary of Findings

The following results applied to both the French and Latins subsamples and were demonstrated irrespective of whether the French and Latin subsamples were analyzed separately or merged for invariance tests:

(1) The findings supported the existence of social comparisons in the formation of domain-specific self-concepts since positive achievement–self-concept relations within the same domains were found irrespective of the domain considered.

(2) Between math and German achievements and self-concepts and between math and English achievements and self-concepts, dimensional comparisons seem to operate. These dimensional comparisons lead to contrast effects as there were negative relations between math achievement on the one hand and German and English self-concepts on the other hand. In addition, the findings showed negative relations between German and English achievements on the one hand and math self-concept on the other hand.

(3) The relations between German and English self-concepts and achievements were also characterized by contrast effects although they were only evident from negative relations between German achievement and English self-concept, while the relations between English achievement and German self-concept were not significant.

So far, the results of our study replicated the findings of previous studies supporting contrast effects between math and verbal domains (i.e., languages), but also, albeit weaker, contrast effects between German and English as two languages (Arens et al., 2016; Marsh et al., 2015; Niepel et al., 2014).

The major contribution of our study addresses the inclusion of a SFL while considering both French and Latin as two SFLs with different characteristics. The findings of the invariance tests hinted at invariant achievement–self-concept relations for students learning French and students learning Latin as the SFL. SFL self-concepts were unrelated to students' German achievement, but negatively related to math and English achievements. Hence, dimensional comparisons seem to be also involved in the formation of SFL self-concepts and they rather lead to contrast effects than to assimilation effects. This observation refutes the original assumptions based on the Marsh/Shavelson model of academic self-concept and on the subsequently established DCT that assimilation effects should occur between domains located close to each other (e.g., near the verbal endpoint) on the math-verbal continuum (Marsh, 1990; Marsh et al., 2014) and between domains which might base on similar abilities (Möller et al., 2015). Moreover, this observation replicated the findings by

Marsh et al. (2015) indicating that contrast effects among languages exist even when involving a SFL. We extended this conclusion by demonstrating that contrast effects among languages occur irrespective of the specific SFL considered.

Yet, the notion of contrast effects involved in the formation of SFL self-concept has to be somewhat qualified since some differences emerged between the French and Latin subsamples when considering them separately. In fact, the findings revealed a significantly negative relation between math achievement and French self-concept indicating the expected contrast effect between math and language domains. The relation between math achievement and Latin self-concept was not significant. This lacking contrast effect might be due to the logical nature of both math and Latin.

Moreover, considering the relations among languages, the relation between English achievement and Latin self-concept was significantly negative while the relation between English achievement and French self-concept was not significant. This finding might support the assumption that the passive nature of Latin evokes a contrast effect to English though this effect was absent between English and French as two actively spoken foreign languages.

These two findings point to slight differences in the pattern of findings regarding SFL self-concept formation contingent upon whether French or Latin were considered as the SFL. The slightly different modes of operation of French and Latin also became visible as we found a significantly positive relation between French achievement and German self-concept, but a non-significant relation between Latin achievement and German self-concept. This finding might support the assumption of a stronger verbal nature of French than Latin leading to an assimilation effect to German. These differences in the pattern of achievement–self-concept relations between the French and Latin subsamples should not be overinterpreted given the results of the invariance tests. Still, the respective insights may redirect attention towards the role of (dis)similarity between the domains considered. Stronger dissimilarity between domains might enhance the chance of contrast effects. Similarity between domains, in turn,

might lead to assimilation effects but might also simply weaken contrast effects. At least, our findings demonstrated that the effects of dimensional comparisons (i.e., contrast or assimilation effects) might vary with the language considered. Based on these insights, we outline some directions for future research.

Directions for Future Research

To hypothesize on potential differences in the achievement–self-concept relations when involving French or Latin as two different SFLs, we considered the objective characteristics including the similarities and dissimilarities of both languages. Yet, the objective differences in the characteristics and affordances of languages have to be distinguished from students' perceived, subjective differences and similarities between languages. Hence, future studies should include students' perceptions of languages. Students should thus be asked, for instance, whether they perceive different languages as more verbal-like, as more math-like, or as (dis)similar to each other. In an experimental study, relative to higher perceived similarity, lower perceived similarity (i.e., higher dissimilarity) was found to be associated with larger mean level differences between domain-specific self-concepts reflecting a contrast effect (Helm, Müller-Kalthoff, Nagy, & Möller, 2016). Hence, students' perceptions of similarity between domains seem to matter for the strength of contrast effects. In this context, qualitative and introspective studies might also be useful since they can give some indication of students' perceptions of domains, students' explicit and deliberate use of social and dimensional comparisons in the formation of language self-concepts, and the interrelation between students' domain-specific perceptions and the effects of dimensional comparisons (Möller & Husemann, 2006).

In this study considering the languages of German, English, French, and Latin, we derived our assumptions based on three criteria of (dis)similarity between these languages: language origin, active/spoken versus passive/non-spoken nature of the language, and the logical approach to the subject. It has to be resolved which criteria of (dis)similarity between

domains in general and between languages in particular are especially responsible for the operation and effects of dimensional comparisons. Here, other criteria of (dis)similarity have to be taken into account, such as orthographic systems. Moreover, our study was restricted to comparisons between math, German, English, and French or Latin. Although we thus integrated a variety of languages which can be categorized on a range of criteria of (dis)similarity, future studies need to consider other languages which are even more similar to each other (e.g., French and Spanish) or which are highly dissimilar (e.g., Chinese and English).

Differences in the pattern of achievement–self-concept relations regarding languages might also originate from the different periods of time students had been learning a language. Naturally, the students had been learning the SFL for a shorter period of time than the FFL. Students' domain-specific self-concepts are assumed to be formed through their experiences with specific domains (Marsh & Ayotte, 2003; Shavelson et al., 1976). The longer students study a SFL, the more students might realize potential differences and similarities between different languages. Hence, dimensional comparisons might become more salient and apparent after a longer period of SFL learning. In our study, students might have already established an English self-concept which is well separable from their self-concepts in other school subjects, thus invoking a contrast effect to math as well as to German. Students might have not yet formed a differentiated self-concept regarding the SFL leading to the observed weaker or missing contrast effects to math, German, and English. Further insights might be gained from systematic investigations of the impact of students' learning experiences (e.g., duration, frequency of lessons) with a specific subject on the formation of the corresponding self-concept.

In our study, all participating students had German as the LOI and English as the FFL, but the students had either chosen French or Latin as the SFL. It was thus not possible to disclose whether the sequence of language learning might have affected the findings. Hence,

future studies should systematically account for different combinations and sequences of language learning. Moreover, other SFLs learned by German students (e.g., Spanish which is gaining popularity as a SFL with German students) were not considered and should be included in future studies.

Theoretical Implications

From the perspective of research on the I/E model, this study fits research on the GI/E model since one extension of the original I/E model addresses the inclusion of multiple school subjects (Marsh et al., 2015). Extensions of the original I/E and thus studies within the framework of the GI/E model help refine DCT (Möller & Marsh, 2013). Along with the findings of other studies on I/E models including multiple languages (Arens et al., 2016; Marsh et al., 2015; Niepel et al., 2014), our findings might help to reformulate DCT to presume contrast effects rather than assimilation effects among languages. Still, as outlined above and as indicated by the slight differences between the French and Latin subsamples as found in our study, further research is necessary to specify the role of language similarity and further specific conditions for the occurrence of contrast and assimilation effects in the formation of language self-concepts. Respective future research might help resolve the heterogeneous results regarding contrast and assimilation effects among languages, since few studies actually found the originally assumed assimilation effect between achievements and self-concepts related to students' LOI and FFL (Marsh et al., 2014; Möller et al., 2006).

Practical Implications

Along with other studies (Arens & Jansen, 2016; Marsh et al. 2015; Niepel et al., 2014), the present study offers empirical evidence that students display separate self-concepts related to different languages including LOI, FFL, and SFL. Moreover, from this study, we learned that dimensional comparisons mainly leading to contrast effects are involved in the formation of language self-concepts. Still, teachers as well as parents might infer that a student who is performing well in one language perceives him/herself to perform well and to

be talented not only in that but also in other languages. From a teachers' and parents' perspective, this is comprehensible given the substantive correlations among language achievements and the well-known transfer effects between languages (Chen et al., 2010; Cunningham & Graham, 2000; Gebauer et al., 2013; Gottardo et al., 2001; Kellerman, 1995). In addition, previous studies demonstrated that teachers and parents do not take dimensional comparisons into account when they are asked to infer their students' or children's academic self-concepts (Dai, 2002; Helm, Müller-Kalthoff, Mukowski, & Möller, 2018). Hence, parents and teachers might underestimate the domain specificity of language6 self-concepts as well as the operation of dimensional comparisons in their formation. We would like to raise teachers' and parents' awareness for the existence of dimensional comparisons in students' formation of language self-concepts. Moreover, we would like to point out that students' language self-concepts as the subjective representations of one's language achievement behave differently compared to students' language achievements. Hence, teachers and parents should not solely focus on the enhancement of one language self-concept and feel confident that the positive intervention effects will spill over to other language self-concepts. Each domain-specific language self-concept should rather be treated and fostered separately in interventions. Moreover, the above mentioned insights might help practitioners understand that an individual student performing well in two languages can nevertheless display a relatively high self-concept in one language, but a relatively low self-concept in another.

Practitioners might also deliberately take advantage of the operation of dimensional comparisons in the formation of students' language self-concepts. Teachers and parents may intentionally emphasize similarities of different languages in order to boost an individual students' self-concept related to a specific language in the case that this student has already performed well in another language. Hence, in this case, transfer effects between languages should be initiated in order that students' good achievement in one language does not only positively affect students' self-concept in the same but also in another language (i.e.,

triggering assimilation effects and reducing contrast effects). In the event that an individual student has not succeeded in language learning so far, it might be prudent to underline the dissimilarities between the so far learned language and a new language. Thereby, contrast effects should be initiated which help to establish or preserve a high level of self-concept related to one language even in the case of poor achievement in another language.

Limitations

Our study contributes to research on the role of social and dimensional comparisons in the formation of language self-concepts. Yet, some shortcomings have to be mentioned. The data are cross-sectional and thus preclude temporal or causal inferences in order that further longitudinal and experimental studies are needed (see for example Möller & Köller, 2001). With longitudinal data, research has demonstrated that achievement and self-concept are reciprocally interrelated across time (reciprocal effects model, REM; Marsh & Craven, 2006). Yet, reciprocal relations between achievement and self-concept were predominantly examined within one domain only. The reciprocal I/E (RI/E) model combines the REM with the I/E model and examines the longitudinal relations between achievement and self-concepts with the same and across different domains (Möller, Retelsdorf, Köller, & Marsh, 2011). In a study with German students, Niepel et al. (2004) included two languages (German as the LOI, English as the FFL) in a RI/E model. Further studies are needed to replicate these findings and to integrate a SFL.

We considered German secondary school students attending the academic track who are obliged to learn a SFL. Academic track students are a positively biased sample of high-achieving students. Therefore, the results need to be tested with respect to their generalizability to other student samples. The findings of the present study only apply to German samples learning English as the FFL, and either French or Latin as the SFL. Hence, the study suffers from limited generalizability as it is unclear whether the findings are applicable to students from other countries or educational systems and to students with

another LOI and learning different languages as the FFL and the SFL. Studies on foreign language learning and foreign language motivation generally suffer from this limitation of generalizability and always relate to specific student samples with a specific LOI and specific FFLs and SFLs.

Conclusion

This study is in line with recent research on extending the I/E model to multiple school subjects (e.g., Marsh et al., 2015) and research on dimensional comparisons (Möller & Marsh, 2013). Its innovative contribution is due to the inclusion of three languages (LOI, FFL, and SFL), whereby we integrated French and Latin as two SFLs. Our findings indicated that social comparisons consistently operate in the formation of students' language self-concepts, including FFL and SFL self-concepts. Dimensional comparisons also take place in the formation of students' language self-concepts. Dimensional comparisons primarily led to contrast effects not only between math and languages but even among different languages. Although invariance tests implied similar achievement–self-concept relations across students learning French and students learning Latin as the SFL, some small differences became apparent when scrutinizing the French and Latin subsamples separately. Based on the complex pattern of findings, we identified some factors, encompassing (dis)similarities between languages, which might affect the results from I/E model studies including different languages. Hence, we proposed some guidelines for future research and indicated refinements on DCT. Finally, we advised practitioners to consider the domain specificity of students' language self-concepts as well as the role of dimensional comparisons in their formation.

Footnotes

[1] The gender ratio differed across the subsamples ($\chi(1, n=968)=19.674$, $p=.000$). T-tests demonstrated no differences in the school grades between the French and Latin subsamples except for higher mean chemistry grades in the Latin subsample than in the French subsample (t (623)=-2.419, $p<.05$).

[2] Given that the models analyzed here contained more free parameters than number of clusters, the standard errors of the model parameters may not be trustworthy. We therefore conducted all models without any cluster variables (see Tables S6 to S12 of the Online Supplements). The pattern of findings was similar although the significance level varied in some cases.

References

Arens, A.K., & Jansen, M. (2016). Self-concepts in reading, writing, listening, and speaking: A multidimensional and hierarchical structure and its generalizability across native and foreign languages. *Journal of Educational Psychology, 108,* 646–664. http://doi.org/10.1037/edu0000081

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

Barber, G.E. (1985). Latin as a practical study. *The Classical Journal, 81*, 158–160. https://www.jstor.org/stable/3296745

Becker, M., Neumann, M., & Dumont, H. (2017). Recent developments in school tracking practices in Germany: An overview and outlook on future trends. *ORBIS SCHOLAE*, *10*, 9–25. https://doi.org/10.14712/23363177.2017.8

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.

Chen, X., Xu, F., Nguyen, T.-K., Hong, G., & Wang, Y. (2010). Effects of cross-language transfer on first-language phonological awareness and literacy skills in Chinese children receiving English instruction. *Journal of Educational Psychology*, *102*, 712–728. https://doi.org/10.1037/a0018802

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

Cunningham, T.H., & Graham, C.R. (2000). Increasing native English vocabulary recognition through Spanish immersion: Cognate transfer from foreign to first language. *Journal of Educational Psychology, 92,* 37–49. http://dx.doi.org/10.1037/0022-0663.92.1.37

Dai, D.Y. (2002). Incorporating parent perceptions: A replication and extension study of the internal-external frame of reference model of self-concept development. *Journal of Adolescent Research*, *17*, 617–645. https://doi.org/10.1177/074355802237467

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

Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Language Learning, 53*, 3–32. doi:10.1111/1467-9922.53222

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

Fleckenstein, J., Möller, J., & Baumert, J. (in press). Ist jünger immer besser? Frühes Fremdsprachenlernen in der Grundschule [Is younger always better? Early foreign language learning at primary school]. *Zeitschrift für Pädagogische Psychologie*.

Gardner, R.C. (2010). *Motivation and second language acquisition. The socio-educational model.* New York: Peter Lang.

Gebauer, S.K., Zaunbauer, A.C.M., & Möller, J. (2013). Cross-language transfer in English immersion programs in Germany: Reading comprehension and reading fluency. *Contemporary Educational Psychology, 38*, 64–74. https://doi.org/10.1016/j.cedpsych.2012.09.002

Gottardo, A., Yan, B., Siegel, L.S., & Wade-Woolley, L. (2001). Factors related to English reading performance in children with Chinese as a first language: More evidence of

cross-language transfer of phonological processing. *Journal of Educational Psychology*, *93*, 530–542. https://doi.org/10.1037/0022-0663.93.3.530

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

Haag, L., & Stern, E. (2000). Non scholae sed vitae discimus? Auf der Suche nach globalen und spezifischen Effekten des Lateinunterrichts [Non scholae sed vitae discimus? In search of global and specific effects of learning Latin]. *Zeitschrift für Pädagogische Psychologie, 14*, 146–157. https://doi.org/10.1024//1010-0652.14.23.146.

Helm, F., & Möller, J. (2017). Der Einfluss von Erhebungsbedingungen auf Kontrasteffekte dimensionaler Vergleiche [The influence of manner of assessment on contrastive effects of dimensional comparisons]. *Psychologie in Erziehung und Unterricht, 64*, 24–34. http://dx.doi.org/10.1016/10.2378/peu2017.art03d

Helm, F., Müller-Kalthoff, H., Mukowski, R., & Möller, J. (2018). Teacher judgment accuracy regarding students' self-concepts: Affected by social and dimensional comparisons? *Learning and Instruction*, *55*, 1–12. https://doi.org/10.1016/j.learninstruc.2018.02.002

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. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424–453. http://dx.doi.org/10.1037/1082-989X.3.4.424

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

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.

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.

Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? *Annual Review of Applied Linguistics*, *15*, 125–150. https://doi.org/10.1017/S0267190500002658

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

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. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, *82*, 623–636. http://dx.doi.org/10.1037/0022-0663.82.4.623

Marsh, H.W. (2007). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester, UK: British Psychological Society.

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

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

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.

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., Kong, C.-K., & Hau, K.-T. (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. http://dx.doi.org/10.1037/0022-0663.93.3.543

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. & O'Mara, A.J. (2008). Self-concept is as multidisciplinary as it is multidimensional. In H.W. Marsh, R.G. Craven & D.M. McInerney (Eds.), *Self-processes, learning, and enabling human potential. Dynamic new approaches* (pp. 87-115). Charlotte, NC: Information Age.

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

Masciantonio, R. (1977). Tangible benefits of the study of Latin: A review of research. *Foreign Language Annals*, *10*, 375–382. https://doi.org/10.1111/j.1944-9720.1977.tb02999.x

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525–543.

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

Möller, J., Helm, F., Müller-Kalthoff, H., Nagy, N., & Marsh, H.W. (2015). Dimensional comparisons and their consequences for self-concept, motivation, and emotion. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (pp. 430–436). Oxford: Elsevier.

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. (2016). The generalized internal/external frame of reference model: An extension to dimensional comparison theory. *Frontline Learning Research, 4,* 1–11. https://doi.org/10.1037/0033-295X.84.6.485

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

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

Möller, J., Streblow, L., Pohlmann, B., & Köller, O. (2006). An extension to the internal/external frame of reference model to two verbal and numerical domains. *European Journal of Psychology of Education*, *11*, 467–487. https://doi.org/10.1007/BF03173515

Moller, 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

Muthén, L.K., & Muthén, B.O. (1998-2018). *Mplus User's Guide. Eighth Edition*. Los Angeles, CA: Muthén & Muthén.

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

Ortner, T.M., Asanger, M., Kubinger, K.D., & Proyer, R.T. (2008). Zur Frage nach Auswirkungen von Lateinunterricht auf die kognitive Fähigkeit „Reasoning" [On effects of school education in Latin on the cognitive ability "reasoning"]. *Psychologie in Erziehung und Unterricht, 55*, 189–195.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407–441. https://doi.org/10.3102/00346543046003407

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

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

Zimmermann, F., Möller, J., & Köller, O. (2018). When students doubt their teachers' diagnostic competence: Moderation in the internal/external frame of reference model. *Journal of Educational Psychology, 110,* 46–57. http://dx.doi.org/10.1037/edu0000196

Table 1

Goodness-of-fit Indices

Model Description	χ^2	df	CFA	TLI	RMSEA	SRMR
CFA with the French subsample[1]	297.500	122	.967	.949	.054	.043
CFA with the Latin subsample[1]	249.427	122	.978	.965	.047	.039
CFA with SFL subsamples as a grouping variable, configural invariance	548.210	244	.973	.957	.051	.042
CFA with SFL subsamples as a grouping variable, loading invariance[2]	570.305	256	.972	.958	.050	.044
I/E model with invariant factor loadings and invariant path coefficients across the SFL subsamples	597.383	272	.971	.959	.050	.050
I/E model with invariant factor loadings and invariant path coefficients for within-domain achievement–self-concept relations across the SFL subsamples	582.660	260	.971	.957	.051	.048
I/E model with invariant factor loadings and invariant path coefficients for cross-domain achievement–self-concept relations across the SFL subsamples	588.996	268	.971	.959	.050	.048

Note. CFA = confirmatory factor analyses; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean squared residual; SFL = second foreign language. All models were conducted with the MLR estimator. All χ^2 values are significant ($p < .001$). I/E model = internal/external frame of reference model.

[1] The CFA models for the French and Latin subsamples are statistically equivalent and thus result in the same fit as the models estimating the extended I/E model including multiple languages in the French and Latin subsamples.

[2] This model is statistically equivalent and thus results in the same fit as a model freely estimating the extended I/E model including multiple languages in both SFL subsamples.

Table 2

Path Coefficients from the Regression Models with the French and Latin Subsamples

Predictors	
	Outcome: German self-concept
German achievement	.651***/.680***
English achievement	-.102/.024
Math achievement	-.156***/-.201***
SFL achievement	.121*/.015
	Outcome: English self-concept
German achievement	-.180***/-.197***
English achievement	.731***/.806***
Math achievement	-.128**/-.149**
SFL achievement	.066/-.060
	Outcome: Math self-concept
German achievement	-.115***/-.157***
English achievement	-.101***/-.163***
Math achievement	.860***/.854***
SFL achievement	-.023/-.020
	Outcome: SFL self-concept
German achievement	.006/-.030
English achievement	-.064/-.109*
Math achievement	-.116**/-.050
SFL achievement	.753***/.832***

Note. All parameters are standardized. SFL = Second foreign language. The first coefficient refers to the French subsample, the second coefficient refers to the Latin subsample.
*** $p < .001$; ** $p < .01$; * $p < .05$.

Table 3

Path Coefficients from the Regression Models Stating Invariance of all Achievement–Self-concept Relations across SFL Groups

Predictors	
	Outcome: German self-concept
German achievement	.662***/.669***
English achievement	-.040/-.042
Math achievement	-.180***/-.190***
SFL achievement	.062/.077
	Outcome: English self-concept
German achievement	-.193***/-.173***
English achievement	.788***/.738***
Math achievement	-.141***/-.132***
SFL achievement	.001/.001
	Outcome: Math self-concept
German achievement	-.134***/-.130***
English achievement	-.131***/-.133***
Math achievement	.850***/.866***
SFL achievement	-.025/-.030
	Outcome: SFL self-concept
German achievement	-.014/-.012
English achievement	-.099**/-.091*
Math achievement	-.085**/-.078**
SFL achievement	.773***/.824***

Note. All parameters are standardized. SFL = Second foreign language. The first coefficient refers to the French subsample, the second coefficient refers to the Latin subsample.
*** $p < .001$; ** $p < .01$; * $p < .05$.

Supplemental Material

Social and Dimensional Comparisons in the Formation of German Students' Language Self-concepts

Table S1

Descriptive Statistics of the French and Latin Subsamples

	French Subsample	**Latin Subsample**
n	489	481
n boys	188 (38.4%)	254 (52.8%)
n girls	299 (61.1%)	227 (47.2%)
n students without indicated sex	2 (0.4%)	0 (0.0%)
Age (*M* and *SD*)	16.08 (1.027)	16.03 (0.897)
Years of Learning the SFL (*M* and *SD*)	4.57 (1.255)	4.85 (1.008)
Number of Schools	9	9
Number of Classes	63	61
n of Students with German as the mother tongue	449 (91.8%)	456 (94.8%)
n of Students attending Grade 9	115 (23.5%)	91 (18.9%)
n of Students attending Grade 10	175 (35.8%)	228 (47.4%)
n of Students attending Grade 11	199 (40.7%)	162 (33.7%)
M (SD) Grade point average	4.36 (0.653)	4.42 (0.649)
M (SD) School grade in German	4.28 (0.878)	4.36 (0.829)
M (SD) School grade in English	4.47 (0.904)	4.42 (0.894)
M (SD) School grade in the SFL	4.31 (1.018)	4.30 (1.144)
M (SD) School grade in math	4.19 (1.105)	4.31 (1.095)
M (SD) School grade in social studies	4.40 (0.837)	4.50 (0.800)
M (SD) School grade in physics	4.26 (0.952)	4.39 (0.928)
M (SD) School grades in biology	4.42 (0.910)	4.43 (0.887)
M (SD) School grades in chemistry	4.41 (0.950)	4.59 (0.951)
M (SD) School grades in geography	4.36 (0.901)	4.44 (0.848)
M (SD) School grades in history	4.35 (0.826)	4.45 (0.901)
Reliability of the German self-concept scale (α/ω)	.802/.834	.819/.845
Reliability of the English self-concept scale (α/ω)	.855/.886	.897/.895
Reliability of the math self-concept scale (α/ω)	.914/.914	.900/.903
Reliability of the SFL self-concept scale (α/ω)	.863/.864	.901/.906

Note. SFL = Second foreign language. α = Cronbach's alpha reliability estimate. ω = McDonald's omega reliability estimate.
School grades and grade point average are reversely coded so that higher values represent higher achievement levels.

Table S2

Standardized Factor Loadings of the Self-concept Measures

	French Subsample	Latin Subsample
German Self-concept		
I can achieve at most things in German.	.844	.768
Nobody's perfect but I'm just not good at German.	.792	.834
With some of the topics in German, I know from the start that I just won't get them.	.416	.522
I am good at German.	.883	.864
English Self-concept		
I can achieve at most things in English.	.888	.896
Nobody's perfect but I'm just not good at English.	.797	.823
With some of the topics in English, I know from the start that I just won't get them.	.652	.672
I am good at English.	.897	.894
Math Self-concept		
I can achieve at most things in math.	.921	.924
Nobody's perfect but I'm just not good at English.	.811	.788
With some of the topics in math, I know from the start that I just won't get them	.744	.695
I am good at math.	.925	.923
Second Foreign Language Self-concept		
I can achieve at most things in French/Latin.	.908	.947
Nobody's perfect but I'm just not good at French/Latin.	.666	.778
With some of the topics in French/Latin, I know from the start that I just won't get them.	.644	.679
I am good at French/Latin.	.893	.938

Note. For all $p < .001$.

Table S3

Factor Correlations of the Confirmatory Factor Analysis (CFA) Models for the French and Latin Subsamples

	German self-concept	English self-concept	Math self-concept	SFL self-concept	German achievement	English achievement	Math achievement
English self-concept	.129**/.232***						
Math self-concept	-.029/-.165**	-.082/-.119*					
SFL self-concept	.330***/.310***	.125**/.093	.101*/.141**				
German achievement	.602***/.626***	.163***/.157**	.098*/.061	.250***/.314***			
English achievement	.225***/.348***	.635***/.632***	.073*/-.029	.233***/.240***	.488***/.545***		
Math achievement	.065/.065	.043/-.032	.787***/.744***	.138***/.260***	.317***/.371 **	.280***/.267***	
SFL achievement	.291***/.286***	.262***/.145**	.193***/.186***	.686***/.746***	.414***/.507***	.434***/.456***	.359***/.421***

Note. SFL = Second foreign language. The first coefficient refers to the French subsample, the second coefficient refers to the Latin subsample.
*** $p < .001$; ** $p < .01$; * $p < .05$.

Table S4

Factor Correlations from the Regression Models with the French and Latin Subsamples

	German self-concept	English self-concept	SFL self-concept
English self-concept	.108*/.184**		
SFL self-concept	.231***/.266***	-.052/.094	
Math self-concept	.030/-.128*	-.069/.088	.120*/.055
	German achievement	English achievement	SFL achievement
English achievement	.488***/.545***		
SFL achievement	.414***/.507***	.434***/.456***	
Math achievement	.317***/.371***	.280***/.267***	.359***/.421***

Note. SFL = Second foreign language. The first coefficient refers to the French subsample, the second coefficient refers to the Latin subsample.
*** $p < .001$; ** $p < .01$; * $p < .05$.

Table S5

Factor Correlations from the Regression Models Stating Invariance of all Achievement–Self-concept Relations across SFL Groups

	German self-concept	English self-concept	SFL self-concept
English self-concept	.116*/.185**		
SFL self-concept	.229***/.252***	-.049/.082	
Math self-concept	.028/-.133*	-.070/.088	.120*/.056
	German achievement	English achievement	SFL achievement
English achievement	.488***/.545***		
SFL achievement	.417***/.507***	.437***/.453***	
Math achievement	.317***/.371***	.280***/.266***	.355***/.425***

Note. SFL = Second foreign language. The first coefficient refers to the French subsample, the second coefficient refers to the Latin subsample.
*** $p < .001$; ** $p < .01$; * $p < .05$.

Analyses without using students' classes as a cluster variable

Given that the models analyzed here contained more free parameters than number of clusters (classes), the standard errors of the model parameters may not be trustworthy. We therefore conducted all models without any cluster variables.

Table S6

Standardized Factor Loadings of the Self-concept Measures from CFA Models without Using Students' Classes as a Cluster Variable

	French Subsample	Latin Subsample
German Self-concept		
I can achieve at most things in German.	.844	.768
Nobody's perfect but I'm just not good at German.	.792	.834
With some of the topics in German, I know from the start that I just won't get them.	.416	.522
I am good at German.	.883	.864
English Self-concept		
I can achieve at most things in English.	.888	.896
Nobody's perfect but I'm just not good at English.	.797	.823
With some of the topics in English, I know from the start that I just won't get them.	.652	.672
I am good at English.	.897	.894
Math Self-concept		
I can achieve at most things in math.	.921	.924
Nobody's perfect but I'm just not good at English.	.811	.788
With some of the topics in math, I know from the start that I just won't get them	.744	.695
I am good at math.	.925	.923
Second Foreign Language Self-concept		
I can achieve at most things in French/Latin.	.908	.947
Nobody's perfect but I'm just not good at French/Latin.	.666	.778
With some of the topics in French/Latin, I know from the start that I just won't get them.	.644	.679
I am good at French/Latin.	.893	.938

Note. CFA = confirmatory factor analyses. For all $p < .001$.

Table S7

Factor Correlations of the CFA Models for the French and Latin Subsamples without Using Students' Classes as a Cluster Variable

	German self-concept	English self-concept	Math self-concept	SFL self-concept	German achievement	English achievement	Math achievement
English self-concept	.129**/ .232***						
Math self-concept	-.029/ -.165**	-.082/ -.119**					
SFL self-concept	.330***/ .310***	.125*/ .093	.101/ .141**				
German achievement	.602***/ .626***	.163***/ .157**	.098*/ .061	.250***/ .314***			
English achievement	.225***/ .348***	.635***/ .632***	.073/ -.029	.233***/ .240***	.488***/ .545***		
Math achievement	.065/ .065	.043/ -.032	.787***/ .744***	.138**/ .260***	.317***/ .371***	.280***/ .267***	
SFL achievement	.291***/ .286***	.262***/ .145**	.193***/ .186***	.686***/ .746***	.414***/ .507***	.434***/ .456***	.359***/ .421***

Note. CFA = confirmatory factor analyses. SFL = Second foreign language. The first coefficient refers to the French subsample (χ^2 (122) = 301.989, CFI = .966. TLI = .948, RMSEA = .055, SRMR = .043), the second coefficient refers to the Latin subsample (χ^2 (122) = 247.627, CFI = .978, TLI = .965, RMSEA = .046, SRMR = .039).
*** $p < .001$; ** $p < .01$; * $p < .05$.

Table S8

Path Coefficients from the Regression Models with the French and Latin Subsamples without Using Students' Classes as a Cluster Variable

Predictors	
	Outcome: German self-concept
German achievement	.651***/.680***
English achievement	-.102/.024
Math achievement	-.156*/-.201***
SFL achievement	.121*/.015
	Outcome: English self-concept
German achievement	-.180***/-.197***
English achievement	.731***/.806***
Math achievement	-.128**/-.149***
SFL achievement	.066/-.060
	Outcome: Math self-concept
German achievement	-.115***/-.157***
English achievement	-.101**/-.163***
Math achievement	.860***/.854***
SFL achievement	-.023/-.020
	Outcome: SFL self-concept
German achievement	.006/-.030
English achievement	-.064/-.109**
Math achievement	-.116**/-.050
SFL achievement	.753***/.832***

Note. All parameters are standardized. SFL = Second foreign language. The first coefficient refers to the French subsample (χ^2 (122) = 301.989, CFI = .966. TLI = .948, RMSEA = .055, SRMR = .043), the second coefficient refers to the Latin subsample (χ^2 (122) = 247.627, CFI = .978, TLI = .965, RMSEA = .046, SRMR = .039).

*** $p < .001$; ** $p < .01$; * $p < .05$.

Table S9

Factor Correlations from the Regression Models with the French and Latin Subsamples without Using Students' Classes as a Cluster Variable

	German self-concept	English self-concept	SFL self-concept
English self-concept	.108/.184**		
SFL self-concept	.231***/.266***	-.052/.094	
Math self-concept	.030/-.128*	-.069/.088	.120*/.055
	German achievement	English achievement	SFL achievement
English achievement	.488***/.545***		
SFL achievement	.414***/.507**	.434***/.456***	
Math achievement	.317***/.371***	.280***/.267***	.359***/.421**

Note. SFL = Second foreign language. The first coefficient refers to the French subsample, the second coefficient refers to the Latin subsample.
*** $p < .001$; ** $p < .01$; * $p \leq .05$.

Table S10

Goodness-of-fit Indices of the Invariance Models without Using Students' Classes as a Cluster Variable

Model Description	χ^2	df	CFA	TLI	RMSEA	SRMR
CFA with French and Latin subsamples as a grouping variable, configural invariance	549.567	244	.972	.957	.051	.042
CFA with French and Latin subsamples as a grouping variable, loading invariance	573.045	256	.971	.957	.051	.044
I/E model, freely estimated across French and Latin subsamples	573.045	256	.971	.957	.051	.044
I/E model, invariance across French and Latin subsamples	601.458	272	.970	.958	.050	.050

Note. CFA = Confirmatory factor analyses; I/E model = internal/external frame of reference model; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean squared residual. All models were conducted with the MLR estimator. All χ^2 values are significant ($p < .001$).

Table S11

Path Coefficients from the Regression Models Stating Invariance of all Achievement–Self-concept Relations across SFL Groups without Using Students' Classes as a Cluster Variable

Predictors	
	Outcome: German self-concept
German achievement	.662***/.669***
English achievement	-.040/-.042
Math achievement	-.180***/-.190***
SFL achievement	.062/.077
	Outcome: English self-concept
German achievement	-.193***/-.173 ***
English achievement	.788***/.738***
Math achievement	-.141***/-.132***
SFL achievement	.001/.001
	Outcome: Math self-concept
German achievement	-.134***/-.130***
English achievement	-.131***/-.133**
Math achievement	.850***/.866***
SFL achievement	-.025/ -.030
	Outcome: SFL self-concept
German achievement	-.014/ -.012
English achievement	-.099**/-.091**
Math achievement	-.085**/-.078**
SFL achievement	.773***/.824***

Note. All parameters are standardized. SFL = Second foreign language. The first coefficient refers to the French subsample, the second coefficient refers to the Latin subsample.
*** $p < .001$; ** $p < .01$; * $p < .05$.

Table S12

Factor Correlations from the Regression Models Stating Invariance of all Achievement–Self-concept Relations across SFL Groups without Using Students' Classes as a Cluster Variable

	German self-concept	English self-concept	SFL self-concept
English self-concept	.116*/.185**		
SFL self-concept	.229***/.252***	-.049/.082	
Math self-concept	.028*/ -.133*	-.070/.088	.120*/.056
	German achievement	English achievement	SFL achievement
English achievement	.488***/.545***		
SFL achievement	.417***/.507***	.437***/.453***	
Math achievement	.317***/.371***	.280***/.266***	.355***/.425***

Note. SFL = Second foreign language. The first coefficient refers to the French subsample, the second coefficient refers to the Latin subsample.
*** $p < .001$; ** $p < .01$; * $p \leq .05$.