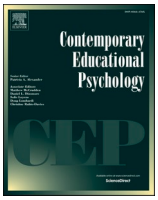
Contemporary Educational Psychology 68 (2022) 102037

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych



The generalized internal/external frame of reference model with academic self-concepts, interests, and anxieties in students from different language backgrounds

Lindie van der Westhuizen [a,*], A. Katrin Arens [b], Samuel Greiff [a], Antoine Fischbach [c], Christoph Niepel [a]

[a] Department of Behavioural and Cognitive Sciences, University of Luxembourg, Esch-sur-Alzette, Luxembourg
[b] DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany
[c] Luxembourg Centre for Educational Testing, University of Luxembourg, Esch-sur-Alzette, Luxembourg

ARTICLE INFO

Keywords:
Generalized internal/external frame of reference (GI/E) model
Dimensional comparisons
Academic self–concept
Interest
Anxiety

ABSTRACT

Student motivation and affect play an important role in successful language learning. To investigate the formation of language learning motivation and affect, this study extended the generalized internal/external frame of reference (GI/E) model framework to multiple languages (German and French, along with math) and multiple motivational-affective outcomes (academic self-concept, interest, and anxiety). We examined whether social and dimensional comparisons play similar roles in the formation of students' self-concepts, interests, and anxieties concerning different languages and whether dimensional comparisons result in contrast or assimilation effects. Moreover, we tested the generalizability of the GI/E model assumptions across students with different language backgrounds. Using a data set comprising virtually all ninth-grade students (N = 6275; 48.0% female) from Luxembourg's multilingual educational system, our findings indicated (1) clear contrast effects in the formation of self-concept and interest in math, German, and French, and (2) a combination of contrast, assimilation, and no effects in the formation of anxiety in math, German, and French. Using a subsample of 5837 students with valid language information (48.0% female), invariance tests demonstrated that the GI/E achievement–outcome relations operated equivalently across students from different home language backgrounds.

1. Introduction

Multilingualism has become commonplace among individuals and societies (Aronin, 2019). Educational practices have been reformed to reflect the global movement towards multilingualism (King, 2017). Most children today will need to learn and speak multiple languages at school and/or at home. The extent to which students successfully learn a language is (partly) determined by their motivation and affect (Dörnyei, 2002; Gardner, 2010; Lasagabaster et al., 2014). Motivation and affect for language learning is a multidimensional construct (Gardner & MacIntyre, 1992, 1993) encompassing academic self–concept (ASC), academic interest, and academic anxiety, among other factors (e.g., Brantmeier, 2006; Eidswick, 2009; Papi, 2010). Given the importance of motivation and affect in language learning, it is necessary to understand their formation to inform educational interventions seeking to positively influence student motivation and affect. To this end, the present study builds on the generalized internal/external frame of reference (GI/E) model (Möller, Müller-Kalthoff, Helm, Nagy, & Marsh, 2016), which is based on its precursor, the internal/external frame of reference (I/E) model (Marsh, 1986, 1990). The GI/E model assumes that social and dimensional comparisons are involved in the formation of motivation and affect in specific academic domains, including languages. The present study expands existing research on the GI/E model by (1) extending it to multiple language domains (German and French) in addition to the

* Corresponding author at: University of Luxembourg, Belval Campus, 2 Avenue de l'Université, L-4365 Esch-sur-Alzette, Luxembourg.
E-mail address: lindie.vanderwesthuizen@uni.lu (L. van der Westhuizen).

https://doi.org/10.1016/j.cedpsych.2021.102037

Available online 22 November 2021

domain of math; (2) extending it to multiple motivational–affective outcome variables beyond ASC (ASC, interest, and anxiety); and (3) testing the role of students' home language background on the relations assumed by the GI/E model by means of invariance testing. To this end, we drew on a large, diverse, and fully representative sample including virtually all ninth-grade students ($N = 6275$) from Luxembourg's multilingual educational system. To our knowledge, our study is one of the first to use the GI/E model framework to simultaneously investigate social and dimensional comparisons in the formation of ASC, interest, and anxiety in two language domains and math. We are also the first to examine the generalizability of the GI/E model across students from different home language backgrounds.[1]

2. The generalized internal/external frame of reference model

The GI/E model is as an overarching framework that describes the formation of domain–specific motivational–affective constructs. It is based on the original internal/external (I/E) frame of reference model (Marsh, 1986, 1990) explaining the formation of students' subjective beliefs about their abilities in different academic domains (i.e., ASC; Shavelson, Hubner, & Stanton, 1976). The original I/E model assumes that verbal and math ASCs are established through the joint operation of social and dimensional comparisons. Social comparisons occur when students compare their own achievement in one domain with the achievement of other students in the same domain. According to the I/E model, if an individual student perceives their math achievement to be lower than the math achievement of their classmates, their math ASC will also be lower (i.e., social comparisons). In dimensional comparisons, students compare their own achievement in one domain with their own achievement in another domain. Accordingly, when an individual student perceives their verbal achievement to be higher in comparison to their own math achievement, it (simultaneously) strengthens their verbal ASC while weakening their math ASC. In regression models, social comparisons are empirically supported by positive relations between achievements and ASCs from matching domains (e.g., math achievement and math ASC); dimensional comparisons are evidenced by negative relations between achievements and ASCs across nonmatching domains (e.g., math achievement and verbal ASC). Due to the relative weight of these two comparison processes, verbal and math achievements are positively correlated, while verbal and math ASCs are weakly correlated or approximate zero. The assumptions of the I/E model concerning the formation of verbal and math ASCs have been repeatedly supported in a diverse range of empirical studies, including experimental (Möller & Köller, 2001), longitudinal (Niepel, Brunner, & Preckel, 2014a), and meta–analytic studies (Möller, Zitzmann, Helm, Machts, & Wolff, 2020). Möller et al. (2020) reported average effect sizes of −0.20 and −0.17 for the cross-domain path coefficients (indicating dimensional comparison effects) and stronger average effect sizes of 0.46 and 0.57 for the within-domain path coefficients (i.e., indicating social comparison effects).

The I/E model stimulated interest in dimensional comparisons and led to the establishment of dimensional comparison theory (Möller & Marsh, 2013), which addresses the occurrence and effects of dimensional comparisons. Dimensional comparison theory serves as the theoretical foundation for empirical research on dimensional comparisons and led to the development of the GI/E model (Möller et al., 2016). The GI/E model, in turn, serves as the empirical framework for testing and further developing dimensional comparison theory. As such, the GI/E model permits the inclusion of domains beyond the verbal and math domains as well as the inclusion of outcome variables beyond ASC.

2.1. The extension of the GI/E model to multiple verbal domains

Recent extensions of the GI/E model expanded the verbal domain to include multiple languages. In the original I/E model, the verbal domain was represented by the language of instruction (or the native language of the country where the study took place; Möller, Pohlmann, Köller, & Marsh, 2009). To date, extensions have included first foreign languages (e.g., Marsh & Yeung, 2001) and second foreign languages (Arens et al., 2019).[2]

When examining domains other than a single math and a single verbal domain, two types of effects can emerge from dimensional comparisons: contrast or assimilation effects. In regression models, contrast effects are supported when the achievement–outcome relations across nonmatching domains are negative (like the cross–domain achievement–ASC relations in the original I/E model). Assimilation effects are indicated when the achievement–outcome relations across nonmatching domains are positive. The occurrence of contrast versus assimilation effects depends on the perceived similarity of the academic domains in question (Möller et al., 2016). When students perceive school subjects as similar, they may believe that achievement in these subjects relies on shared underlying abilities. In this case, assimilation effects are likely to occur. In contrast, when subjects are dissimilar (i.e., when achievement in two domains are believed to be unrelated), contrast effects are likely to occur. Similarly, the Marsh/Shavelson model of ASC (Marsh, 1990), posits that domain-specific ASCs can be placed on a continuum ranging from pure verbal to pure math. When comparing far or dissimilar domains (i.e., domains that are located farther apart on the continuum), contrast effects are likely to occur. When students engage in dimensional comparisons of similar or close domains, assimilation effects are likely to occur.

Thus, extending the GI/E model to include multiple language domains, language ASCs are likely to be located on the verbal endpoint of the verbal–math continuum and students may assume that common verbal abilities underlie language achievement in general. Moreover, students' proficiency in one language domain has been found to facilitate learning in another language domain (e.g., Kellerman, 1995; Prevoo, Malda, Mesman, & van IJzendoorn, 2016; Wawire & Kim, 2018). These considerations might lead to the assumption of assimilation effects among languages. In support of this, Marsh et al. (2014) found assimilation effects between English achievement and Dutch ASC in a Dutch sample of high school students, while Möller, Streblow, Pohlmann, and Köller (2006) found small and positive relations between English achievement and German ASC in a German sample. Other studies have found evidence of contrast effects between languages. For example, studies conducted with German students using German as the language of instruction and (typically) English as the first foreign language found contrast effects (e.g., Arens, Möller, & Watermann, 2016; Marsh, Lüdtke, et al., 2015; Niepel et al., 2014a). A contrast effect also emerged in a study with Luxembourgish secondary school students examining French and German achievement and ASCs (Brunner et al., 2010). In a Hong Kongese sample, Marsh, Kong, and Hau (2001) found a significant contrast effect between English achievement and Chinese ASC. Moreover, some studies have found no support for either contrast or assimilation effects between languages (e.g., Gaspard et al., 2018; Xu et al., 2013). Hence, studies based on the GI/E model examining two or more language subjects have found a mixture of contrast effects, assimilation effects, and nonsignificant effects, making further

[1] Drawing on our cross-sectional data, we also tested the following supplementary research aims: (a) whether domain-specific ASCs mediate the relation between domain-specific achievements on the one hand, and domain-specific interests and anxieties on the other hand; and (b) whether the aforementioned mediation model generalizes across home language groups. The theory, analyses, and findings related to the supplementary research aims are reported in the online supplementary material.

[2] The majority of these extensions to multiple verbal domains include German or Chinese as native language and English as additional language (see Table S1 in the online supplementary material for an overview).

investigation necessary.

2.2. Extension of the model beyond self–concept

In contrast to the original I/E model, which only included ASC as an outcome variable, the GI/E model allows for the inclusion of a wider range of outcome variables. For example, motivational–affective variables such as achievement goals (Dörendahl et al., 2021), student ratings of self–regulated learning (Miller, 2000), enjoyment (Goetz, Frenzel, Hall, & Pekrun, 2008), intrinsic motivation (Marsh, Abduljabbar, et al., 2015), intrinsic value (Arens & Preckel, 2018), interest (Köller, Daniels, & Baumert, 2000; Schurtz, Pfost, Nagengast, & Artelta, 2014), and anxiety (Arens, Becker, & Möller, 2017) have all been found to be affected by dimensional comparisons (see Wigfield, Eccles, & Möller, 2020 for a recent integration of dimensional comparisons with expectancies and values). Although the majority of the aforementioned studies generally found support for the typical pattern assumed by the GI/E model (i.e., social comparison effects in matching domains and contrast effects in nonmatching domains), most of them were restricted to a single math and single verbal domain (e.g., Arens et al., 2017; Nagy et al., 2008) or to math/science domains (e.g., Guo, Marsh, Parker, Morin, & Dicke, 2017; Nagy, Trautwein, Baumert, Köller, & Garrett, 2006; Marsh, Abduljabbar, et al., 2015). Still, Gaspard et al. (2018) examined a GI/E model with school grades (as predictors), expectancy (defined by ASC), intrinsic value (i.e., enjoyment gained from a task), attainment value (i.e., the importance of a task in relation to satisfying key life goals or values), utility value of the task (i.e., its perceived usefulness for future plans and goals), and cost (i.e., negative aspects resulting from engaging in a task) in the German, English, math, biology, and physics domains using a nonrepresentative sample of 857 German secondary school students. The results strongly supported contrast effects for the cross-domain achievement–outcome relations between math-like and verbal subjects. The majority of the cross-domain achievement–outcome relations between the two verbal domains were nonsignificant and close to zero. Overall, studies on the GI/E model integrating other motivational-affective variables in multiple verbal domains remain scarce, therefore warranting further investigation.

2.3. The role of student home language background

Previous studies extended the GI/E model to different languages by considering multiple languages learnt at school as additional domains (e.g., Arens, Helm, Wolff, & Möller, 2019; Gaspard et al., 2018; Xu et al., 2013). Given the prevalence of multilingualism in modern societies, it is surprising that differences in students' home language backgrounds have not been taken into account empirically in GI/E model studies to date. It has thus not been examined whether the relations between achievement and motivational–affective constructs, as assumed in the GI/E model, vary across students from different home language backgrounds. More specifically, it has not been established whether social and dimensional comparisons operate equally for students from monolingual and multilingual language backgrounds.

Students born into multilingual environments may have been learning a language for a longer period and may use their languages in a range of academic and nonacademic settings. They may therefore exhibit greater proficiency compared to students from monolingual environments, who often only acquire an additional language at school and whose use of the additional language may be restricted mostly to the academic context. Experiences in specific domains influence the formation of students' domain–specific ASCs, since more experience in certain domains may strengthen the domain specificity of ASCs (Marsh & Ayotte, 2003). Thus, multilingual students may notice the differences and similarities between different languages to a greater extent than monolingual students. This might cause multilingual students to differentiate more strongly between domain-specific ASCs, invoking stronger contrast effects of dimensional comparisons. However, one might also assume that if students identify greater similarities between two languages, stronger assimilation effects of dimensional comparisons are likely to occur. Thus, depending on whether multilingual students perceive greater differences or greater similarities between the languages in question, multilinguals may show either stronger contrast or stronger assimilation effects. This remains an open research question, creating a need to further investigate whether the achievement–outcome relations proposed in the GI/E model generalize across students from different language backgrounds.

In summary, the two advancements of the GI/E model (i.e., extension beyond the verbal and math domains and extension beyond ASC) have often been examined separately, but have rarely been investigated simultaneously in a single study (but see Gaspard et al., 2018). Moreover, multilingual students may demonstrate stronger (or weaker) dimensional comparisons due to greater exposure to different languages. However, it has yet to be examined whether the achievement–outcome relations proposed in the GI/E model vary across students from different language backgrounds. Thus, there is a need for studies on the GI/E model examining social and dimensional comparisons in the context of both extensions as well as examining the generalizability of these relations across students from different language backgrounds. The present study aimed to address these gaps in the literature.

2.4. The present study

The overarching aim of the present study was to determine whether social and dimensional comparisons are involved in the formation of language ASCs, interests, and anxieties. Such motivational-affective constructs play an important role in successful language learning (Dörnyei, 2002, 2003; Gardner, 2010). Insights into their formation can benefit educational practice and interventions by revealing how students' language motivation and affect should be targeted in order to contribute to successful language learning. Therefore, understanding the sources and formation of language motivation and affect by means of the GI/E model could have direct practical implications in a globalized society in which multilingualism is increasingly valued.

To this end, we extended current research on the GI/E model in three ways. First, we extended the GI/E model by simultaneously investigating two verbal domains, namely French and German, while also including math as a third domain, to gain further insight into contrast and assimilation effects of dimensional comparisons between languages (Research Aim 1). Second, we extended the GI/E model beyond ASC by simultaneously examining three outcomes, namely ASC, interest, and anxiety (Research Aim 2). Third, we tested the generalizability of the GI/E model across students from different home language backgrounds, which has, to our knowledge, never been tested to date (Research Aim 3).

To achieve these three research aims, we capitalized on the diversity of a large and fully representative sample consisting of virtually every ninth-grade student in Luxembourg in 2017 ($N = 6275$). Luxembourg, a country with an official multilingual educational policy, has three official languages: French, German, and Luxembourgish, with the latter considered linguistically very close to German (Gilles & Trouvain, 2013). Almost half (48%) of the population are foreign residents, of which 16% are Portuguese nationals (Grand Duchy of Luxembourg, 2019a). The three official languages are used to a varying extent in everyday life, while the most widely spoken home language is Luxembourgish (74%), followed by French (32%) and Portuguese (15%; Grand Duchy of Luxembourg, 2019b). Luxembourg thus provides an ideal setting to assess the invariance of the GI/E model across students from different language backgrounds.

We expected social comparisons to hold for the relations between achievement and ASC, interest, and anxiety within matching domains; we expected dimensional comparisons to hold for these same relations across nonmatching domains. Concerning the GI/E model relations across the verbal and math domains, in their meta-analysis, Möller et al.

(2020) observed a similar pattern of I/E model relations between math and first language domains and between math and second language domains (βs ranging between −0.20 and −0.15). Thus, we expected contrast effects between the math and (each of) the verbal domains for the relations between achievement and ASC. Against the background of the GI/E model, the anticipated contrast effects between math and language domains were also expected for the relations between achievement and interest and between achievement and anxiety (e.g., Arens et al., 2019; Brunner et al., 2010; Marsh et al., 2015; Marsh, Kong, & Hau, 2001).

We were particularly interested in the GI/E model relations between the two language domains (French and German) due to the mixture of contrast effects, assimilation effects, and nonsignificant effects that have been found in previous studies based on the GI/E model (e.g., Arens et al., 2019; Jansen, Schroeders, Lüdtke, & Marsh, 2015; Marsh et al., 2015; Möller et al., 2006). The majority of previous studies on the GI/E model have largely taken place in contexts that are conceptually and/or institutionally perceived as monolingual (e.g., Germany, Australia, USA; Adle & Beyer, 2017; Australian Government, 2021; see Möller et al., 2020). Our study, however, took place in a multilingual context in which multiple languages are used to a considerably greater extent in everyday life (see also Brunner et al., 2010).

Lastly, as Research Aim 3 is more exploratory in nature, we have no specific expectations regarding the generalizability of achievement–outcome relations across home languages. As outlined in more detail in the paragraphs above, students from multilingual home language backgrounds could show stronger contrast (or assimilation) effects between domain–specific language ASCs, interests, and anxieties (compared to students from monolingual home language backgrounds).

To address the aforementioned research aims, we used structural equation modeling (SEM; Bollen, 1989; Jöreskog & Sörbom, 1993) and specified a comprehensive GI/E model simultaneously containing domain-specific achievements, ASCs, interests, and anxieties across the math, German, and French domains (see Fig. 1 for a graphical representation).

3. Method

3.1. Sample and procedure

The data analyzed in the present study was obtained from the large-scale, Luxembourg School Monitoring Programme, the *Épreuves Standardisées* (ÉpStan; Fischbach, Ugen, & Martin, 2014; see also epstan.lu), which is implemented annually by the Luxembourg Centre for Educational Testing (LUCET). ÉpStan examines students' achievement of key educational outcomes every year at the beginning of each new learning cycle of mandatory schooling in Grades 1, 3, 5, 7, and 9. The ÉpStan has been approved by the national committee for data protection.

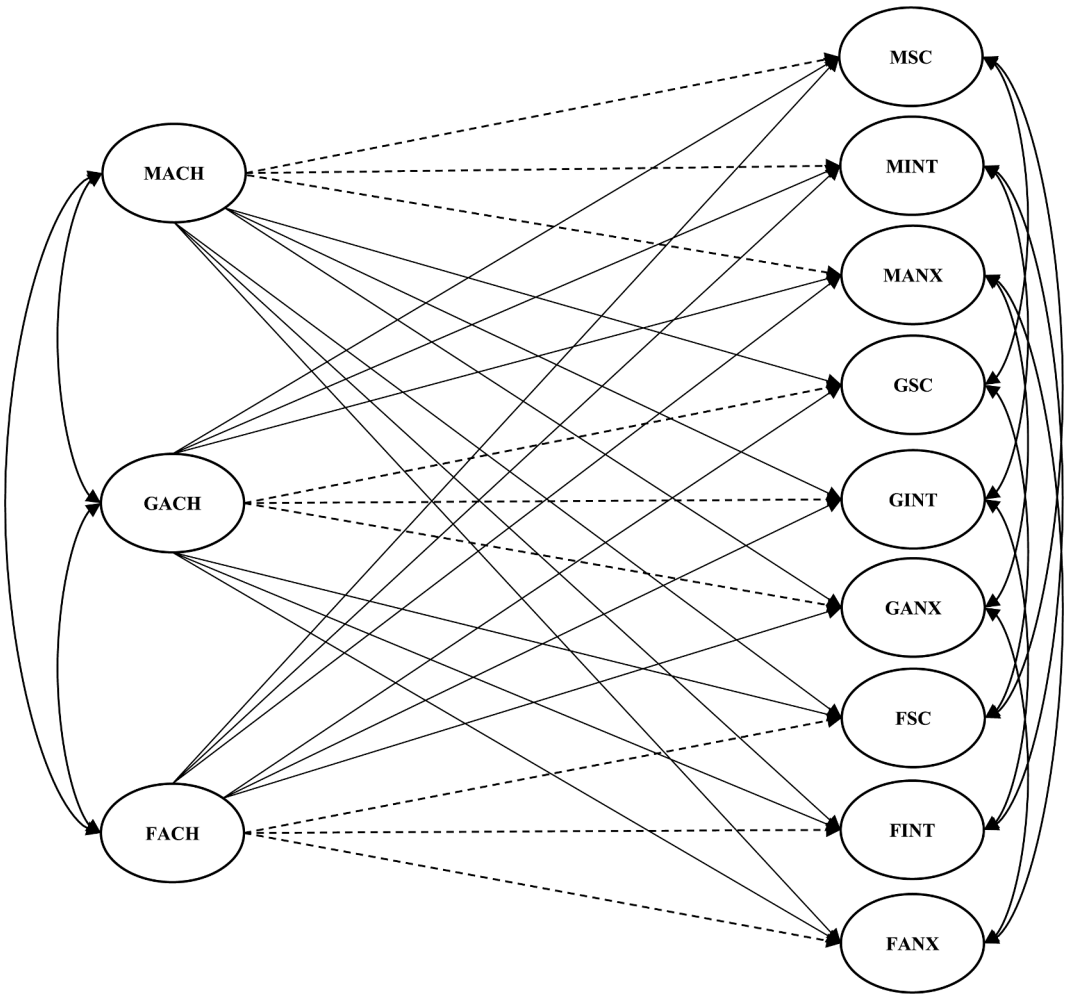


Fig. 1. The comprehensive generalized internal/external frame of reference (GI/E) model with self-concept, interest and anxiety as outcomes. *Note*. MACH = math achievement; GACH = German achievement; FACH = French achievement; MSC = math self-concept; GSC = German self-concept; FSC = French self-concept; MINT = math interest; GINT = German interest; FINT = French interest; MANX = math anxiety; GANX = German anxiety; FANX = French Anxiety. Dashed lines indicate social comparison paths (i.e., within-domain relations); solid lines indicate dimensional comparison paths (i.e., cross-domain relations).

Participation is compulsory for all students, however, participants and their parents or legal guardians were duly informed prior to data collection and were presented with the possibility to opt-out. A trusted-third-party-solution ensured the privacy of the participants and pseudonymized data was used in all statistical analyses.

A brief explanation of the multilingual nature of the Luxembourgish school system is required. Compulsory education starts with kindergarten at the age of four, where Luxembourgish serves as the language of instruction. At the age of six, primary school begins, and, in the first year, children begin to read and write in German. German is the language of instruction in primary school, regardless of whether the students' home language is Luxembourgish, German, French, or a different language. At the age of seven, in Grade 2, students start learning French in a content and language integrated learning (CLIL) approach. CLIL involves learning content through an additional language, thus teaching both the subject and the language (Coyle, 2007). By Grade 3, the teaching of French constitutes an integral component of the curriculum. In secondary school, the language of instruction shifts to French, although most subjects continue to be taught in German during the first years of secondary education—with the exception of math, which is taught in French from Grade 7 onwards. In the highest academic track secondary schools, the language of instruction for all subjects other than language courses progressively shifts to French, while German partially prevails in the vocationally-oriented secondary tracks.

The data for this study comprises the entire ninth-grade student cohort from 2017. The sample consisted of $N = 6325$ students (47.9% female) from 351 classes in 34 schools; the mean self-reported age was 15.01 years ($SD = 1.03$; median = 15.00; range = 12–25). The questionnaires used in our study as well as the achievement tests were administered in German and French, and students could choose their response language for each item. Only a small number ($N = 121$) of students switched between languages while responding to some of the motivational-affective items in the questionnaire. These students were removed from the sample to first test the measurement invariance between the different language versions, then reinserted after measurement invariance was established. A further 50 students were removed from the dataset as they had no responses on the ASC, interest, and anxiety items. This resulted in a final data set of $N = 6275$ students (henceforth called the "full sample"), which was used to investigate Research Aims 1 and 2. In the full sample, the majority (44.3%) of participants were native Luxembourgers (students who were born in Luxembourg and had at least one parent born in Luxembourg), while 24.2% were first generation immigrants (students born outside Luxembourg and whose parents were also born outside Luxembourg) and 31.3% were second generation immigrants (students born in Luxembourg but whose parents were born in another country). The International Socio-Economic Index of Occupational Status (ISEI-08; Ganzeboom, 2010; Ganzeboom & Treiman, 1996) was used to obtain an indication of students' socioeconomic status. ISEI scores range from 10 to 90, with higher values indicating higher socio-economic status. This information was obtained via student responses regarding their parents' occupation. For each student, the highest ISEI (HISEI) value of either the mother or father was used. The mean HISEI was 42.65 ($SD = 16.98$), with a range between 10 and 89. Luxembourg has three secondary school tracks. In the full sample, 12.6% ($N = 788$) of students were from the lower track, 59.4% ($N = 3724$) were from the intermediate track, and 28.1% ($N = 1763$) were from the highest track.

To address Research Aim 3, the generalizability of our GI/E models across students from different home language backgrounds, we allocated students to home language groups based on the languages spoken at home (i.e., self-reported language mostly spoken with their mother and/or father; see Table S2 in online supplementary material for a summary of the language groups). We first extracted students speaking one or more of the official languages and schooling languages at home, resulting in the following groups: (1) LUX/DE group ($N = 2145$), comprising students who speak Luxembourgish and/or German to both parents; (2) FR group ($N = 386$), comprising students who speak French to both parents; (3) FR-LUX/DE group ($N = 295$), a multilingual home language group encompassing students who speak French to one parent and Luxembourgish/German to the other parent; and (4) LUX/DE/FR-Other group ($N = 523$), a multilingual home language group encompassing students who speak Luxembourgish, German, or French to one parent and another language (excluding Luxembourgish, German, French, and Portuguese) to another parent. Next, given the size of the Portuguese-speaking population in Luxembourg, we extracted the two main Portuguese-speaking groups: (5) POR ($N = 1475$), comprising students who speak Portuguese to both parents; and (6) POR-LUX/DE/FR ($N = 309$), a multilingual home language group encompassing students who speak Portuguese to one parent and Luxembourgish, German, or French to the other parent. The remaining students were divided into two groups: (7) Other-Matched ($N = 485$), comprising students who speak the same language (excluding Luxembourgish, German, French, and Portuguese) to both parents; and Other-Mixed ($N = 219$), comprising students who speak a different language (excluding Luxembourgish, German, and French) with each parent. The sample used to investigate Research Aim 3 thus comprised 5837 students (henceforth called the "reduced sample" and excluded only 438 students due to insufficient or missing language information). See Table S3 in the online supplementary material for a comparison of the samples used to address Research Aims 1, 2, and 3.

3.2. Measures

Domain-specific ASC, interest, and anxiety were assessed for the domains of math, German, and French as part of the larger ÉpStan questionnaire. Each scale underwent extensive pilot testing prior to its implementation in the ÉpStan. ÉpStan scales use a four point Likert scale ranging from *disagree* (1) to *agree* (4). The questionnaires were administered on desktop and laptop computers. Descriptive statistics for the measures are shown in Table 1. The items are presented in Table S4 in the online supplementary material.

3.2.1. Academic self-concept

The three-item ASC scales we used were applied in Gogol, Brunner, Preckel, Goetz, and Martin (2016; McDonald's ω reliability estimates ranging from 0.89 to 0.91) and are also a short form of the five item scale used by Brunner et al. (2010) and the four item scale by Gogol, Brunner, Martin, Preckel, and Goetz (2017). The items were based on the Self-Description Questionnaire (SDQ; Marsh & O'Neill, 1984) and adapted

Table 1
Descriptive statistics based on the full sample ($N = 6275$).

	N	*M*	*SD*	Reliability coefficient
Math self-concept	6268	2.631	0.921	0.899
German self-concept	5820	2.816	0.874	0.881
French self-concept	6272	2.566	0.960	0.918
Math interest	6240	2.684	0.928	0.881
German interest	5803	2.607	0.884	0.885
French interest	6261	2.641	0.920	0.871
Math anxiety	6262	2.244	0.896	0.805
German anxiety	5816	1.927	0.841	0.815
French anxiety	6262	2.110	0.881	0.819
Math achievement	6269	510.800	97.784	0.796
German achievement	5826	525.380	101.340	0.838
French achievement	6271	511.330	118.412	0.839

Note. Means for the self-concept, interest, and anxiety scales were calculated by summing the item scores and dividing the sum by the number of items. Listwise deletion in SPSS was employed when calculating these descriptive statistics. The number of observations for each scale/indicator may thus vary. McDonald's ω reliabilities are reported for the self-concept, interest, and anxiety scales. Responses on the self-concept, interest, and anxiety scales were measured on a 4-point Likert scale. Weighted mean likelihood estimation (WLE) reliabilities are reported for the achievement scales.

for our three academic domains based on the instructions of Marsh (1990). Three-item ASC scales derived from the SDQ have been shown to possess acceptable psychometric properties for use in educational research (Gogol et al., 2014) and are frequently used in ASC research (e.g., Marsh et al., 2018; Möller, Retelsdorf, Köller, & Marsh, 2011; Niepel, Brunner, & Preckel, 2014b). Items were formulated in parallel across the three domains of math, German, and French (e.g., "I am good at mathematics/German/French"). McDonald's ω reliability estimates for the math, German, and French ASC scales were good (Table 1).

3.2.2. Academic interest

We used the three-item interest questionnaire applied by Gogol et al. (2016) and Gogol et al. (2017; McDonald's ω reliability estimates ranging from 0.86 to 0.94). One item in the scale assessed feelings of personal importance, another item assessed emotional value, and the last item was a global interest item that sought to optimally represent the essence of academic interest. Items were worded analogously across the three domains (e.g., "I am interested in mathematics/German/French"). McDonald's ω reliabilities for the interest scales were good for math, German, and French (Table 1).

3.2.3. Academic anxiety

The three-item anxiety scale used in this study was applied in Gogol et al. (2016) and Gogol et al. (2017; McDonald's ω reliability estimates ranging from 0.78 to 0.83). It is based on the Hebrew adaptation of the Test Anxiety Inventory (Spielberger, 1980; Zeidner, Nevo, & Lipschitz, 1988) and the Academic Emotions Questionnaire (AEQ; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). The items for math, German, and French had parallel wording (e.g., "I am afraid of mathematics/German/French class"). The McDonald's ω reliabilities for the anxiety scales were good for math, German, and French (Table 1).

3.2.4. Academic achievement

Academic achievement in math, German, and French was assessed with the math, German, and French ÉpStan achievement tests, which were developed in accordance with the official national curriculum of the Luxembourg Ministry of Education. The tests are used annually to provide feedback to students, teachers, schools, and the Ministry of Education on Luxembourg's educational outcomes (University of Luxembourg & SCRIPT, 2018). Pre-testing and subsequent psychometric evaluation are conducted annually prior to the ÉpStan to validate the content, format, and feasibility of the test material. Testing for overlap between theoretical and empirical item difficulty, differential item functioning, and Rasch compliance is also conducted. There are different test versions for each subject for use in the different school tracks. To allow for comparability across academic tracks, the test material comprises a certain minimal number of anchoring tasks included in the different achievement tests created for each subject for each of the three academic tracks. Responses on the achievement tests were scaled in R (R Core Team, 2019) with the TAM package (Robitzsch, Kiefer, & Wu, 2019) using a unidimensional Rasch model. The analyses reported here were conducted using weighted likelihood estimates (WLE) scores (Warm, 1985). Higher scores on the achievement tests indicated higher achievement. WLE reliability estimates were also calculated with the TAM package. Good WLE reliability estimates were obtained for math, German, and French achievement (Table 1). More information on the content of the achievement tests are presented in Table S5 in the online supplementary material.

3.3. Statistical analyses

All analyses were performed using Mplus 8.3 (Muthén & Muthén, 1998–2019). Robust maximum likelihood estimation (MLR) was used to estimate all models. MLR is robust against violations of normality assumptions (Beauducel & Herzberg, 2006). As students were nested in classes, the "type = complex" command in Mplus was used to correct for possibly biased standard errors (Raudenbush & Bryk, 2002). Full information maximum likelihood (FIML), a reliable method for treating missing data, was used to account for missing values (Enders, 2010). Missing values ranged between 0.35% and 8.27% for interest and anxiety items, between 0.14% and 8.57% for ASC items, and between 0.06% and 7.16% for achievement test scores. Correlated uniquenesses were specified between items which were worded in parallel across domains to capture shared variances due to item wording (Marsh et al., 2013).

A series of confirmatory factor analysis (CFA) models were run first to assess the integrity of the ASC, interest, anxiety, and achievement measures (Brown, 2006). As students could respond to these measures in French or German, we tested whether the assumption of measurement invariance holds across the German-language and French-language versions (Millsap, 2011) so that we could use both language versions simultaneously in the subsequent analysis. The same procedure was used to test weak measurement invariance across the two language versions as when testing for measurement invariance across home language backgrounds, which is described later in this section.

We next tested a series of structural GI/E models, ranging from models containing only ASC (as outcome) and achievement (as predictor) across three domains (math, German, and French) to a more comprehensive model (Fig. 1) simultaneously containing ASC, interest, and anxiety (as outcomes), as well as achievement (as predictor) across three domains. The domain-specific ASC, interest, and anxiety factors were defined by their respective sets of items. Domain-specific achievements were included as factors with standardized single items as indicators, assuming no measurement error.

A combination of goodness-of-fit indices were considered to evaluate the fit of the CFA models. The comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) were considered. We used the following criteria by Hu and Bentler (1999) to deem a model as having a good fit: CFI and TLI > 0.95, RMSEA < 0.06, and SRMR < 0.08.

To investigate the generalizability of the GI/E model relations across the eight home language groups, we proceeded by testing a series of increasingly restrictive invariance models (Meredith, 1993; Millsap, 2011). First, it was necessary to establish weak measurement invariance (i.e., invariance of the factor structure and factor loadings; Millsap, 2011). We therefore compared a model in which all parameters were freely estimated across groups (i.e., configural invariance or invariance of the factor structure) to a model in which factor loadings were constrained to the same value across groups (weak or metric invariance). Evidence of weak measurement invariance implies that the different groups responded to the items in the same way. This is a precondition for the next step, namely, testing for the invariance of factor (i.e., GI/E model) relations (Millsap, 2011). To this end, we first estimated the comprehensive GI/E structural model (Fig. 1) freely across the eight subsamples. We then compared this model to a model in which the achievement–outcome relations were constrained to be equal across groups.

To evaluate the invariance tests, we followed the recommendations for satisfactory change in model fit statistics by Cheung and Rensvold (2002) and Chen (2007): ΔCFI $\geq$ –0.010, ΔRMSEA ≤ 0.015, and ΔSRMR ≤ 0.030. We used the ΔCFI as the main criterion as RMSEA and SRMR are more likely to be affected by sample size and model complexity and tend to overreject invariant models when the sample size is small (Chen, 2007).

4. Results

4.1. Preliminary analyses

Weak measurement invariance could be demonstrated across the two language versions of the instruments (Table 2, Models 1–3). Consequently, the sample was pooled and the analyses continued on the full sample (i.e., $N = 6275$, including the 121 language switchers). A CFA

Table 2
Invariance across questionnaire language version and students' home language background.

Model		*N*	χ^2	*df*	CFI	TLI	RMSEA	SRMR	Com-parison	ΔCFI	ΔRMSEA/ ΔSRMR
1a	German sample: CFA with self-concept, interest, anxiety and achievement	4344	2427.437	315	0.968	0.956	0.039	0.035			
1b	French sample: CFA with self-concept, interest, anxiety and achievement	1810	1068.094	315	0.967	0.954	0.036	0.038			
2	Configural invariance across German and French language version	6154	3496.477	630	0.968	0.955	0.038	0.036			
3	Weak invariance across German and French language version	6154	3615.362	648	0.967	0.955	0.039	0.037	2 vs. 3	−0.001	+0.001/ +0.001
4	Configural invariance across eight language groups	5837	5840.189	2520	0.962	0.948	0.042	0.042			
5	Invariant factor loadings (weak invariance) across eight language groups	5837	6142.005	2667	0.960	0.948	0.042	0.044	4 vs. 5	−0.002	0.000/ +0.002
6	GI/E model with invariant factor loadings and free path coefficients across eight language groups	5837	6142.003	2667	0.960	0.948	0.042	0.044			
7	GI/E model with invariant factor loadings and invariant path coefficients across eight language groups	5837	6538.023	2856	0.958	0.949	0.042	0.051	6 vs. 7	−0.002	0.000/ +0.007

Note. *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean squared residual. All models integrated correlated uniquenesses between parallel-worded items. All χ^2 values are significant ($p < .001$).

model containing ASC, interest, anxiety, and achievement in all three domains (math, German, and French) obtained a good fit [χ^2 (315) = 3465.292; CFI = 0.968; TLI = 0.955; RMSEA = 0.040; SRMR = 0.039], attesting to the integrity of the measures. The ASC, interest, and anxiety items had significant factor loadings on their respective factors (λs ranging from 0.561 to 0.908, *p*s < 0.001; see Table S4 in the online supplementary material). Domain-specific achievements were positively correlated (r = 0.501 to 0.589), as were domain-specific anxieties (r = 0.350 to 0.486) and interests (r = 0.067 to 0.234; see Table 3). The intercorrelations for domain-specific ASCs ranged from –.328 to .015.

4.2. Extending the GI/E model to multiple domains and outcomes

To address Research Aims 1 and 2, we tested the comprehensive GI/E structural model (Fig. 1) in which math, German, and French ASC, interest, and anxiety (all outcome variables) were regressed on math, German, and French achievement. The model obtained a good fit [χ^2 (315) = 3465.292; CFI = 0.968; TLI = 0.955; RMSEA = 0.040; SRMR = 0.039].[3] The typical I/E model relations involving the paths between verbal and math achievement and ASC could be replicated (see Table 4): The within-domain paths from achievement to ASC were significantly positive (βs ranging from 0.437 to 0.588, *p*s < .001). This suggests a social comparison effect, as students who exhibited higher achievement in one domain (e.g., math) also exhibited higher ASC in the same corresponding domain. Across domains, math achievement was negatively and significantly related to German ASC (β = −0.083, $p < .001$) and French ASC (β = −0.157, $p < .001$). German achievement was found to be negatively and significantly related to math ASC (β = −0.175, $p < .001$), and French achievement was found to be negatively and significantly related to math ASC (β = −0.118, $p < .001$). This implies that students with higher math achievement exhibited lower verbal ASC and that students with higher verbal achievement exhibited lower math ASC. With regard to the relations between the French and German domains, German achievement was found to be negatively and significantly related to French ASC (β = −0.389, $p < .001$), while French achievement was found to be negatively and significantly related to German ASC (β = −0.354, $p < .001$), indicating a contrast effect. It appears that higher achievement in one language (e.g., French) was associated with students experiencing a lower ASC in the other language (e.g., German).

Concerning the extension of the GI/E model to interest as an outcome (Research Aim 2), the within-domain paths between achievement and interest were also significant and positive, albeit descriptively somewhat smaller in magnitude compared to the relations between achievement and ASC (βs ranging from 0.286 to 0.471, *p*s <0.001). This indicates the operation of social comparison effects, as students with higher achievement in one domain experienced higher interest in that same domain. Examination of the cross-domain paths showed that math achievement was negatively and significantly related to German interest (β = −0.158, $p < .001$) and French interest (β = −0.190, $p < .001$). German achievement was significantly and negatively related to math interest (β = −0.173, $p < .001$), and French achievement was negatively and significantly related to math interest (β = −0.069, $p < .001$). This implies that higher achievement in math was associated with students experiencing lower interest in each of the two languages, while higher achievement in each of the two languages was associated with lower interest in math. With regard to the two languages, German achievement was significantly and negatively related to French interest (β = −0.352, $p < .001$), while French achievement was negatively and significantly related to German interest (β = −0.211, $p < .001$). This contrast effect suggests that higher achievement in one language was associated with students experiencing lower interest in the other language.

Importantly, when considering the relations between achievement and anxiety (Research Aim 2), the signs of the GI/E model relations are reversed: Negative within-domain paths indicate social comparisons, positive cross-domain paths indicate contrast effects, and negative cross-domain paths indicate assimilation effects. For anxiety, significantly negative relations were observed within the domains of math, German, and French (βs ranging from −0.441 to −0.251, *p*s < 0.001). Therefore, in corresponding domains, higher achievement is linked to lower anxiety, suggesting the operation of social comparison effects. The cross-domain achievement–anxiety paths showed that math achievement was negatively and significantly related to German anxiety (β = −0.050, $p = .008$), but not related to French anxiety (β = 0.004, $p = .811$). Furthermore, French achievement was found to be significantly and positively related to math anxiety (β = 0.175, $p < .001$), but the relation between German achievement and math anxiety was nonsignificant (β = 0.029, $p = .153$). This implies that higher achievement in math was related to *lower* anxiety in German (but not to higher/lower anxiety in French) and that higher achievement in French was associated with *higher* anxiety in math. Lastly, with regard to the two languages, French achievement was found to be significantly and positively related to German anxiety (β = 0.297, $p < .001$), and German achievement was significantly and positively related to French anxiety (β = 0.150, $p < .001$). This suggests that higher achievement in one language (e.g., French) was associated with students experiencing higher anxiety in the other language (e.g., German), indicating a contrast effect. Overall, the

[3] The CFA measurement model and SEM regression model (testing the GI/E model) are statistically equivalent, as the factor correlations (in the CFA model) are replaced by path coefficients (in the SEM model).

Table 3
Standardized latent correlations.

	1 MSC	2 GSC	3 FSC	4 MINT	5 GINT	6 FINT	7 MANX	8 GANX	9 FANX	10 MACH	11 GACH	12 FACH
1	—	−0.046*	0.015	0.818**	0.013	0.040*	−0.523**	0.094**	0.050**	0.275**	0.022	0.011
2		—	−0.328**	0.009	0.756**	−0.254**	0.087**	−0.544**	0.232**	0.052**	0.301**	−0.123**
3			—	0.056**	−0.179**	0.802**	0.176**	0.376**	−0.381**	−0.091**	−0.180**	0.310**
4				—	0.210**	0.234**	−0.275**	0.198**	0.154**	0.150**	−0.040**	−0.015
5					—	0.067**	0.185**	−0.207**	0.282**	−0.095**	0.086	−0.143**
6						—	0.246**	0.455**	−0.183**	−0.162**	−0.223**	0.196**
7							—	0.453**	0.486**	−0.230**	−0.078**	0.023
8								—	0.350**	−0.161**	−0.318**	0.045*
9									—	−0.034*	0.023	−0.173**
10										—	0.589**	0.501**
11											—	0.512**
12												—

Note. Latent correlations are provided above the diagonal and were calculated using full information maximum likelihood (FIML) estimation. MSC = Math self-concept; GSC = German self-concept; FSC = French self-concept; MINT = Math interest; GINT = German interest; FINT = French interest; MANX = Math anxiety; GANX = German anxiety; FANX = French anxiety; MACH = Math achievement; GACH = German achievement; FACH = French achievement.
* $p < .05$; ** $p < .01$.

results indicate that the GI/E model could successfully be extended to math and two verbal domains as well as to interest and anxiety as outcomes (Research Aims 1 and 2).

4.3. The role of students' home language background

To address Research Aim 3, we next tested for weak invariance (i.e., factor loading invariance in the CFA model containing achievement, ASC, interest, and anxiety across three domains) across students from the eight home language backgrounds (Model 5 in Table 2). Compared to the model of configural invariance (Model 4 in Table 2), the ΔCFI, ΔRMSEA, and ΔSRMR values did not exceed the recommended cut-offs of −0.010, 0.015, and 0.030, respectively, indicating weak invariance. Next, we estimated the comprehensive GI/E model (Fig. 1) freely across the eight subsamples (Model 6 in Table 2). When comparing this freely estimated comprehensive GI/E model to a model in which the path coefficients for all achievement–ASC, achievement–interest, and achievement–anxiety relations (within and across domains) were constrained to be equal (Model 7 in Table 2), the ΔCFI, ΔRMSEA, and ΔSRMR values changed by −0.002, 0.000, and 0.007, respectively, indicating that the GI/E model relations for ASC, interest, and anxiety were found to be invariant across students from different home language backgrounds.[4]

5. Discussion

Most extensions of the I/E model within the GI/E model framework either expand the number of academic domains considered or expand the outcome domains to constructs beyond ASC. Rarely do they include both extensions (see Arens & Preckel, 2018; Gaspard et al., 2018 for exceptions) and such an examination using a diverse and fully representative sample has been missing to date. The extension of the GI/E model to two language domains is particularly important, since the evidence is rather mixed with regard to the existence of contrast versus assimilation effects between languages. In addition, the generalizability of the functional GI/E model relations has not yet been established across students from different language backgrounds. In our study, we simultaneously extended the GI/E model to multiple academic domains (math, German, and French) and multiple outcome domains (ASC, interest, and anxiety), particularly focusing on the interplay between achievement and outcomes across the two language domains. Furthermore, we examined whether the obtained GI/E model results generalized across students from different home language backgrounds.

5.1. The GI/E model extensions and generalizability across language backgrounds

5.1.1. Extension to multiple domains and outcomes

In general, our results supported the role of social comparisons in the formation of domain–specific ASCs, interests, and anxieties as positive relations between achievement and ASC and interest, and negative relations between achievement and anxiety within the same domain were found—irrespective of the domain considered. Contrast effects were found for achievement–ASC relations between the verbal and math domains, thereby replicating the assumptions of the original I/E model. Moreover, contrast effects of dimensional comparisons also appear to characterize achievement–interest relations for the cross–domain paths between math and French as well as between math and German. Thus, contrast effects of dimensional comparisons can be replicated for the achievement–interest relations between math and two verbal domains. This finding is in line with previous research on the (G)I/E model (e.g., Arens & Preckel, 2018, Gaspard et al., 2018; Goetz et al., 2008; Möller et al., 2006; Schurtz et al., 2014) as well as theory about far comparisons within the Marsh/Shavelson model of ASC (Marsh, 1990) and dimensional comparison theory's assumptions of dissimilar underlying abilities (Möller & Marsh, 2013).

With regard to the two language domains, contrast effects appear to operate between German achievement, on the one hand, and French ASC and French interest on the other, as well as between French achievement, on the one hand, and German ASC and German interest on the other. This contradicts the nonsignificant effects found by Gaspard and colleagues (2018) for the achievement–intrinsic value relations between German (native language) and English (foreign language) domains. Overall, our results suggest that dimensional comparisons are involved in the formation of verbal ASCs and interests and lead to contrast effects rather than assimilation effects. Our results are in line with a growing number of studies to find evidence supporting contrast effects across verbal domains (e.g., Arens et al., 2016; Brunner et al., 2010; Marsh, Lüdtke, et al., 2015; Niepel et al., 2014a). The recent meta-analysis by Möller et al. (2020) found small significant contrast effects for achievement–ASC relations between first language (L1) and second language (L2) domains (β = −0.10; β = −0.13). We observed much stronger contrast effects between the two verbal domains in our sample (ASC: β = −0.389 to −0.354; interest: β = −0.352 to −0.211). The sociocultural context in Luxembourg could account for these relatively strong contrast effects (see also Brunner et al., 2010). The relations among verbal ASCs may depend on the importance of the languages

[4] We also tested the invariance of the comprehensive GI/E model across the three school tracks using the full sample ($N = 6275$). The GI/E model relations were found to be invariant across the three tracks. See Models I1 to I4 in Table S6 in the online supplementary material.

Table 4
Standardized path coefficients for the comprehensive GI/E model.

Predictor	Outcome								
	Self-concept			Interest			Anxiety		
	German	French	Math	German	French	Math	German	French	Math
Achievement: German	0.531 (0.02)**	−0.389 (0.02)**	−0.175 (0.02)**	0.286 (0.02)**	−0.352 (0.02)**	−0.173 (0.02)**	−0.441 (0.02)**	0.150 (0.02)**	0.029 (0.02)
Achievement: French	−0.354 (0.02)**	0.588 (0.01)**	−0.118 (0.02)**	−0.211 (0.02)**	0.471 (0.02)**	−0.069 (0.02)**	0.297 (0.02)**	−0.251 (0.02)**	0.175 (0.02)**
Achievement: Math	−0.083 (0.02)**	−0.157 (0.02)**	0.437 (0.02)**	−0.158 (0.02)**	−0.190 (0.02)**	0.287 (0.02)**	−0.050 (0.02)**	0.004 (0.02)	−0.335 (0.02)**
R^2	0.199	0.266	0.115	0.070	0.201	0.051	0.162	0.047	0.079

Note. The comprehensive GI/E model includes achievement in math, German, and French as predictor variables, and academic self-concept, interest and anxiety in math, German, and French as outcome variables. The standard errors of the path coefficients are reported in parentheses.

** $p < .01$.

learnt at school in the specific society and curriculum under investigation. German and French are both essential in Luxembourgish society and the school curriculum, consequently representing salient sources of information for dimensional comparisons (Marsh & Yeung, 2001; Marsh et al., 2001). When only one language is highly salient in a society (Yeung & Lau, 1998) or the school context (Xu et al., 2013), verbal ASCs may not be subject (as strongly) to dimensional comparisons. Furthermore, the stronger contrast effects might also be due to the gradual shift in language of instruction from German to French in the seventh to ninth grade. Increased exposure to French in other academic domains could more acutely highlight differences between French and German, thereby decreasing students' perceived subject similarity. This may, in turn, engender stronger contrast effects of dimensional comparisons (Helm, Müller-Kalthoff, Nagy, & Möller, 2016; Möller et al., 2016). Indeed, in their examination of the differential stability of general and domain-specific components of ASC, interest, and anxiety in a sample of Luxembourgish students from the seventh to the ninth grade, Gogol et al. (2016) found that although French and German ASC, interest, and anxiety were moderately stable, students' profile differences between French and German became greater over this period (e.g., students with high French ASC retained their level of French ASC, but developed weaker German ASC). Altogether, our study adds to mounting evidence contradicting the assumption of assimilation effects between verbal domains posited by the Marsh/Shavelson model of ASC (Marsh, 1990) and by dimensional comparison theory (Möller & Marsh, 2013).

While our pattern of findings is consistent for the relations between achievement and ASC as well as achievement and interest, the results for the relations between achievement and anxiety are more diverse. At first glance, the findings from our study do not match previous GI/E model studies involving anxiety (Arens et al., 2017). However, closer inspection reveals that the contrast effect between French achievement (the language of instruction in our study) and math anxiety is similar to the result found by Arens et al. (2017) in their GI/E model including anxiety. More specifically, Arens and colleagues observed a contrast effect between German achievement (the language of instruction in their study) and math anxiety. We also found that the paths between math achievement and French anxiety and between German achievement and math anxiety were nonsignificant. Similarly, Arens et al. (2017) also found a nonsignificant relation between math achievement and German anxiety. The significant assimilation effect we observed for the relation between math achievement and German anxiety should not be overinterpreted, as the effect size was small. Similarly, the assimilation effect observed between math achievement and German anxiety was small. With regard to the two language domains, our study was, to the best of our knowledge, the first to examine achievement–anxiety relations between two languages. We found these relations to be characterized by contrast effects. The contrast effects for the cross-domain achievement–anxiety paths between the two languages is akin to the pattern of contrast effects found for the achievement–ASC and achievement–interest relations across the two languages. Overall, all cross-domain relations (for all outcomes) between the two language domains represented contrast effects, as did the majority of the cross-domain relations (for all outcomes) between math and the verbal domains and between the verbal domains and math (with the exception of the path between math achievement and German anxiety, the path between math achievement and French anxiety, and the path between German achievement and math anxiety).

Lastly, while a few studies have found evidence of contrast effects for achievement–ASC relations between languages, our study is one of the first to simultaneously investigate and find contrast effects for achievement–interest and achievement–anxiety relations between two languages. We thus emphasized the importance of considering contrasting dimensional comparison effects when examining students' language learning motivation and affect.

5.1.2. Generalizability across home language backgrounds

Another major contribution and unique feature of our study addresses the generalizability of the GI/E model across students from different home language backgrounds. Apparently, only Marsh and Yeung (2001) have previously hinted at the potential impact of students' home language backgrounds on GI/E model relations given their study's heterogeneous ethnic sample, but they did not empirically test for these differences. The findings from our study showed invariant achievement–ASC, achievement–interest, and achievement–anxiety relations for math, German, and French for students from a variety of monolingual and multilingual home language backgrounds. For the first time, we have thus shown that the GI/E model applies similarly to students from different language backgrounds and that students' home language background therefore does not seem to affect the GI/E model relations between achievement, ASC, interest, and anxiety. Beyond student home language background, previous research has also examined the generalizability of the (G)I/E model across language of instruction. In particular, Xu et al. (2013) examined differences in the GI/E model relations between Hong Kongese students with either English or Chinese language of instruction. Similar to our finding of equivalence, Xu et al. (2013) also found that their I/E model was fully invariant across the two different languages of instruction. The existing findings thus seem to indicate invariance across student home and school language background, thereby supporting the high level of generalizability of the GI/E model assumptions and the universality of social and dimensional comparisons in the formation of motivational-affective variables. However, research is scarce and further studies are needed.

5.2. Limitations, strengths and directions for future research

Our study adds to the literature on the role of social and dimensional comparisons in the formation of verbal ASCs, interests, and anxieties. A few limitations should, nevertheless, be mentioned. The cross–sectional nature of our study does not allow for causal inferences. More longitudinal research is needed to examine the role of social and dimensional comparisons in the formation of domain-specific achievements, ASCs, interests, and anxieties over time (e.g., Niepel et al., 2014a). Longitudinal designs should also be used to test whether ASC mediates the relation between achievement and interest and between achievement and anxiety. The mediating role of ASC is in line with the assumptions of expectancy–value theory (Wigfield & Eccles, 2000) and control-value theory (Pekrun & Stephens, 2010; Pekrun, 2006) and has been empirically validated in numerous studies (Arens et al., 2017; Gaspard et al., 2018; Goetz et al., 2008; Guo et al., 2017; Köller et al., 2000; Schurtz et al., 2014). These studies were, however, predominantly cross-sectional in nature and thus not fully appropriate for testing the causal ordering implicit in mediation (Mitchell & Maxwell, 2013). Longitudinal research would provide stronger tests of the mediating role of ASC in the relations between achievement, interest, and anxiety.

Our interest and anxiety scales consisted of three items each and did not distinguish between different subfacets of each construct. For example, anxiety is typically composed of a worry and an emotionality subfacet (Liebert & Morris, 1967; Zeidner, 2007), while interest is often differentiated into a personal importance subfacet and emotional value subfacet (Krapp, 2002; Renninger, 2000; Schiefele, 1991). Our short measures may fail to capture the differential subfacet relations that may potentially exist with achievement within and across domains (see Arens et al., 2017; Guo et al., 2017). Thus, future studies should include more comprehensive motivational–affective instruments to better measure these relations.

Our results suggest that the assumptions of dimensional comparison theory (Möller & Marsh, 2013) and the Marsh/Shavelson model of ASC (Marsh, 1990) do not currently provide a satisfactory explanation for the emergence of assimilation versus contrast effects between verbal domains. Thus, the mechanisms underlying the emergence of these effects require further investigation and refinement. For example, Helm et al. (2016) ran three experimental studies analyzing the effects of experimentally induced higher or lower perceived subject similarity on ASC differences. They found that students' perceptions of similarity between domains determined the strength of contrast effects. Future GI/E model studies including multiple verbal domains could thus benefit from including perceived subject similarity as a moderator in the model. Apart from perceived subject similarity, other factors could also impact students' language-related social and dimensional comparisons. For example, a sense of identity in relation to the languages being learnt is posited to facilitate language acquisition (Lambert, 1992; Mercer, 2011). A strong (or weak) identity in relation to one (or more) languages, could moderate dimensional comparison effects and might explain the strong contrast effects found in our sample. This interaction could be more complex if students' native language(s) and school languages do not overlap and might also depend on whether the student views the additional language as replacing their native language (see Marsh, Hau, & Kong, 2002). Future GI/E studies using multiple verbal domains could include identity variables as moderators and include the verbal domains (that are not included in the school curriculum) of language minority students. Furthermore, future studies could also include other factors as moderators, such as students' perceived importance of the languages learnt at school, linguistic distance between languages (Arens et al., 2019), age of acquisition, the number of languages known, and contextual factors associated with language acquisition and use (see Hoff, 2006). Complementary to such quantitative studies on the GI/E model, qualitative research investigating the role of the context (and the other aforementioned factors) in the formation of language motivation and affect could provide valuable insights into the mechanisms underlying dimensional comparisons (e.g., Mercer, 2011; Möller & Husemann, 2006).

Due to the cross-sectional nature of our study, we could not assess if, how, and to what extent the change in language of instruction from German to French affected the formation of students' math, German, and French ASCs. Future studies should employ longitudinal designs to investigate the impact of a change in language of instruction on students' ASC over time. This may be particularly relevant in societies that are increasingly multilingual and geographically mobile.

The findings of our study are in line with motivational-affective theory and previous research. However, in contrast to many previous (G)I/E model studies using (nonrepresentative) convenience samples (e.g., Arens, Helm, Wolff, & Möller, 2019; Bong, 2001; Gaspard et al., 2018; Möller, Streblow, & Pohlmann, 2006), we used a fully nationally representative dataset encompassing virtually all ninth-grade students in Luxembourg, thereby covering a diverse range of sociodemographic variables including socioeconomic status, immigration background, school track, and home language background. While it should be taken into account that certain characteristics of Luxembourg's multilingual educational system may not easily generalize to other countries or student samples, such multilingual educational systems might arguably become more commonplace in the future due to rising multilingualism and globalization. As such, a deeper understanding of the potential impact of students' linguistic background on the formation of language learning motivation and affect is required. Empirically testing and establishing the invariance of the GI/E model relations across students from different home language backgrounds represents an important first step in achieving this. Using a fully representative student sample resulted in home language groups that were unequal in size. However, the potential impact of unbalanced group sizes on invariance results with more than two groups remains an open question to be dealt with in future studies (Yoon & Lai, 2018). Replicating our results with completely balanced language groups will thus provide further robustness to our findings.

6. Implications and conclusions

The present study builds on previous research on the GI/E model with the aim of extending and informing educational theory and practice

in the domain of language learning, with the ultimate goal of improving students' language–related motivation and affect in a globalized society in which multilingualism is increasingly valued. In terms of theoretical implications, the results of our study could help refine dimensional comparison theory (Möller & Marsh, 2013) to assume contrast rather that assimilation effects among languages. Moreover, the results of our invariance tests implied similar achievement–ASC, achievement–interest, and achievement–anxiety relations across students from various monolingual and multilingual home language backgrounds. The GI/E model relations thus appear to function similarly, regardless of students' linguistic backgrounds.

Concerning the practical implications, previous research has shown that teachers and parents may overlook the domain specificity of languages and the role of dimensional comparisons in ASC formation (Dai, 2002; Helm, Müller-Kalthoff, Mukowski, & Möller, 2018). The GI/E model could help teachers and parents understand how an individual student who is achieving well in two languages can have a relatively high ASC in one language and a relatively low ASC in the other language. Similarly, the GI/E model can explain how interventions aimed at enhancing students' ASC in one language domain may not transfer to their ASCs in other language domains. Thus, language ASCs should be regarded as domain-specific and targeted through interventions that consider the complex interplay between achievement, ASC, interest, and anxiety within and across various academic domains. Furthermore, the finding that the functional GI/E model relations generalize across students from different home language backgrounds could facilitate the design of interventions aiming to improve students' language-related ASCs, interests, and anxieties. Educationalists working with students from different home language groups can assume that similar functional mechanisms (i.e., dimensional comparisons) apply to the formation of these students' ASCs, interests, and anxieties. For example, if a student from a monolingual (or multilingual) household is performing well in one language, parents and teachers could improve the student's ASC in the other language by mitigating the potential unwanted effects of dimensional comparisons. This could be achieved by downplaying the differences between the languages (thereby influencing the student's perceived similarity of the two languages), which, in turn, might minimize contrast effects (Helm et al., 2016). In conclusion, our study highlights the generalizability of the GI/E model to multiple domains, multiple outcomes, and across different student groups, and offers directions for future research in this area.

7. Author note

We have no conflicts of interest to disclose. This research was supported by grants from the University of Luxembourg (project "enhanCe") and from the Luxembourg National Research Fund (C16/SC/11333571). We would like to thank the national school monitoring team from the Luxembourg Centre for Educational Testing for providing access to the Épreuves Standardisées database.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cedpsych.2021.102037.

References

Adle, Astrid, & Beyer, Rahel (2017). Languages and language policies in Germany. Contributions to the EFNIL Conference 2017 in Mannheim. In *National language institutions and national languages* (pp. 221–242). Budapest: Research Institute for Linguistics, Hungarian Academy of Sciences.

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

Arens, A. K., Helm, F., Wolff, F., & Möller, J. (2019). Social and dimensional comparisons in the formation of German students' language self-concepts. *The Journal of Experimental Education, 88*(2), 245–264. https://doi.org/10.1080/00220973.2019.1635069

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

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

Aronin, L. (2019). What is Multilingualism? In D. Singleton, & L. Aronin (Eds.), *Twelve Lectures on Multilingualism* (pp. 3–34). Multilingual Matters.

Australian Government. (2021). Australian education system. https://www.studyaustralia.gov.au/english/australian-education/education-system.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal, 13*(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2

Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley. https://doi.org/10.1002/9781118619179

Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task value, and achievement goals. *Journal of Educational Psychology, 93*(1), 23–34. https://doi.org/10.1037/0022-0663.9

Brantmeier, C. (2006). Toward a multicomponent model of interest and L2 reading: Sources of interest, perceived situational interest, and comprehension. *Reading in a Foreign Language, 18*, 89–115.

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

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology, 102*(4), 964–981. https://doi.org/10.1037/a0019644

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Coyle, D. (2007). Content and language integrated learning: Towards a connected research agenda for CLIL pedagogies. *International Journal of Bilingual Education and Bilingualism, 10*(5), 543–562. https://doi.org/10.2167/beb459.0

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*(6), 617–645. https://doi.org/10.1177/074355802237467

Dörendahl, J., Scherer, R., Greiff, S., Martin, R., & Niepel, C. (2021). Dimensional comparisons in the formation of domain-specific achievement goals. *Motivation Science*. https://doi.org/10.1037/mot0000203. Advance online publication.

Dörnyei, Z. (2002). The motivational basis of language learning tasks. In P. Robinson (Ed.), *Individual differences in second language acquisition* (pp. 137–158). John Benjamins. https://doi.org/10.1075/lllt.2.10dor.

Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Language Learning, 53*(S1), 3–32. https://doi.org/10.1111/1467-9922.53222

Eidswick, J. (2009). The influence of interest on reading comprehension in EFL students. *Annual Research Report of the Language Center, 12*, 25–38, Kwansei Gakuin University.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Fischbach, A., Ugen, S., & Martin, R. (2014). *ÉpStan Technical Report*. University of Luxembourg.

Ganzeboom, H. B. G. (2010). *International Standard Classification of Occupations* [ISCO-08 with ISEI-08 Scores]. http://www.harryganzeboom.nl/isco08/isco08_with_isei.pdf.

Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research, 25*(3), 201–239. https://doi.org/10.1006/ssre.1996.0010

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

Gardner, R. C., & MacIntyre, P. D. (1992). A student's contributions to second language learning. Part I: Cognitive variables. *Language Teaching, 25*(4), 211–220. https://doi.org/10.1017/S026144480000700X

Gardner, R. C., & MacIntyre, P. D. (1993). A student's contributions to second-language learning. Part II: Affective variables. *Language Teaching, 26*(1), 1–11. https://doi.org/10.1017/S0261444800000045

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

Gilles, P., & Trouvain, J. (2013). Luxembourgish. *Journal of the International Phonetic Association, 43*(1), 67–74. https://doi.org/10.1017/S0025100312000278
Goetz, T., Frenzel, A. C., Hall, N. C., & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology, 33*(1), 9–33. https://doi.org/10.1016/j.cedpsych.2006.12.002
Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., … Preckel, F. (2014). "My Questionnaire is Too Long!": The assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology, 39*(3), 188–205. https://psycnet.apa.org/doi/10.1016/j.cedpsych.2014.04.002.
Gogol, K., Brunner, M., Martin, R., Preckel, F., & Goetz, T. (2017). Affect and motivation within and between school subjects: Development and validation of an integrative structural model of academic self-concept, interest, and anxiety. *Contemporary Educational Psychology, 49*, 46–65. https://doi.org/10.1016/j.cedpsych.2016.11.003
Gogol, K., Brunner, M., Preckel, F., Goetz, T., & Martin, R. (2016). Developmental dynamics of general and school-subject-specific components of academic self-concept, academic interest, and academic anxiety. *Frontiers in Psychology, 7*, 356. https://doi.org/10.3389/fpsyg.2016.00356
Grand Duchy of Luxembourg. (2019a). *10 Fact sheets about the Grand Duchy of Luxembourg.* https://luxembourg.public.lu/content/dam/luxembourg/publications/10-fiches-sur-le-grand-duche-de-luxembourg/10-fiches-sur-le-grand-duche-de-luxembourg-en.pdf.
Grand Duchy of Luxembourg. (2019b). *Demographics*. https://luxembourg.public.lu/en/society-and-culture/population/demographics.html.
Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., & Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in science: Dimensional comparison processes and expectancy-by-value interactions. *Learning and Instruction, 49*, 81–91. https://doi.org/10.1016/j.learninstruc.2016.12.007
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*(2), 1–9. https://doi.org/10.1177/2332858416650624
Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*(1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002
Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118
Jansen, Malte, Schroeders, Ulrich, Lüdtke, Oliver, & Marsh, Herbert, W. (2015). Contrast and assimilation effects of dimensional comparisons in five subjects: An extension of the I/E model. *Journal of Educational Psychology, 107*(4), 1086–1101. https://doi.org/10.1037/edu0000021
Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International; Lawrence Erlbaum Associates, Inc.
Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? *Annual Review of Applied Linguistics, 15*, 125–150. https://doi.org/10.1017/S0267190500002658
King, L. (2017). *The impact of multilingualism on global education and language learning.* Cambridge Assessment English Perspectives. Cambridge English Language Assessment.
Köller, O., Daniels, Z., & Baumert, J. (2000). *Multiple frames of reference, academic interests, and coursework selection in upper secondary schools in Germany*. Annual Meeting of the American Educational Research Association. https://pdfs.semanticscholar.org/f82a/22e59dc12c90cc687fe68395cb644c0cfdb1.pdf?_ga=2.136569968.1563034321.1587671253-1699252056.1571158001.
Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction, 12*(4), 383–409. https://doi.org/10.1016/S0959-4752(01)00011-1
Lambert, W. E. (1992). Challenging established views on social issues: The power and limitations of research. *American Psychologist, 47*(4), 533–542.
Lasagabaster, D., Doiz, A., & Sierra, J. M. (Eds.). (2014). *Motivation and foreign language learning: From theory to practice.* John Benjamins.
Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports, 20*(3), 975–978.
Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*(1), 129–149. https://doi.org/10.2307/1163048
Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology, 82*(4), 623–636. https://doi.org/10.1037/0022-0663.82.4.623
Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J. S., Abdelfattah, F., Leung, K. C., Xu, M. K., Nagengast, B., & Parker, P. D. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology, 105*, 108–128. https://doi.org/10.1037/a0029907
Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F., Nagengast, B., Möller, J., & Abu-Hilal, M. M. (2015). The internal/external frame of reference model of self-concept and achievement relations. *American Educational Research Journal, 52*(1), 168–202. https://doi.org/10.3102/0002831214549453
Marsh, H. W., & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology, 95*(4), 687–706. https://doi.org/10.1037/0022-0663.95.4.687
Marsh, H. W., Hau, K.-T., & Kong, C.-K. (2002). Multilevel causal ordering of academic self-concept and achievement: Influence of language of instruction (English compared with Chinese) for Hong Kong students. *American Educational Research Journal, 39*(3), 727–763. https://doi.org/10.3102/00028312039003727
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*(3), 543–553. https://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*(4), 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'Neill, R. (1984). Self Description Questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement, 21*(2), 153–174. https://doi.org/10.1111/jedm.1984.21.issue-210.1111/j.1745-3984.1984.tb00227.x
Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology, 54*(2), 263–280. https://doi.org/10.1037/dev0000393
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*(3), 389–420. https://doi.org/10.1207/S15327906389-420
Mercer, S. (2011). Towards an understanding of language learner self-concept. In *, 12. Educational Linguistics*. Springer. https://doi.org/10.1007/978-90-481-9569-5.
Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525–543. https://doi.org/10.1007/BF02294825
Miller, J. W. (2000). Exploring the source of self-regulated learning: The influence of internal and external comparisons. *Journal of Instructional Psychology, 27*, 47–52.
Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. Routledge. https://doi.org/10.4324/9780203821961
Mitchell, M. A., & Maxwell, S. E. (2013). A comparison of the cross-sectional and sequential designs when assessing longitudinal mediation. *Multivariate Behavioral Research, 48*(3), 301–339. https://doi.org/10.1080/00273171.2013.784696
Möller, J., & Husemann, N. (2006). Internal comparisons in everyday life. *Journal of Educational Psychology, 98*(2), 342–353. https://doi.apa.org/doi/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*(4), 826–835. https://doi.org/10.1037/0022-0663.93.4.826
Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review, 120*(3), 544–560. https://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*(2), 1–11. https://doi.org/10.14786/flr.v4i2.169
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*(3), 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. *American Educational Research Journal, 48*(6), 1315–1346. https://doi.org/10.3102/0002831211419649
Möller, J., Streblow, L., & Pohlmann, B. (2006). The belief in a negative interdependence of math and verbal abilities as determinant of academic self-concepts. *The British Journal of Educational Psychology, 76*, 57–70. https://doi.org/10.1348/000709905X37451
Möller, J., Streblow, L., Pohlmann, B., & Köller, O. (2006). An extension to the internal/external frame of reference model to two verbal and numerical domains. *European Journal of Psychology of Education, 21*(4), 467–487. https://doi.org/10.1007/BF03173515
Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research, 90*(3), 376–419. https://doi.org/10.3102/0034654320919354
Muthén, L. K., & Muthén, B. O. (1998–2019). *Mplus Version 8.3 [statistical software]*. Muthén & Muthén.
Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., & Eccles, J. S. (2008). Gendered high school course selection as a precursor of gendered careers: The mediating role of self-concept and intrinsic value. In H. M. G. Watt, & J. S. Eccles (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 115–143). American Psychological Association. https://doi.org/10.1037/11706-004.
Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation, 12*(4), 323–345. https://doi.org/10.1080/13803610600765687
Niepel, C., Brunner, M., & Preckel, F. (2014a). Achievement goals, academic self-concept, and school grades in mathematics: Longitudinal reciprocal relations in above

average ability secondary school students. *Contemporary Educational Psychology, 39* (4), 301–313. https://doi.org/10.1016/j.cedpsych.2014.07.002

Niepel, C., Brunner, M., & Preckel, F. (2014b). 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*(4), 1170–1191. https://doi.org/10.1037/a0036307

Papi, M. (2010). The L2 motivational self system, L2 anxiety, and motivated behavior: A structural equation modeling approach. *System, 38*(3), 467–479. https://doi.org/10.1016/j.system.2010.06.011

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational Psychology, 36*(1), 36–48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Pekrun, R., & Stephens, E. J. (2010). Achievement emotions: A control-value approach. *Social and Personality Psychology Compass, 4*, 238–255. https://doi.org/10.1111/j.1751-9004.2010.00259.x

Prevoo, M. J. L., Malda, M., Mesman, J., & van IJzendoorn, M. H. (2016). Within- and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background. *Review of Educational Research, 86*, 237–276. https://doi.org/10.3102/0034654315584685

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/.

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

Renninger, K. A. (2000). Individual interest and its implications for understanding of intrinsic motivation. In Carol Sansone, & Judith M. Harackiewicz (Eds.), *Intrinsic and Extrinsic Motivation: The search for optimal motivation and performance* (pp. 373–404). Academic Press. https://doi.org/10.1016/B978-012619070-0/50035-0.

Robitzsch, A., Kiefer, T., & Wu, M. (2019). *Test analysis modules (V3.3-10)*. https://cran.r-project.org/web/packages/TAM/TAM.pdf.

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychology, 26*(3), 299–323. https://doi.org/10.1207/s15326985ep2603&4_5

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

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*(3), 407–441. https://doi.org/10.3102/00346543046003407

Spielberger, C. D. (1980). *Test anxiety inventory: Preliminary professional manual*. Consulting Psychologists Press.

University of Luxembourg, & SCRIPT (Eds.). (2018). *Nationaler Bildungsbericht Luxemburg 2018*. University of Luxembourg. http://www.men.public.lu/catalogue-publications/themes-transversaux/statistiques-analyses/bildungsbericht/2018/de.pdf.

Warm, T. A. (1985). *Weighted likelihood estimation of ability in item response theory with tests of finite length* [Doctoral dissertation, The University of Oklahoma], ShareOK Archive. https://shareok.org/bitstream/handle/11244/5356/8524078.PDF?sequence=1&isAllowed=y.

Wawire, B. A., & Kim, Y.-S. (2018). Cross-language transfer of phonological awareness and letter knowledge: Causal evidence and nature of transfer. *Scientific Studies of Reading, 22*(6), 443–461. https://doi.org/10.1080/10888438.2018.1474882

Wigfield, A., Eccles, J. S., & Möller, J. (2020). How dimensional comparisons help to understand linkages between expectancies, values, performance, and choice. *Educational Psychology Review, 32*(3), 657–680. https://doi.org/10.1007/s10648-020-09524-2

Wigfield, Allan, & Eccles, Jacquelynne S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

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*(2), 489–503. https://doi.org/10.1037/a0031333

Yeung, A. S., & Lau, I. C. (1998). *The verbal academic self-concept structure of university students*. International Conference of the University of Malaysia Faculty of Language. https://files.eric.ed.gov/fulltext/ED431332.pdf.

Yoon, M., & Lai, M. H. C. (2018). Testing factorial invariance with unbalanced samples. *Structural Equation Modeling: A Multidisciplinary Journal, 25*, 201–213. https://doi.org/10.1080/10705511.2017.1387859

Zeidner, M. (2007). Test anxiety in educational contexts: Concepts, findings, and future directions. In P. Schutz, & R. Pekrun (Eds.), *Emotion in Education* (pp. 165–184). Academic Press. https://doi.org/10.1016/B978-012372545-5/50011-3.

Zeidner, M., Nevo, B., & Lipschitz, H. (1988). *The Hebrew version of the test anxiety inventory*. University of Haifa.