Jansen, Malte; Becker, Michael; Neumann, Marko

Dimensional comparison effects on (gendered) educational choices

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
The Journal of educational psychology 113 (2020) 2, S. 330-350

Bitte verwenden Sie in der Quellenangabe folgende URN oder DOI /
Please use the following URN or DOI for reference:
urn:nbn:de:0111-pedocs-220459
10.25656/01:22045

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-220459
https://doi.org/10.25656/01:22045

Dimensional Comparison Effects on (Gendered) Educational Choices

Malte Jansen, Michael Becker & Marko Neumann

Authors' version (last submitted version)

Abstract

Expectancy-value theory (EVT) proposes that students' appraisals of success expectancy and task value are the main drivers of their study and career choices. Dimensional comparison theory (DCT) proposes that these beliefs are themselves affected by students comparing their ability across different domains. However, only a few studies have aimed to integrate these approaches and clarify the role of dimensional comparisons within EVT. Using longitudinal data, we aimed to fill this gap by studying within- and cross-domain effects of achievement (grades and test scores), academic self-concept (as a surrogate for expectancy beliefs), and values on German adolescents' (N = 519) high school course choices and their intentions to major in a STEM subject at university. We show that (a) self-concepts predicted course choices, whereas values predicted STEM study intentions, (b) dimensional comparison patterns (positive within-domain and negative across-domain relations) were present, (c) gender differences in course choices were mediated by differences in achievement, self-concept and value, and (d) there was an incremental gender effect on STEM study intentions above and beyond achievement, self-concept, value, and previous course choices. Furthermore, overall, a model incorporating cross-domains paths representing dimensional comparisons fit the data better than a model without these paths. We conclude that direct and indirect dimensional comparison effects contribute to predicting choices of high school courses and university majors and to understanding gender differences in these choices. We recommend that studies in the EVT framework include cross-domains effects.

Keywords: academic self-concept, I/E model, dimensional comparison theory, course choice, STEM, longitudinal study

Educational Impact and Implications Statement

We show that students' domain-specific self-concepts (i.e., their self-perceived ability in a domain) and values (e.g., their enjoyment and interest in a domain) are related to their choices and intentions (expectancy-value-theory; EVT). When making course choices, students' self-perceived ability seems to play a major role, whereas intentions regarding the domain of study after graduation are more strongly related to the values students ascribe to different domains. Furthermore, our study emphasizes that, when choosing advanced courses or forming study intentions, students consider their achievement, self-concept and value in several domains thus weighting their strengths and weaknesses against each other (dimensional comparison theory; DCT). Our results may be important for educators who counsel students about their course and post-school choices. The assumed processes (i.e., differences in self-concept and value) partly, but not fully explain gender differences in STEM study intentions, indicating a need for further research.

Dimensional Comparison Effects on (Gendered) Educational Choices

How do high school students approaching graduation decide what to study at university and what kind of career to choose? How are these decisions affected by the students' achievement, their motivation, and their past course choices? And why are there substantial gender differences in educational and subsequent career choices related to the STEM domains (Science, Technology, Engineering, and Mathematics) even though no substantial gender differences have been found in mathematics and science achievement (Cheryan, Ziegler, Montoya, & Jiang, 2017; Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Wang, Eccles, & Kenny, 2013; Zell, Krizan, & Teeter, 2015)?

Expectancy-value theory (EVT; Eccles, 2011; Wigfield & Eccles, 2000) describes the current understanding of such decision processes. It proposes that differences in domain-specific academic self-concepts and values are the main drivers of differential educational choices. Indeed, there is strong evidence for stereotypical differences in these motivational constructs (for an overview, see Schoon & Eccles, 2014): female students tend to prefer and feel more competent in language domains, and male students tend to prefer and feel more competent in mathematics (Eccles, Wigfield, Harold, & Blumenfeld, 1993; Gaspard et al., 2015; Hyde, Fennema, Ryan, Frost, & Hopp, 1990) and some science domains such as physics and chemistry (but not biology; Jansen, Schroeders, & Lüdtke, 2014; Jansen, Schroeders, Lüdtke, & Marsh, 2019).

There is a great deal of literature on two aspects of self-concepts and values: (a) how they affect aspirations and choices (based on the EVT framework) and (b) how different antecedents such as comparison processes affect them. In the second line of research, dimensional comparison processes have been studied extensively—they occur when students internally compare their achievement across different study domains (dimensional comparison theory (DCT), see Möller & Marsh, 2013).

However, until recently, few studies have connected EVT and DCT and studied the mediational chain that extends from individual characteristics (e.g., gender) to the antecedents of self-concepts and values to choices (Lauermann, Chow, & Eccles, 2015; Nagy, Trautwein, Baumert, Köller, & Garrett, 2006; Parker et al., 2012). Furthermore, not all these studies were longitudinal and few included more than one choice indicator (e.g., course or study choice) or more than one achievement indicator (i.e. either teacher-assigned grades, that are students' most salient achievement feedback at school and tap into their performances and behaviors in class, or scores from standardized tests, that tap into their domain-specific abilities). Studying how different achievement measures within and across several domains affect expectancy and value, which may then affect subsequent choices, would shed more light on the role of mediated effects and of cross-domain comparisons within EVT. Furthermore, it would provide evidence for the practical implications of DCT (if dimensional comparisons were to affect choices). Finally, it could also help to understand the well-established gender differences in study intentions better: Are women, for example, less likely to choose a STEM-related subject because of their motivational characteristics (expectancy and value), because of their ability or school performance in different domains, or because they previously already chose a different specialization through their course selection in high school? Are there additional direct effects that are not explained by these factors? To address these questions, studying both direct and indirect effects along a mediational chain is helpful. With this study, we aimed to implement such a design. Using longitudinal data, we examined within- and cross-domain effects of German students' achievement (both grades and test-scores), self-concept and value on their advanced mathematics vs. language course choices in upper secondary school, and their intentions to choose a STEM major at university. In addition, we tested whether differences in achievement, expectancy, and value would mediate and partly explain gender differences in course choices and study intentions.

Theoretical Background

Expectancy-Value Theory as a Framework for Predicting Educational Choices

The main goal of EVT is to serve as an overall framework for explaining students' achievement motivation and their educational aspirations and choices, both in general and with regard to gender differences in these choices in particular (Eccles, 2009; Eccles et al., 1983; Wigfield & Eccles, 2000). According to EVT, achievement motivation and subsequently aspirations and choices are mainly influenced by *expectancy beliefs* (i.e., what students think they can do) and *values* (i.e., what students like to do and what they see value in). When studying motivation, aspirations, and choices with regard to academic domains (e.g., mathematics, languages), the expectancy component is typically operationalized as students' domain-specific academic self-concept because the two constructs have been shown to be empirically indistinguishable on domain level (Gaspard et al., 2018; Guo, Marsh, Parker, Morin, & Dicke, 2017; Nagengast et al., 2011; Trautwein et al., 2012) . Therefore, we will also use the terms interchangeably in the following when reporting previous results. By contrast, the value component is assumed to be differentiated into four facets. High *intrinsic value* would be indicated by experiencing flow and enjoyment in an academic subject. Furthermore, students may feel that doing well in a domain is personally important for them (*attainment value*). Students may also consider doing well in a domain to have high *utility value* for them (e.g., because this might increase their chances to study their preferred subject at university or do well on the job market). Finally, when making decisions about engaging in certain tasks, students may also consider the *cost* (i.e., negative consequences from pursuing this path over other alternatives). Recent studies have suggested that these four categories of values can be further differentiated (for details, see Gaspard et al., 2015, 2018).

Previous studies have shown how expectancies and the different value components can contribute to educational achievement (Guo, Marsh, Parker, Morin, & Yeung, 2015; Guo, Parker, Marsh, & Morin, 2015; Trautwein et al., 2012), career or coursework aspirations

(Nagengast et al., 2011; Watt et al., 2012), and actual educational choices such as coursework and major selection (Guo, Parker, et al., 2015; Lauermann, Tsai, & Eccles, 2017). Furthermore, expectancy and value are not only assumed to show independent main effects, but also to interact such that the combination of high expectancy beliefs and high values leads to an incremental positive effect. Some recent studies tested and found this positive interaction on coursework aspirations in science (Guo et al., 2017) as well as course choices (Guo, Parker, et al., 2015) and career attainment in mathematics (Lauermann et al., 2017).

Gender Differences in Motivation and Choice

The STEM domain has been a particularly relevant field of application for EVT because of the "leaking pipeline" to STEM occupations and the lower likelihood of women to choose such occupations (Ball, Huang, Cotten, & Rikard, 2017; Eccles, 2007; Marsh et al., 2019; Wang & Degol, 2013). In the school setting, male students typically show higher self-concepts in mathematics and the "hard" sciences such as physics, whereas female students tend to show higher language self-concepts (Arens & Jansen, 2016; Eccles et al., 1993; Hyde, Fennema, Ryan, et al., 1990; Jansen et al., 2014; Watt, 2004). The differences in mathematics self-concept are not substantiated by gender differences in mathematics ability or achievement, which are small at best (Else-Quest, Hyde, & Linn, 2010; Hyde, 2014; Hyde, Fennema, & Lamon, 1990; Reilly, Neumann, & Andrews, 2014). With regard to value facets, the findings are a little more nuanced with similar stereotypical differences found mostly for intrinsic value (and academic interest), and smaller or no differences for attainment and utility value (Gaspard et al., 2015; Gaspard, Häfner, Parrisius, Trautwein, & Nagengast, 2017; Watt, 2004).

Gender differences in mathematics and science related self-concept and values are assumed to result from a variety of influences including differences in role orientations, identity, and self-stereotype matching to different domains (Heyder & Kessels, 2013; Heyder, Kessels, & Steinmayr, 2017; Kessels, 2005), differences in participation in science and mathematics related school activities (Simpkins, Davis-Kean, & Eccles, 2006), as well as gender-differential beliefs and expectations of parents (Muntoni & Retelsdorf, 2019; Simpkins, Fredricks, & Eccles, 2012) and teachers (Gentrup & Rjosk, 2018; Muntoni & Retelsdorf, 2018).

Overall, there is clear evidence for stereotypical gender differences in the expectancy and value components, which should act as mediators between students' gender and their educational and career choices. This mediational chain has been documented empirically such that indirect effects of gender on educational aspirations and course choices have been found to be mediated by mathematics self-concept (Guo, Marsh, et al., 2015). Longitudinal studies have also found, however, that gender effects on career plans in mathematics are usually direct and not (fully) mediated by mathematics-related self-concept and value (Lauermann et al., 2015, 2017). Similarly, Parker et al. (2012) and Marsh et al. (2019) found that gender differences in choice of university major persisted after they controlled for achievement as well as mathematical and verbal self-concept.

The Role of Dimensional Comparisons in Shaping Educational Choices

Comparison processes are assumed to be amongst the most important sources of academic self-concepts. Students compare their achievement in a given domain with at least three standards (Marsh et al., 2018; Möller & Marsh, 2013; Wolff, Helm, Zimmermann, Nagy, & Möller, 2018): the achievement of their peers in the same domain (social comparisons), their own previous achievement in the same domain (temporal comparisons), and their own achievement in other domains (dimensional comparisons). According to DCT, students with similar levels of achievement in one domain (e.g., math) often show different self-concepts in that domain because they have different achievement levels in other domains (e.g., English), and they tend to engage in dimensional comparison processes. When it comes to explaining educational aspirations and choices, dimensional comparisons may be particularly relevant, for example, when a decision to take advanced coursework in one subject is consequently a decision not to pursue advanced coursework in other subjects or when a decision for or against a study domain is also affected by the self-perceived expectancies and values as well as the vocational interests related to other domains (Eccles, 2009; Perera & McIlveen, 2018).

Empirically, the classic paradigm for studying dimensional comparison effects is the internal/external frame of reference model (I/E model; Marsh, 1986; Möller, Pohlmann, Köller, & Marsh, 2009). It is based on the analysis of achievement and self-concept measures in two domains (typically mathematics and a language domain) using path modeling. After controlling for the positive effect of achievement on self-concept in the same domain (e.g., mathematics achievement on mathematics self-concept), dimensional comparison effects can be observed as a negative effect of achievement on self-concept across domains (e.g., German achievement on mathematics self-concept). This pattern of relations for mathematics and the native language domain has been found consistently in cross-cultural observational studies (Marsh & Hau, 2004), longitudinal studies (e. g., Möller, Retelsdorf, Köller, & Marsh, 2011), experimental studies (Möller & Köller, 2001), and a meta-analytical review (Möller et al., 2009).

Whereas this line of research was initially primarily focused on academic self-concept, recent research applied it to other constructs including a range of cognitive, affective, and motivational constructs that, like academic self-concept, may mediate the relation between achievement and motivated behavior, such as intentions and choices (generalized I/E model, see Arens, Becker, & Möller, 2017; Arens & Möller, 2016; Van Zanden et al., 2017). This also includes several studies that have found dimensional comparison effects on values and interests, showing effect patterns that are very similar to those found for academic self-concept (Goetz, Frenzel, Hall, & Pekrun, 2008; Guo et al., 2017; Schurtz, Pfost, Nagengast, & Artelt, 2014). A recent study including expectancy and several facets of value measured in five academic domains found that dimensional comparison effects varied across different value facets with the strongest effects for intrinsic value (Gaspard et al., 2018).

As mentioned above, one key argument for the importance of dimensional comparison effects is their potential role in educational and career choices. Dimensional comparisons (based on students' achievement in several domains) may affect students' choices in two ways: indirectly through the effects on self-concept and values, or directly. However, the direct path seems implausible to a certain extent because there should always be motives, cognitive processes, and dispositions that lead to a choice (with respect to the assumptions of EVT, resulting in expectancies and values that then explain the choices). Furthermore, it is assumed that students engage in dimensional comparisons explicitly to obtain information about themselves and to develop a profile of strengths and weaknesses (Möller & Marsh, 2013). This mechanism suggests an effect on academic self-concepts first before choices are affected.

Dimensional comparisons and coursework selection. Nagy and colleagues (2006) examined the coursework selection of students. Using a longitudinal data set with one measurement point before courses were chosen (Grade 10) and one after (Grade 12), they found evidence for dimensional comparison effects. More specifically, mathematics self-concept and intrinsic value showed positive effects on the choice of advanced mathematics courses and negative effects on the choice of advanced biology courses. Gender differences in course choices were completely mediated by achievement, self-concept, and value (along various within- and cross-domain paths). This research was further expanded in the two domains of mathematics and English showing dimensional comparison effects along a meditational path from achievement to self-concept and value to course choices (Nagy et al., 2008). Another study focused on younger students who had to choose at the end of Grade 7 between biology and chemistry as their science course in the next school year (Dickhäuser, Reuter, & Hilling, 2005). The authors found contrasting dimensional comparison effects from self-concept in one domain on course choice in the other domain. This is in line with a more recent study focusing on coursework aspirations (Guo, Marsh, Parker, Morin, & Dicke, 2017).

Dimensional comparison and study choices. Contrary to the abovementioned studies, which focused on coursework in secondary school, a few studies examined post-school choices. Lauermann, Chow, and Eccles (2015) showed cross-domain (i.e., dimensional comparison) effects of values on career plans in a sample of twelfth-graders. For example, valuing mathematics was positively related to planning a mathematics-related career and negatively related to planning a human services career with the opposite pattern for English values. A gender effect on career plans remained above and beyond the effects of self-concept and value. However, the study did not include domain-specific achievement measures. Another longitudinal study using German and English samples aimed to predict university entry and university major selection (Parker et al., 2012). The authors found that mathematics self-concept and achievement negatively affected major selection in the humanities, law, and biomedical sciences relative to selecting a math-intensive major such as physics or engineering. Self-concept and achievement in English showed the opposite effects (a negative effect on choosing math-intensive majors, a positive effect on choosing other majors), indicating a pattern of dimensional comparison effects. Gender differences in major selection were partly mediated through self-concepts. However, domain-specific value facets were not considered. Guo, Parker, Marsh, and Morin (2015) showed dimensional comparison effects of reading achievement on mathematics self-concept and value which later influenced the choice of college mathematics courses in an Australian sample. However, no self-concept or value facets in other domains were considered.

The Present Study

Until recently there was a relative lack of studies integrating perspectives from EVT and DCT even though “internal comparison processes” have also been suggested to be a central source of expectancy beliefs (Eccles, 2009, p. 82). There are now a few studies moving towards that integration, but there are still gaps. Only some studies integrating EVT and DCT were longitudinal, and not all of these measured achievement, self-concept, and value before the choice outcomes. Furthermore, no previous studies have examined both high school course choices and the study intentions that might follow from these choices (i.e., choice of major at university). Also, most studies focusing on the integration of EVT and DCT used either teacher-assigned grades or achievement, but not both.

In this study, we focus on the longitudinal prediction of course choices (advanced mathematics vs. German course) and the intention to study a STEM subject at university in a sample of academic-track students from Berlin, Germany. Our two measurement points take place before (Grade 9) and after (Grade 12) the transition to upper secondary school, which happens after Grade 10. At this transition, students decide which two advanced courses they will specialize in. Students will receive more lessons in these courses (five weekly lessons compared with three in basic courses), and their achievement in these courses will carry double the weight in contributing to overall GPA. Because of this specialization, it has been argued that the choice of these courses “often determines the students’ field of study at college” (Nagy et al., 2008, p. 116) and may serve as a filter for the range of perceived future choices (e.g., advanced mathematics courses for STEM study choices and careers; Ma & Johnson, 2008). A conceptual representation of the model that was estimated can be found in Figure 1. It includes all within-domain and cross-domain direct and indirect effects along the mediational chain from gender to achievement to expectancy and value to advanced course choice to study intentions.

Based on previous theory and findings, we derived a set of hypotheses for the paths of the model. We first assumed that the predictions of the classic I/E model would be valid (Hypothesis 1). That is, for the effects of achievement on academic self-concept and value, we expected positive within-domain effects (e.g., math grades/test-scores on math self-concept/value; H1.a) and negative cross-domain effects (e.g. math grades/test-scores on German self-concept/value; H1.b). In addition, we expected differential effects of teacher-assigned grades vs. test-scores due to the different characteristics of these achievement indicators. As discussed in previous research on dimensional comparison theory (Jansen, Schroeders, Lüdtke, & Marsh, 2015; Marsh et al., 2014), grades may be a more salient source of feedback for students; hence and therefore more likely to be used for dimensional comparison processes. Furthermore, they have "high ecological validity" (Marsh et al., 2014, p. 329) even though they are more strongly affected by the social frame of reference of the other students in a class ("grading on a curve"), and thus harder to compare across classrooms. Previous studies including both indicators found higher relations between grades and academic self-concepts compared to test-scores (Jansen et al., 2015; Marsh et al., 2014). Therefore, we would also hypothesize that the relations described above (with and cross-domain effects of achievement) would be stronger for grades than test-scores (H1.c).

We further assumed to find the classic pattern predicted by EVT (Hypothesis 2). That is, domain-specific self-concept, value and their interaction should predict course choices in mathematics (H2.a) and German (H2.b). Furthermore, mathematics self-concept, value, and their interaction should also predict STEM study intentions (H2.c). As mentioned above, the interaction effect has recently been re-introduced as a prediction of EVT (Guo et al., 2017; Nagengast et al., 2011).

Going beyond these well-established predictions, we expected choices to be also affected by dimensional comparisons that would manifest in cross-domain effects of expectancy and value (Hypothesis 3). Mathematics self-concept and value should show a negative cross-domain effect on choosing an advanced German course (H3.a) and vice versa (H3.b). In addition, German self-concept and value should be negatively related to STEM study intentions (H3.c).

Fourth, we expected course choices to affect major choice above and beyond the previously measured expectancy and value. Here, we also expected a pattern of dimensional comparisons. That is, students in advanced math courses should have higher STEM study intentions (H4.a), whereas students in advanced German courses should have lower STEM study intentions (H4.b).

Finally, there should be stereotypical gender differences in self-concept, values, course choices, and study intentions (Hypothesis 5). That is, boys should show higher mathematics self-concepts and values and be more likely to select an advanced mathematics course or intent to study a STEM subject (H5.a), whereas girls should show higher German self-concepts and values and be more likely to select advanced German courses (H5.b). Finally, we expected gender differences in choices to be at least partly mediated by achievement, but mostly by self-concept and value (H5.c).

Due to the high number of direct and indirect paths, we did not make literature-based predictions for some more specific relations. For example, we made no specific predictions whether (a) there would be differential effects of expectancy and value on course choices and STEM intentions, (b) there would be additional direct effects of achievement on choices beyond self-concept and value, and (c) cross-domain dimensional comparison effects on choices should manifest mostly from self-concept or from value. These questions will be explored. Our predictions for each path for the estimated models (number and direction) can also be found in Table 3.

Method

Sample and Study Design

Our analyses were based on a subsample of a longitudinal, multicohort study in Berlin, Germany (Maaz et al., 2013, Neumann et al., 2017)[1]. We started with a cohort of students that was drawn as a representative sample of ninth-grade secondary school students in 2011, and

[1] Due to the conditions specified by the Berlin Senate Administration for Education, Youth and Family Affairs which commissioned the study (these conditions were also part of the informed consent statement), the research data are not yet openly available for replication or secondary use. It will be made available in 2023 when an anonymized, well-documented scientific use file (SUF) will be provided. Until then, the data may be used at personal request and in collaboration with the principal investigators.

then followed them longitudinally. The initial sample consisted of 2,913 students in all school tracks (higher/academic track: $N = 672$, intermediate track: $N = 553$, lower/vocational track: $N = 847$, comprehensive schools: $N = 841$). We then focused on students from the academic track (*Gymnasium*) within this sample because, unlike students from the other tracks, most of them went on to the upper secondary level (Grades 11 and 12 in which course choices can be made) and finished with the university entrance certificate (*Abitur*). The sample that we analyzed included all students who were in the academic track in Grade 9 in 2011 and were in the upper secondary level (Grade 12) at the same schools in 2014 ($N = 519$ students from $N = 29$ schools from the initial $N = 672$ *Gymnasium* students).

This sample consisted of 255 girls and 264 boys. The average age was $M = 15.40$ years ($SD = 0.58$), and 155 students had at least one parent that was born outside of Germany. To assess the socioeconomic background, students were asked to specify the current occupation of their parents. The information was recoded using the International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom, de Graaf & Treiman, 1992). The ISEI is a continuous measure aiming to classify occupations according to their income, prestige, and the required educational status ranging from 16 to 90. The ISEI was measured for both parents. The average higher ISEI of both parents (HISEI) was $M = 61.69$ ($SD = 19.15$). This indicated a slightly above average socioeconomic background (the average HISEI for a representative sample of German ninth-grade students in a recent study was 50.7 with an average of 53.9 for Berlin; Mahler & Kölm, 2019), which is plausible given that we focused on students from the academic track.

We used data from two measurement points. In Grade 9 (measurement point T1), students completed achievement tests in different domains as well as a student questionnaire. In Grade 12 (measurement point T2), students completed only a questionnaire. Additional information such as student gender as well as grades and course choices were reported by the schools. Participation was mandatory for all students in public schools in Berlin. The participation rate (i.e., the rate of initially drawn students who were present on the day[s] of assessment) on the student level was 84.8% in Grade 9 (overall rate across all tracks) and 91.5% in Grade 12 (only academic track students).

Measures

Self-concept. We measured students' mathematics and German self-concepts in Grade 9. In both domains, we used five items, including some items with parallel wording (e.g., "I am good at [mathematics/German]") and other items that were used in only one domain (e.g., "I am good at reading" for German or the negatively worded item "I always have problems with mathematical tasks" for mathematics). Students replied using 4-point Likert scales (1 = *fully disagree*, 4 = *fully agree*). The items were a mixture of items from a German self-concept scale for adolescents and young adults developed by Schwanzer and colleagues (2005) that is based on the Self-Description Questionnaire (SDQ III; Marsh & O'Neill, 1984) as well as items from another large-scale German study, the TRAIN study. The reliability coefficients were very good for both domains (mathematics: $\omega = .92$, German: $\omega = .81$). Furthermore, a two-dimensional confirmatory factor analysis (CFA) model (with the four-category items treated as ordinal variables; details about the modeling approach are described below) fit the data well (CFI = .995, TLI = .993, RMSEA = .063).

Value. Like academic self-concepts, the values were also measured at T1 in Grade 9 with similar Likert-style items. There were four items for each domain, all with parallel wording. The scale was originally used to measure individual interest based on conceptualization by Krapp (2002; for similar items, also see Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005). Due to the conceptual similarity, we argue the items can also be used as measures of domain-specific value. They tapped into different value facets with a focus on intrinsic value ("It is important to me to be good in [German/mathematics]": attainment value; "I am prepared to sacrifice leisure time to study for [German/mathematics]": cost, intrinsic value; "Sometimes when I am working on a task in [German/mathematics], I don't notice time flying by":

intrinsic value; "Working on tasks in [German/mathematics] is fun": intrinsic value). Due to the theoretical framing of this study within EVT, we further refer to this scale as "value".

Because only the above mentioned four items were available, we could not differentiate between different value facets. However, similar short-scale approaches have also been used in other studies (Lauermann et al., 2015; Nagengast et al., 2011; Nagengast, Trautwein, Kelava, & Lüdtke, 2013). A two-dimensional CFA model fit (again, with the items treated as ordinal variables) the empirical data reasonably well with regard to some fit indices (CFI = .945, TLI = .932), but showed a sub-optimal RMSEA with .114. However, given the substantive broadness of the value indicators that were included, the good reliability coefficients (mathematics: $\omega = .83$, German: $\omega = .77$) and the acceptable standardized factor loadings (Min = .60; $M = .76$), we decided to stick to the two-factor model with all items in the following analyses. Furthermore, leaving out the single item related to cost did not increase the model fit (CFI = .924, TLI = 0.86, RMSEA = 0.163).

Achievement. The end-of-year grades from the school year 2010/2011 (Grade 9) were reported by the school officials. In addition, standardized achievement tests were also used at T1. Both the mathematics and reading tests consisted of a selection of test booklets from the German PISA 2006 study with a multimatrix test design (Gonzales & Rutkowski, 2010). A total of 48 items were used for mathematics (11-24 per booklet) and 28 items for German reading literacy (14 per booklet). In our analyses, we used five plausible values (PVs) derived from the raw test scores through one parameter IRT models (the Rasch model and the partial credit model; for details on the usage of PVs, see Von Davier, Gonzalez, & Mislevy, 2009; Wu, 2005). The EAP reliabilities were high for both domains ($r_{\text{reading}} = .89$; $r_{\text{math}} = .90$).

Course choices. As mentioned above, students enter the upper secondary level after Grade 10, and then finish school with their university entrance certificate after Grade 12. For this upper secondary level, a course system is introduced. The courses are chosen at the beginning of Grade 11 for the whole upper secondary level. Thus, these course choices took place before the T2 measurement in Grade 12 where students reported their study intentions (see below). In Berlin, students have to choose exactly two advanced courses within the upper secondary level of which one has to be either mathematics, German, a foreign language, or a science subject. The other subject can be freely chosen. Thus, related to our domains of focus, students could chose both mathematics and German, only one, or none. For our study, the course choices were reported by the school at the end of the upper secondary level (we received coded variables for all information included in the university entrance certificates given out by the schools).

STEM study intentions. In Grade 12 (T2), students were asked in an open-ended question format to specify which subject(s) they planned to study at university. They could name up to three subjects (but most students chose to name only one). We coded whether their answers included a STEM subject or not, resulting in a dichotomous classification. The coding was based on a classification of subject groups by the Federal Statistical Office of Germany. All subjects belonging to the groups "mathematics and natural sciences" (mathematics, physics, chemistry, biology, geology and geography) and "engineering" (several domains of engineering as well as computer science) were coded as STEM. Therefore, for example, psychology, medicine, and anthropology were not coded as STEM. We decided to use a dichotomous classification of STEM because the label "STEM" might be the more important trigger of students' attitudes towards and beliefs about the subjects than perceived math-intensity. In addition, it allowed the classification to be based on a clearly defined criterion (see above). As will be shown in the descriptive statistics, the validity of this coding approach can be seen in positive bivariate relations to mathematics achievement and motivation as well as in the pattern of stereotypical gender differences.

Data Analysis

All analyses were conducted within the framework of structural equation modeling (SEM) using the software Mplus 8. Figure 1 shows the full model that was estimated to test the hypotheses outlined above. It included all direct and mediated paths from gender to mathematics and German achievement (test-scores and grades) to self-concept and value (all measured in Grade 9) to course choices after Grade 10 to study intentions in Grade 12. For mathematics and German self-concept and value, latent factors were estimated from the items. Due to the structure of the items with four response categories, we treated them as ordered categorical variables. Advanced mathematics and German course choice and STEM study intentions were included as dichotomous variables (0 = course not chosen, 1 = course chosen; respectively: 0 = no STEM subject intended, 1 = at least one STEM subject intended). For the achievement measure, we used 5 plausible values as mentioned above. Analogous to the procedure for multiple imputed datasets, all models were estimated for each of the plausible values, and the results were pooled (this was automatically implemented by setting the data type to IMPUTATION in Mplus).

We included the interaction between domain-specific self-concept and value as a predictor of course choices and study intentions in accordance with the assumptions of EVT. That is, for advanced mathematics course choice and STEM study intentions, we included the interaction between mathematics self-concept and value, and for advanced German course choice, we included the interaction between German self-concept and value. We used the latent moderated structural equation approach (LMS; Klein & Moosbrugger, 2000) as implemented in the software Mplus 8 (using the XWITH option) to estimate the interaction between the latent self-concept and value factors. The idea of the distribution-analytic LMS approach is to explicitly model the non-normal distribution produced by a latent interaction term through the mixture of normal distributions (Klein & Moosbrugger, 2000). The LMS approach thereby directly estimates the latent interaction and requires no product-term indicators to identify the latent interaction factor.

We used a robust maximum likelihood estimator (MLR) with a probit link function for the full model. There are no absolute fit indices for maximum likelihood estimation including categorical data because of the required numerical integration. Therefore, the CFAs described above in the measures section as well as the CFA producing the bivariate correlations (see Table 2) were conducted using the WLSMV estimator to obtain absolute fit indices (RMSEA, CFI). This allowed us to check the goodness of fit of the measurement models for the latent factors. In addition, we used relative model fit comparisons based on the AIC and the BIC to compare the full model to a model without dimensional comparisons (i.e., no cross-domain paths; see result section). The input files for all models reported in the manuscript can be found in Supplement A.

To better interpret the probit regression coefficients, we also computed predicted probabilities of course choices and study intentions for a selection of models using the formula provided in the Mplus User's Guide (Muthén & Muthén, 2015, p. 494). To account for the hierarchical structure of the data (students nested within classes), we corrected the standard errors using the TYPE = COMPLEX option in *Mplus* which implements a sandwich estimator (Muthen & Satorra, 1995).

Test of indirect effects. To test hypothesis 5.c, we estimated indirect effects of gender on course choices and the STEM study intention. Because of the plethora of the many, sometimes parallel mediational paths and because of binary variables serving both as outcomes and mediators, the formal test of indirect effects involved estimating a complex mediation model with a relatively small sample. In general, Mplus 8 offers two options for treating binary mediators (in this case, course choice as a mediator for effects on STEM study intentions). Either they can be treated as binary indicators of a continuous latent variable or the analysis can be conducted in the framework of causal mediation based on counterfactuals (Muthén & Asparouhov, 2015). The latter is best suited for single, one-step mediational paths and not for multi-step mediation with many different paths. Thus, we decided to use the continuous latent variable approach. Even for this approach, the computation of indirect effects was too complex to reach convergence with the MLR estimator. Therefore, we switched to a Bayesian estimation framework, with the weakly informative priors defined as a default in Mplus, for the computation of the indirect effects (Option ESTIMATOR = BAYES with MEDIATOR = LATENT and MODEL INDIRECT). This Bayesian approach facilitates the estimation of complex models in small-N scenarios where frequentist approaches no longer work. The Bayesian model reached convergence and produced meaningful estimates for indirect effects that were in line with the effect pattern we found for the overall model using the maximum likelihood approach. We report a selection of indirect effect estimates based on our hypotheses—(a) all total indirect effects of gender on self-concept, value and, choices and (b) all total indirect effects of achievement, self-concept, and value on choices.

Treatment of missing data. Of the 519 students in our sample, the questionnaire and achievement data from the T1 test session were missing for 37 students. Within the group that filled out the questionnaire, item-specific rates of missing data were very low (all < 2%). Grades as well as gender were reported for all students by the school officials. Information on course choices from the Grade 12 final report cards was present for 401 students, and information on study intent was present for 357 students. This was mostly due to report cards not being available for all students (they were collected in a separate step), or students not having a plan/idea yet. Some students also were no longer present in the school. We checked if this missingness at T1 was related to variables at T2. Indeed, for both constructs, we found that students with missing course choice or STEM intention data at T2 had lower grades and reading achievement scores at T1 suggesting a MAR process[2].

We estimated all models using *full information maximum likelihood* (FIML) estimation—a model-based approach for handling missing data. In the FIML procedure, missing data and

[2] We conducted t-tests for independent samples to check for mean differences between students with missing data in (a) course choice at T2 or (b) study intentions at T2 in (1) mathematics grades, (2) German grades, (3) mathematics test-scores and, (4) German test-scores at T1 resulting in a total of 8 comparisons. Students with missing data at T2 had lower grades and reading achievement scores at T1 (for all 6 comparisons: $p < .01$), but no lower mathematics test scores (missingness vs. non-missingness in course choices: $p = .29$; missingness vs. non-missingness in study intentions: $p = .28$)

parameter estimation are combined in a single step (Enders, 2010). Under the missing-at-random (MAR) assumption, FIML is considered superior to traditional methods for treating missing data such as listwise deletion because FIML allows for more a less biased parameter estimation with higher statistical power (Schafer & Graham, 2002). In all longitudinal models (except for the Bayesian model), we included students' socioeconomic background (measured with the HISEI as mentioned above) and cognitive ability as auxiliary variables to improve the FIML process. Cognitive ability was measured using a subscale of the *Kognitiver Fähigkeitstest (KFT)*, a well-established German test of general cognitive ability (Heller & Perleth, 2000).

Results

Descriptive Statistics

Of the 401 students for which information on course choices was available, 78 (23 girls and 55 boys) chose an advanced mathematics course, and 95 (62 girls and 33 boys) chose an advanced German course, whereas the remaining students chose other course options. Furthermore, 115 students (33 girls and 82 boys) of 357 students with valid responses reported the intention to pursue STEM studies at university. Means and standard deviations for achievement, self-concept, and value measures for the overall sample as well as for these groups can be found in Table 1. No ceiling or floor effects could be discerned for any of the variables.

Students with better mathematics respectively German grades as well as higher self-concept and value in Grade 9 were more likely to take an advanced course in that subject in Grade 12 (see Table 1). In addition, students choosing advanced mathematics courses generally showed higher mathematics test scores as well as higher German test-scores in comparison with the overall sample. Students with the intention to pursue a STEM subject after Grade 12 also showed higher mathematics self-concept and value. Furthermore, there were stereotypical gender differences: male students showed higher mathematics self-concepts and values and were more likely to choose advanced mathematics courses and pursue STEM studies, whereas female students showed higher German self-concept and value and were more likely to choose advanced German courses. Thus, these descriptive statistics already provide evidence for Hypotheses 5.a and 5.b.

The patterns described above were also evident in the bivariate correlations (see Table 2). All self-concept and value facets were significantly related to course choice and study intentions with positive correlations in each domain and no or negative correlations between domains (e.g., mathematics value and advanced German course choice). This already indicates that dimensional comparison effects may have played a role. Furthermore, as expected in EVT, self-concept and value facets in the same domain were highly correlated, but separable (e.g., mathematics self-concept and value: $r = .79$).

Dimensional Comparison Effects on Self-Concept and Value (H1)

The path coefficients for the full model (see Figure 1) can be found in Table 3, Model 1. Hypothesis 1 relates to the classic predictions of dimensional comparison theory: a pattern of positive within-domain paths from achievement to self-concept (e.g., mathematics grades to mathematics self-concept; H1.a) and negative cross-domain paths (e.g., German grades to mathematics self-concept and vice versa; H1.b). As expected, we found strong positive within-domain paths from grades to self-concept (math: $\beta = 0.85$, $p < .01$; German: $\beta = 0.75$, $p < .01$; see Table 3, Model 1) and value. There were also negative cross-domain effects from grades to self-concept and value for three of the four paths (all except for the path from German grades to mathematics value; Table 3, Model 1). For example, students with higher mathematics grades showed lower German self-concept and value after controlling for German grades (as well as test-scores and gender). As expected, the relations between test-scores and self-concept/value were generally lower than between grades and self-concept/value (H 1.c). This pertained to both within- and cross-domain relations. Still, positive within-domain paths and negative cross-domains effects could be observed on mathematics self-concept and value.

Whereas it is generally a strength of our study design that both grades and test-scores are included, it leads to a model that includes predictors with substantial correlations. This may have resulted in some counterintuitive results, such as the non-significant marginal effect of the reading test on German self-concept and value. As a robustness check, we also tested two models that included only grades (Table 3, Model 2) and only test-scores (Table 3, Model 3), respectively, as achievement measures. In the latter model, the relations between test-scores and self-concept/value were a bit stronger including a significant effect of the reading test on German self-concept and value.

Summing up, the assumptions of DCT were largely replicated for both self-concept and value (H1.a sustained and H1.b largely sustained) with the effects from grades being stronger than the effects from test-scores (H1.c sustained).

Prediction of Choices and Intentions (H2, H3 & H4)

Advanced course choice. Following EVT, we expected positive within-domain effects of self-concept, value, and their interaction on course choices in mathematics (H 2.a) and German (H2.b). In both domains, self-concept was a significant predictor of course choice, but there were no incremental effects of value and the interaction between self-concept and value (H2.a and H2.b partially confirmed). We further expected negative cross-domain effects from self-concept and value resulting from dimensional comparisons (H 3.a and 3.b). There was a negative cross-domain effect of German self-concept on mathematics course choice. There were no significant dimensional comparison effects on German course choice even though the path coefficient for value in math was substantially negative.

Regarding achievement, there was an incremental direct effect of mathematics test-scores on choosing an advanced math course, but no other direct within- or cross-domain effects of grades or test-scores on course choices. As described in the method section, we used a

Bayesian estimation approach to test a set of indirect effects. The results can be found in Table 4. It includes the direct effects (similar to the path coefficients presented in Table 3 based on the maximum likelihood estimation), the indirect effects, and the total effects (direct + indirect) of test-scores and grades on choices. We found significant positive indirect effects of grades on course choices within domains (e.g., math grades on choosing advanced math courses) and significant negative effects across domains (e.g., math grades on choosing advanced German courses). This can be interpreted as evidence for mediated dimensional comparison effects. For test scores, only the indirect effect of math test scores on advanced math course choice was significant. As can be seen from the full model (see Figure 1), these indirect effects were mediated through the four self-concept and value facets. They also resulted in significant total effects (see Table 4) for all paths from grades to course choices.

The path coefficients reported in Tables 3 and 4 are relatively easy to interpret for the linear parts of the model, but less intuitive for choice outcomes where probit estimates are used. Therefore, as an additional illustration of the total effect of a predictor, we analyzed the model-predicted probabilities of making a choice over a range of predictor values given that the other predictor values were set to zero. Zero refers to (a) the mean for continuous predictors, which were all standardized, or (b) the reference group for categorical predictors. Figure 2 shows the predicted probabilities of choosing advanced mathematics courses. A higher change in the predicted probability with changes in the predictor values (i.e., a steeper curve) can be interpreted as a relatively stronger effect. For example, for an increase in mathematics self-concept from the mean to one standard deviation above the mean (+1 SD), the predicted probability of selecting an advanced mathematics course increased by 16%. On the contrary, an increase in German self-concept by one standard deviation from -1 SD to the mean decreased the probability of selecting an advanced mathematics course by 6%. Figure 2 also shows a relatively strong total effect of mathematics test score on course choice that includes both a significant direct and indirect part (see Table 4). Like the selection of mathematics courses, domain-specific self-concept was also the best predictor of selecting an advanced German course with steep increases in the predicted probability of choosing German as German self-concept increased (see Figure 3).

Overall, we would argue that the results partly support the H2.a, H2.b, and H3.a while not providing evidence for H3.b. The overall pattern of path coefficients is consistent with the theoretical assumptions of positive within-domain and negative cross-domain effects even though the threshold for statistical significance was quite high due to the relatively small sample size and the substantial correlation between predictors.

STEM study intentions. Similar to advanced mathematics course choices, we expected STEM study intentions to be predicted by mathematics self-concept, value, and their interaction (H2.c). Contrary to course choices, STEM study intentions were best predicted by value with no incremental direct effects of academic self-concept or their interaction (H2.c partially confirmed). The negative path coefficient of valuing German was not significant (H3.c not confirmed). Beyond these, there were no significant direct incremental effects of course choice (H4.a and H4.b not confirmed), self-concept, achievement, or the interaction between self-concept and value.

Again, we tested indirect effects in the next step (see Table 4). We found significant indirect effects of self-concept that matched a dimensional comparison pattern (positive effects from mathematics self-concept, negative effects from German self-concept) mediated by advanced course choices. Furthermore, because the (non-significant) indirect effect estimates were in the same direction as the direct effects, there were significant total effects of both mathematics value (positive) and German value (negative) on STEM study intentions.

The predicted probabilities shown in Figure 4 also illustrate these total effects. Both the positive effect of valuing mathematics (an increase of 1 SD from the mean resulted in a predicted increase in the probability of 18%) and the negative effect of valuing German (an increase of 1 SD from the mean resulted in a predicted decrease in the probability of 8%) were quite substantial.

Overall, only H2.c could be partially confirmed, and valuing mathematics was the strongest predictor of STEM study intentions. Regarding dimensional comparison effects, H3.c could not be confirmed as there were no direct effects of self-concept and value in German; there was, however, a significant negative total effect of valuing German.

Direct and Indirect Gender Effects (H5)

The descriptive statistics already provided evidence for hypotheses 5.a and 5.b that focused on stereotypical gender differences (i.e., higher mathematics self-concepts, values, and advanced course choices as well as STEM study intentions for boys compared to girls; higher German self-concept, values, and advanced course choices for girls). Our full model allowed us to additionally study incremental direct effects of gender on high school course and intended university major choices over and above achievement, self-concept, and value and to get an impression of possible mediation patterns (i.e., gender effects being mediated by self-concept and value, H5.c).

First, we found positive effects of female gender on mathematics grades and, to an even larger extent, German grades. We interpret this pattern as an overlap of two well-known findings: stereotypical gender differences (female students showing higher language achievement compared to mathematics with male students showing the opposite pattern) and a general advantage in grades for female students (i.e., a main effect of gender on grades independent of the domain; also see Table 1).

Second, there were stereotypical gender effects on self-concept and value above and beyond achievement. Given similar grades and test-scores, male students showed higher mathematics self-concepts and valued the domain more than female students; female students valued German higher (additional evidence for H5.a and H5.b). In addition to these direct effects of gender, there were also additional indirect gender effects on German self-concept and value mediated through achievement (both in favor of girls; see Table 4).

Third, there were no direct effects of gender on advanced course choices. However, the effect pattern (i.e., mathematics self-concept being the strongest predictor of mathematics course choice and the presence of substantial gender differences in mathematics self-concept) suggested that there might be mediated effects of gender on course choices. Indeed, our estimation of indirect effects (see Table 4) showed that there were substantial indirect effects of gender on course choices in both domains—that is, the gender difference was fully mediated through achievement, self-concept, and value. Due to the high number of paths and the small sample size, we could not compare specific mediational paths, but a comparison of the indirect effects of gender on (a) self-concept and value and (b) course choice gives a first idea of possible mechanisms: The direct effect of gender on mathematics self-concept and value is stronger than the indirect effect through achievement (see Table 4). In addition, as mentioned above, mathematics self-concept has the strongest direct effect on mathematics course choices. Thus, most of the gender effect seems to be mediated through mathematics self-concept. For German, on the other hand, the indirect and direct effects of gender on self-concept are about equal in size. Thus, the gender effect on course choice seems to be mediated both through achievement and self-concept.

Finally, a substantial direct effect of gender on the intention to study a STEM subject remained after controlling for all indirect effects through self-concept, value, and advanced course choices. An additional analysis of the model-predicted probability of intending to study a STEM subject showed that it was half as high for female students (19%) as it was for male students (41%) given similar levels of achievement, self-concept, and value. In addition to the direct effect, there was a total indirect gender effect on STEM study intentions (mediated through achievement, self-concept, value, and course choices; see Table 4). Thus, we found consistent evidence for a full mediation of gender effects on course choices and a partial mediation of gender effects on study intentions (H5.c confirmed) through (combinations of) achievement, expectancy, and value.

Test of the Overall Contribution of Dimensional Comparison Effects

The overall model showed that several cross-domain paths were statistically significant, and an inspection of the predicted probabilities showed that they also contributed substantially to the prediction of the choice outcomes (through a combination of various direct and indirect effects). Still, as a last step, we aimed to do a more global test of the relevance of dimensional comparison effects in a complex longitudinal EVT model such as ours.

Therefore, we estimated another model similar to the full model (Table 3, Model 1), but without any cross-domain paths. That is, mathematics self-concept, for example, was predicted only by mathematics grades and test-scores, and mathematics course choice was predicted only by mathematics self-concept, value, grades, and test-scores, but not German self-concept, value, or achievement. Thus, this model did not include any dimensional comparison effects and had a total of 21 fewer paths than the overall model. The full model that included dimensional comparisons (AIC = 22874.224, BIC (sample-size adjusted) = 23030.490) showed a better fit (i.e., lower values) on both criteria than the model without cross-domain paths (AIC = 22961.266, BIC (sample-size adjusted) = 23094.90). This broad omnibus-style test thus shows that cross-domain paths representing dimensional comparisons add to the overall quality of the model.

Discussion

Summary

Within- and cross-domain effects on self-concept and value (H1). The relation between German and mathematics achievement and self-concept as well as value was characterized by strong within-domain relations (H1.a sustained) and a pattern of dimensional comparisons between domains (H1.b sustained except for the path from German grades to mathematics value) replicating classic predictions of the (generalized) I/E model (Marsh, 1986; Möller, 2016). Teacher-assigned grades showed stronger relations to self-concept and value than test-scores which was in line with our expectations based on the stronger salience of grades as achievement feedback (H1.c sustained).

Within-domain effects on course choice and STEM study intention (H2). Self-concept was positively related to course choices in both domains whereas value did not show incremental effects (H2.a and H2.b partly sustained). Regarding STEM study intentions, we found the opposite pattern with a positive effect of valuing mathematics and no incremental effect of mathematics self-concept (H2.c partly sustained). This pattern did not confirm our prediction of additive effects following the general idea of EVT. However, it may be explained by situational and contextual characteristics of the choices the students had to make. The German upper secondary course system implicates a grouping by achievement. Thus, it seems plausible that their self-perceived achievement would be a salient factor for students when they make course choice decisions. When choosing a subject to study, on the other hand, students have access to a higher range of options and the choice is more closely related to their occupational aspirations and general life goals. Therefore, it can be argued that the expectancy to do well in a subject is only a minimum requirement for choosing a study subject, but that the value attached to that subject should be more directly related to study choice (also see Guo, Parker, et al., 2015). Given that our items were originally intended to measure individual interest, this would also be in line with research arguing for importance of interests in generating intentions based on social cognitive career theory and showing that intentions are reciprocally related to interests (Grigg, Perera, McIlveen, & Svetleff, 2018). We think that this pattern of differential effects of self-concept vs. value on two choice-related outcomes is a central contribution of this study.

In the prediction of all choice outcomes, we also included the expectancy-value interaction and did not find an incremental effect (H2.c not sustained) unlike other research (Guo et al., 2017; Guo, Marsh, et al., 2015; Lauermann et al., 2017). A possible methodological explanation for this difference is that both test-scores and grades (and previous course choices for the prediction of STEM study intentions) in two domains were controlled for in our study. Furthermore, the interaction would usually be expected in addition to two separate positive main effects. As described above, the differential effect pattern we found included only significant main effects of either self-concept or value.

Cross-domain effects on course choice and STEM study intention (H3). Given similar mathematics self-concept (and achievement and value), students with lower German self-concepts were more likely to choose advanced mathematics courses (H3.b sustained). No similar significant effects were found for advanced German courses (H3.a not sustained). However, there were additional indirect dimensional comparison effects from grades to advanced course choices in both domains. Furthermore, given similar mathematics value and all other predictors being equal, students who valued German more were less likely to plan to study STEM subjects (H3.c sustained).

Course choice effects on study intentions (H4). We expected students who had chosen an advanced mathematics course to be more likely to pursue a STEM subject, whereas students in advanced German courses should be less likely. Whereas the bivariate correlations show this pattern suggesting dimensional comparisons, there were no significant incremental effects of course choice beyond achievement, self-concept, value, and gender on STEM study intentions.

Gender effects (H5). There were stereotypical gender differences in self-concepts and value (advantages for male students in mathematics, advantages for female students in German), and female students were less likely to choose advanced mathematics courses and to plan to study a STEM subject than male students. The gender differences in course choices could be explained by differences in domain-specific achievement, self-concepts, and value. However, an incremental direct effect of gender on STEM study intentions remained above and beyond all other predictors.

The Role of Dimensional Comparisons in EVT

Whereas dimensional comparisons have always been an implicit part of EVT (where a choice for one domain is a choice against another), until recently, few studies have modelled cross-domain effects of expectancy and value on educational choices (Cheryan et al., 2017; Wang et al., 2013). Even though some specific hypotheses were only partly supported in the overall model, our result pattern provides evidence that the predictions of EVT (expectancies and/or values influence choices), DCT (negative cross-domain effects of achievement on expectancies and values), and their combination (negative cross-domain effects of expectancies and values on choices) can be well-integrated in a synergistic way to predict choices. More specifically, our results are in line with previous results that showed that dimensional comparisons can affect course choice (Nagy et al., 2008, 2006) and career plans (Lauermann et al., 2015). Furthermore, we provided evidence for the steps in which they occur. We were able to show that there are (a) negative dimensional comparison effects of achievement on expectancy and value and (b) negative dimensional comparison effects of expectancy beliefs on course choice. An analysis of indirect effects showed that, thus, dimensional comparison effects from achievement to course choice are mediated through expectancy and value. That is, showing high achievement in several domains affects choices only if students indeed engage in dimensional comparison processes that affect their self-concepts and value. It should be noted though, that even though we could provide evidence for mediation in a statistical sense, achievement and self-concept/value were assessed at the same time, and this does not reply a causal directionality (also see limitations section below).

Prediction of (Gendered) STEM Study Intentions

Our results again show the importance of values for post-school choice, a finding that goes against the findings for high school course choices where domain-specific self-concepts were the best predictors. Looking at the effects on STEM study intentions, an implication would be to focus interventions on values. This idea is in line with recent research that showed that even short-term value interventions were effective in experimental field trials (Gaspard, Dicke, Flunger, Brisson et al., 2015; Harackiewicz, Canning, Tibbetts, Priniski, & Hyde, 2016; Hulleman & Harackiewicz, 2009).

We were further able to show that gender differences in high school course choices can be fully explained by gender differences in achievement, self-concepts, and values with no direct effects of gender remaining. Through dimensional comparisons, the higher motivation and achievement of female students in language domains also contribute to them being less likely to choose STEM domains. Subjectively, they seem to have a broader spectrum of choices due to a less specific ability profile. This has previously been suggested as a possible reason for gender differences in some STEM domains (Cheryan et al., 2017; Valla & Ceci, 2014; Wang et al., 2013). However, there was still an incremental effect of gender on STEM study intentions above and beyond previous choices, motivation, and achievement—a finding that is in line with recent results based on two Australian samples that also found remaining direct effects of gender on enrollment in STEM subjects (Marsh et al., 2019). Other factors that may affect study intentions and may at least partly explain the remaining gender effects include occupational interests, lifestyle values, personal and collective identity beliefs as well as stereotypes held by students, parents and teachers (e.g., Eccles, 2009; Heyder et al., 2017; Kessels, 2005; Muntoni & Retelsdorf, 2018, 2019; Retelsdorf, Schwartz, & Asbrock, 2015; for an overview, also see Wang & Degol, 2013).

Limitations and Further Research

Our study design had several advantages including (a) a longitudinal approach in which expectancy and values were measured before the course choices were made, (b) the prediction of course choices rather than coursework aspirations and (c) the availability of both school-reported grades and achievement tests. However, it also had a few limitations. First, our study is observational and, for the lack of experimental data, does not provide evidence for causal inferences. In particular, the direction of effects we assumed in our study was based on the assumptions and typical modelling approaches of EVT and DCT. Thus, following DCT, we assumed that achievement affects self-concept and values within and across domains. It should be noted, however, that the relation between achievement and student motivation is likely to be reciprocal (Reciprocal Effects Model; see, e.g. Marsh et al., 2018). Furthermore, EVT assumes that career aspirations and plans are affected by student motivation, but there is also research indicating a reciprocal relation such that career plans and aspirations can affect motivation (Lauermann et al., 2017). Future studies based on larger samples could aim to further expand this multi-domain perspective to a complete cross-lagged panel model (e.g., Marsh et al., 2018) in which achievement, motivation, and career plans are measured at each time point.

Furthermore, we could not study actual choices of careers or university majors because there was no post-school follow-up measurement, and we thus had to focus on students' intentions. More studies with a longitudinal approach across the transition from high school to university are desirable. In addition, STEM subjects vary with regard to the amount of mathematics that is required, and some non-STEM majors (e.g., economics) also require mathematics. Because only a total of 115 students intended to study a STEM major, we could not further differentiate between choice of major because of the sample size. Also, whereas the classification of a subject as part of the STEM field is relatively clear on the basis of the coding scheme we used (classification of subject groups by the Federal Statistical Office of Germany), the math-intensity of subjects is harder to code. Still, this differentiation between "hard" and "soft" STEM is also important for studying gender differences in motivational beliefs, and gender differences may vary substantially across, for example, different science domains such as biology, chemistry, and physics (Hardy, 2014; Jansen et al., 2014; Wang & Degol, 2013). Thus, future studies may include a broader set of STEM domains both on the predictor and the outcome side (e.g., Ochsenfeld, 2016).

Future studies should also include a broader range of value components. In our study, we used only a single, broad value variable corresponding to academic interest. This is a major limitation of our study in the face of the tendency of recent research to include many value facets. Whereas studying more facets would have been fruitful, it should be mentioned that other studies only used a broad measurement approach (Lauermann et al., 2015; Nagengast et al., 2011), and if we had included more, possibly highly interrelated value facets, this would have made an already complex model even harder to estimate, at least with our sample size. One way to reduce the complexity of results from the study of multiple expectancy and value facets in multiple domains as predictors of multinomial choices, could be person-centered analyses. Instead of modelling the effects of motivational characteristics in different domains separately, studies based on this method identify motivational profiles or classes of students and then compare these with regard to other characteristics such as gender, achievement or course/major choices (Chow, Eccles, & Salmela-Aro, 2012; Gaspard, Lauermann, Rose, Wigfield, & Eccles, 2019; Gaspard, Wille, Wormington, & Hulleman, 2019; Guo, Wang, Ketonen, Eccles, & Salmela-Aro, 2018). Depending on the profiles that are compared, dimensional comparisons can also be implicitly tested—for example, Gaspard et al. (2019) show that, given similar achievement, students with a "High Math/Low English" motivational profile were more likely to choose a STEM-related or math-intensive major than students with a "High Math/High English" profile (Gaspard, Wille, et al., 2019).

In our study, we considered mediated effects of gender on educational choices. A few studies also suggest, however, that gender might moderate the relation between expectancy and value facets and decision-making (e.g., utility value playing a stronger role for aspirations of female students regarding STEM; Watt et al., 2012). Whereas our sample is too small to test such hypotheses, it would be a worthwhile endeavor for future research.

The forced-choice nature of the German course system at the upper secondary level may trigger dimensional comparison processes particularly strongly in comparison with school systems in which students can decide to pursue advanced coursework in as many subjects as they want. This means that Germany offers an interesting context for studying dimensional comparison effects, but that generalizations to other contexts such as the US must be carefully validated. As mentioned above, the number of advanced courses a student may choose is limited in Germany, meaning that even students who do well in many subjects have to make a choice and specialize. On the other hand, courses may be chosen without prerequisites because an explicitly tracked secondary school system is already in place, and thus the general level of courses is more similar within schools in the German context. Hence, it would be plausible for self-concept and values to be more predictive of course choices in Germany and for achievement to be more predictive in the US. Nagy et al. (2008) found initial evidence for these hypotheses.

It should further be mentioned that we only examined students in the highest school track. Most students from the other tracks in Germany do not move up to the upper secondary level where the particular course choices we studied are made. Furthermore, choice mechanisms might be different in general in higher achieving student populations (e.g., more options may be available), and thus the results might not generalize to the lower achieving part of the student population. It should, however, be mentioned that we are not studying a particularly exclusive group as about 30 to 45 percent of students in Germany attend the academic track (depending on the federal state; Schipolowski, Stanat, Mahler, & Lenz, 2019), and that both EVT and DCT do not emphasize different choice mechanisms based on the general achievement level of students.

Our main recommendation for future studies on educational and career choices using the EVT framework would be to include cross-domain effects because we were able to show the relevance of dimensional comparisons for both direct and indirect predictions of specific choice outcomes.

References

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. doi:10.1016/j.cedpsych.2017.08.005

Arens, A. K., & Jansen, M. (2016). Self-concepts in reading, writing, listening, and speaking: A multidimensional and hierarchical structure and its generalizability across native and foreign languages. *Journal of Educational Psychology*, *108*(5), 646–664. doi:10.1037/edu0000081

Arens, A. K., & Möller, J. (2016). Dimensional comparisons in students' perceptions of the learning environment. *Learning and Instruction*, *42*, 22–30. doi:10.1016/j.learninstruc.2015.11.001

Ball, C., Huang, K.-T., Cotten, S. R., & Rikard, R. V. (2017). Pressurizing the STEM pipeline: An expectancy-value theory analysis of youths' STEM attitudes. *Journal of Science Education and Technology*, *26*(4), 372–382. doi:10.1007/s10956-017-9685-1

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some stem fields more gender balanced than others? *Psychological Bulletin*, *143*, 1-35. doi:10.1037/bul0000052

Chow, A., Eccles, J. S., & Salmela-Aro, K. (2012). Task value profiles across subjects and aspirations to physical and IT-related sciences in the United States and Finland. *Developmental Psychology*, *48*(6), 1612–1628. doi:10.1037/a0030194

Dickhäuser, O., Reuter, M., & Hilling, C. (2005). Coursework selection: A frame of reference approach using structural equation modelling. *British Journal of Educational Psychology*, *75*(4), 673–688. doi:10.1348/000709905X37181

Eccles, J. (2007). Where are all the women? Gender differences in participation in physical science and engineering. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't more* *women in science?: Top researchers debate the evidence* (pp. 199–210). Washington, DC, US: American Psychological Association.

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist*, *44*(2), 78–89. doi:10.1080/00461520902832368

Eccles, J. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development*, *35*(3), 195–201. doi:10.1177/0165025411398185

Eccles, J., Adler, T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 76–146). San Francisco, CA: W. H. Freeman.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*(3), 830–847. doi:10.1111/j.1467-8624.1993.tb02946.x

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103–127. doi:10.1037/a0018053

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

Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, *21*(1), 1–56. doi:10.1016/0049-089X(92)90017-B

Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Nagengast, B., & Trautwein, U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology*, *51*(9), 1226–1240. doi:10.1037/dev0000028

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

Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., & Nagengast, B. (2017). Assessing task values in five subjects during secondary school: Measurement structure and mean level differences across grade level, gender, and academic subject. *Contemporary Educational Psychology*, *48*, 67–84. doi:10.1016/j.cedpsych.2016.09.003

Gaspard, H., Lauermann, F., Rose, N., Wigfield, A., & Eccles, J. S. (2019). Cross-domain trajectories of students' ability self-concepts and intrinsic values in math and language arts. *Child Development*, cdev.13343. doi:10.1111/cdev.13343

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. doi:10.1016/j.cedpsych.2017.10.003

Gaspard, H., Wille, E., Wormington, S. V., & Hulleman, C. S. (2019). How are upper secondary school students' expectancy-value profiles associated with achievement and university STEM major? A cross-domain comparison. *Contemporary Educational Psychology*, *58*, 149–162. doi:10.1016/j.cedpsych.2019.02.005

Gentrup, S., & Rjosk, C. (2018). Pygmalion and the gender gap: Do teacher expectations contribute to differences in achievement between boys and girls at the beginning of schooling? *Educational Research and Evaluation*, *24*(3–5), 295–323. doi:10.1080/13803611.2018.1550840

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. doi:10.1016/j.cedpsych.2006.12.002

Gonzales, E., & Rutkowski, L. (2010). Principles of multiple matrix booklet designs and parameter recovery in large-scale assessments. *IERI Monograph Series*, *3*, 125–156.

Grigg, S., Perera, H. N., McIlveen, P., & Svetleff, Z. (2018). Relations among math self efficacy, interest, intentions, and achievement: A social cognitive perspective. *Contemporary Educational Psychology*, *53*, 73–86. doi:10.1016/j.cedpsych.2018.01.007

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. doi:10.1016/j.learninstruc.2016.12.007

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., & Yeung, A. S. (2015). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences*, *37*, 161–168. doi:10.1016/j.lindif.2015.01.008

Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology*, *51*(8), 1163–1176. doi:10.1037/a0039440

Guo, J., Wang, M.-T., Ketonen, E. E., Eccles, J. S., & Salmela-Aro, K. (2018). Joint trajectories of task value in multiple subject domains: From both variable- and pattern-centered perspectives. *Contemporary Educational Psychology*, *55*, 139–154. doi:10.1016/j.cedpsych.2018.10.004

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing achievement gaps with a utility-value intervention: disentangling race and social class.

Journal of Personality and Social Psychology, *111*, 745-765. doi:10.1037/pspp0000075

Hardy, G. (2014). Academic self-concept: modeling and measuring for science. *Research in Science Education*, *44*(4), 549–579. doi:10.1007/s11165-013-9393-7

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision: KFT 4-12+ R* [Cognitive ability test for grades 4 to 12]. Beltz-Test.

Heyder, A., & Kessels, U. (2013). Is School Feminine? Implicit Gender Stereotyping of School as a Predictor of Academic Achievement. *Sex Roles*, *69*(11–12), 605–617. doi:10.1007/s11199-013-0309-9

Heyder, A., Kessels, U., & Steinmayr, R. (2017). Explaining academic-track boys' underachievement in language grades: Not a lack of aptitude but students' motivational beliefs and parents' perceptions? *British Journal of Educational Psychology*, *87*(2), 205–223. doi:10.1111/bjep.12145

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting Interest and Performance in High School Science Classes. *Science*, *326*(5958), 1410–1412. doi:10.1126/science.1177067

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, *65*(1), 373–398. doi:10.1146/annurev-psych-010213-115057

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*(2), 139–155.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender Comparisons of Mathematics Attitudes and Affect. *Psychology of Women Quarterly*, *14*(3), 299–324. doi:10.1111/j.1471-6402.1990.tb00022.x

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*(5888), 494–495. doi:10.1126/science.1160364

Jansen, M., Schroeders, U., & Lüdtke, O. (2014). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences*, *30*, 11–21. doi:10.1016/j.lindif.2013.12.003

Jansen, M., Schroeders, U., Lüdtke, O., & Marsh, H. W. (2015). Contrast and assimilation effects of dimensional comparisons in five subjects: An extension of the I/E model. *Journal of Educational Psychology*, *107*(4), 1086–1101. doi:10.1037/edu0000021

Jansen, M., Schroeders, U., Lüdtke, O., & Marsh, H. W. (2019). The dimensional structure of students' self-concept and interest in science depends on course composition. *Learning and Instruction*, *60*, 20–28. doi:10.1016/j.learninstruc.2018.11.001

Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education*, *20*(3), 309–323. doi:10.1007/BF03173559

Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*(4), 457–474. doi:10.1007/BF02296338

Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, *12*(4), 383–409. doi:10.1016/S0959-4752(01)00011-1

Lauermann, F., Chow, A., & Eccles, J. S. (2015). Differential effects of adolescents' expectancy and value beliefs about math and English on math/science-related and human services-related career plans. *International Journal of Gender, Science and Technology*, *7*(2), 205–228.

Lauermann, F., Tsai, Y.-M., & Eccles, J. S. (2017). Math-related career aspirations and choices within eccles et al.'s expectancy–value theory of achievement-related behaviors. *Developmental Psychology.53,* 1540-1559. doi:10.1037/dev0000367

Ma, X., & Johnson, W. (2008). Mathematics as the critical filter: Curricular effects on gendered career choices. In H. M. G. Watt & J. S. Eccles (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences.* (pp. 55–83). Washington: American Psychological Association. doi:10.1037/11706-002

Maaz, K., Baumert, J., Neumann, M., Becker, M., & Dumont, H. (Eds.) (2013). *Die Berliner Schulstrukturreform. Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen* [The secondary school reform in Berlin: Acceptance by the stakeholders and consequences of the new transition system from primary to secondary school]. Münster, Germany: Waxmann.

Mahler, N., & Kölm, J., (2019). Soziale Disparitäten [Social Disparities]. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich & S. Henschel (Eds..), IQB-Bildungstrend 2018. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I im zweiten Ländervergleich (S. 337-354). Münster: Waxmann.

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, *23*(1), 129–149. doi:10.3102/00028312023001129

Marsh, H. W., & Hau, K.-T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology*, *96*(1), 56–67. doi:10.1037/0022-0663.96.1.56

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. doi:10.1016/j.cedpsych.2014.08.003

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.

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. doi:10.1037/dev0000393

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416.

Marsh, H. W., Van Zanden, B., Parker, P. D., Guo, J., Conigrave, J., & Seaton, M. (2019). Young women face disadvantage to enrollment in university stem coursework regardless of prior achievement and attitudes. *American Educational Research Journal*, *56*(5), 1629–1680. doi:10.3102/0002831218824111

Möller, J. (2016). The generalized internal/external frame of reference model: an extension to dimensional comparison theory. *Frontline Learning Research*, *4*(4), 1–11. doi:10.14786/flr.v4i2.169

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. doi:10.1037//0022-0663.93.4.826

Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, *120*(3), 544–560. doi:10.1037/a0032459

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, *79*(3), 1129–1167. doi:10.3102/0034654309337522

Möller, J., Retelsdorf, J., Köller, O., & Marsh, H. W. (2011). The reciprocal internal/external frame of reference model an integration of models of relations between academic achievement and self-concept. *American Educational Research Journal*, *48*(6), 1315–1346. doi:10.3102/0002831211419649

Muntoni, F., & Retelsdorf, J. (2018). Gender-specific teacher expectations in reading—The role of teachers' gender stereotypes. *Contemporary Educational Psychology*, *54*, 212–220. doi:10.1016/j.cedpsych.2018.06.012

Muntoni, F., & Retelsdorf, J. (2019). At their children's expense: How parents' gender stereotypes affect their children's reading outcomes. *Learning and Instruction*, *60*, 95–103. doi:10.1016/j.learninstruc.2018.12.002

Muthén, B., & Asparouhov, T. (2015). Causal effects in mediation modeling: an introduction with applications to latent variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(1), 12–23. doi:10.1080/10705511.2014.935843

Muthen, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*, 267. doi:10.2307/271070

Muthén, L. K., & Muthén, B. O. (2015). *Mplus User's Guide. . Seventh Edition.* Los Angeles, CA: Muthén & Muthén.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K.-T., & Trautwein, U. (2011). Who took the "x" out of expectancy-value theory?: A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science*, *22*(8), 1058–1066. doi:10.1177/0956797611415540

Nagengast, B., Trautwein, U., Kelava, A., & Lüdtke, O. (2013). synergistic effects of expectancy and value on homework engagement: the case for a within-person perspective. *Multivariate Behavioral Research*, *48*(3), 428–460. doi:10.1080/00273171.2013.775060

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). Washington, DC, US: American Psychological Association.

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. doi:10.1080/13803610600765687

Neumann, M., Becker, M., Baumert, J., Maaz, K., & Köller, O. (Eds.). (2017). Zweigliedrigkeit im deutschen Schulsystem. Potenziale und Herausforderungen in Berlin [Tracking in the German school system: Opportunities and Challenges in Berlin]. Münster: Waxmann.

Ochsenfeld, F. (2016). Preferences, constraints, and the process of sex segregation in college majors: A choice analysis. *Social Science Research*, *56*, 117–132. doi:10.1016/j.ssresearch.2015.12.008

Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., & Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: A multicontext study. *Developmental Psychology*, *48*(6), 1629–1642. doi:10.1037/a0029167

Perera, H. N., & McIlveen, P. (2018). Vocational interest profiles: Profile replicability and relations with the STEM major choice and the Big-Five. *Journal of Vocational Behavior*, *106*, 84–100. doi:10.1016/j.jvb.2017.11.012

Reilly, D., Neumann, D. L., & Andrews, G. (2014). Sex Differences in Mathematics and Science Achievement: A Meta-Analysis of National Assessment of Educational Progress Assessments. *Journal of Educational Psychology*. doi:10.1037/edu0000012

Retelsdorf, J., Schwartz, K., & Asbrock, F. (2015). “Michael can’t read!” Teachers’ gender stereotypes and boys’ reading self-concept. *Journal of Educational Psychology*, *107*(1), 186–194. doi:10.1037/a0037107

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. doi:10.1037//1082-989X.7.2.147

Schipolowski, S., Mahler, N., Stanat, P., & Lenz, S. (2019). Kontextinformationen zu den Schulsystemen der Länder in der Sekundarstufe I [Context information about the school systems at the lower secondary level in the federa states]. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich & S. Henschel (Eds..), IQB-Bildungstrend 2018. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I im zweiten Ländervergleich (S. 337-354). Münster: Waxmann.

Schoon, I., & Eccles, J. S. (Eds.). (2014). *Gender Differences in Aspirations and Attainment: A Life Course Perspective*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139128933

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

Schwanzer, A. D., Trautwein, U., Lüdtke, O., & Sydow, H. (2005). Entwicklung eines Instruments zur Erfassung des Selbstkonzepts junger Erwachsener [Development of an instrument to assess self-concepts of young adults]. *Diagnostica*, *51*(4), 183–194. doi:10.1026/0012-1924.51.4.183

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology*, *42*(1), 70–83. doi:10.1037/0012-1649.42.1.70

Simpkins, S. D., Fredricks, J. A., & Eccles, J. S. (2012). Charting the Eccles' expectancy-value model from mothers' beliefs in childhood to youths' activities in adolescence. *Developmental Psychology*, *48*(4), 1019–1032. doi:10.1037/a0027468

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology*, *104*(3), 763–777. doi:10.1037/a0027470

Valla, J. M., & Ceci, S. J. (2014). Breadth-based models of women's underrepresentation in stem fields an integrative commentary on Schmidt (2011) and Nye et al. (2012). *Perspectives on Psychological Science*, *9*(2), 219–224. doi:10.1177/1745691614522067

Van Zanden, B., Marsh, H. W., Seaton, M., Parker, P. D., Guo, J., & Duineveld, J. J. (2017). How well do parents know their adolescent children? Parent inferences of student self-concepts reflect dimensional comparison processes. *Learning and Instruction*, *47*, 25–32. doi:10.1016/j.learninstruc.2016.09.002

Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. *IERI Monograph Series*, *2*, 9–36.

Wang, M.-T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*(4), 304–340. doi:10.1016/j.dr.2013.08.001

Wang, M.-T., Eccles, J. S., & Kenny, S. (2013). Not lack of ability but more choice individual and gender differences in choice of careers in Science, Technology, Engineering, and

Mathematics. *Psychological Science*, *24*(5), 770–775. doi:10.1177/0956797612458937

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*(3), 427–450.

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th- through llth-grade Australian students. *Child Development*, *75*(5), 1556–1574. doi:10.1111/j.1467-8624.2004.00757.x

Watt, H. M. G., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., & Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from Australia, Canada, and the United States. *Developmental Psychology*, *48*(6), 1594–1611. doi:10.1037/a0027838

Wigfield, A., & Eccles, J. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. doi:10.1006/ceps.1999.1015

Wolff, F., Helm, F., Zimmermann, F., Nagy, G., & Möller, J. (2018). On the Effects of Social, Temporal, and Dimensional Comparisons on Academic Self-Concept. *Journal of Educational Psychology*, 1005–1025. doi:10.1037/edu0000248

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, *31*(2–3), 114–128. doi:10.1016/j.stueduc.2005.05.005

Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist*, *70*(1), 10–20. doi:10.1037/a0038208

Table 1
Means and Standard Deviations across Course Groups

	Total sample (***N*** = 519)		Math AC Grade 12(*N* = 78[a])		German AC Grade 12 (***N*** = 95[a])		STEM study intentions Grade 12 (***N*** = 115)		Female (N =255)		Male (N = 264)	
Variable	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Math ASC (Grade 9)	2.94	0.76	3.72	0.36	2.65	0.76	3.28	0.66	2.76	0.78	3.12	0.70
German ASC (Grade 9)	3.12	0.55	3.04	0.58	3.46	0.52	3.09	0.59	3.31	0.50	3.07	0.57
Math VAL (Grade 9)	2.55	0.75	3.16	0.64	2.20	0.70	2.93	0.72	2.47	0.75	2.63	0.75
German VAL (Grade 9)	2.51	0.64	2.25	0.57	2.82	0.65	2.37	0.60	2.71	0.62	2.32	0.60
Math grade (Grade 9)	4.14	0.97	4.86	0.86	4.11	0.81	4.45	0.88	4.24	0.97	4.05	0.98
German grade (Grade 9)	4.25	0.89	4.40	0.90	4.54	0.77	4.22	0.94	4.50	0.80	4.00	0.89
Math test (Grade 9)	0.00	1.00	0.74	0.89	-0.14	0.98	0.35	0.15	-0.17	0.96	0.17	1.01
German test (Grade 9)	0.00	1.00	0.33	0.88	0.10	0.96	1.08	1.06	0.09	0.97	-0.09	1.02
Gender female	49% (N = 255)		29% (N = 23)		65% (N = 62)		29% (33)		/	/	/	/

Note. Means and standard deviations for self-concept and value refer to manifest scale means. ASC = academic self-concept, VAL = value, AC = advanced course students attended in Grade 12 (course choice made after Grade 10), math = mathematics. Grades originally ranged from 1 (*excellent*) to 6 (*insufficient*), but were reverse-coded so that higher values here and in all further analyses represent higher achievement. Test scores refer to an IRT-based measure of person ability (five plausible values; see Method section); the plausible values were standardized after IRT scaling.

[a] information on course choice was available for only 401 students.

Table 2
Correlations and Standard Errors (in Parentheses)

	Math ASC	German ASC	Math VAL	German VAL	Math grade	German grade	Math Test	Reading Test	Math AC	German AC	STEM Intention
Math ASC											
German ASC	-.08 (.05)										
Math VAL	.79 (.02)*	-.14 (.05)*									
German VAL	-.25 (.05)*	.68 (.03)*	-.03 (.06)								
Math grade	.59 (.04)*	.09 (.05)	.43 (.05)*	-.05 (.05)							
German grade	.12 (.04)*	.51 (.03)*	.07 (.04)	.36 (.04)*	.50 (.04)*						
Test math	.43 (.05)*	.09 (.05)	.25 (.06)*	-.15 (.05)*	.41 (.04)*	.30 (.05)*					
Test German	.14 (.05)*	.26 (.04)*	.03 (.05)	.07 (.06)	.29 (.05)*	.39 (.05)*	.63 (.04)*				
Math AC	.81 (.04)*	-.24 (.08)*	.62 (.06)*	-.34 (.08)*	.43 (.07)*	-.01 (.08)	.50 (.05)*	.18 (.09)*			
German AC	-.36 (.08)*	.46 (.07)*	-.39 (.08)*	.41 (.08)*	-.17 (.08)*	.13 (.07)	-.16 (.08)	.00 (.09)	-.70 (.16)*		
STEM intentions	.42 (.06)*	-.21 (.08)*	.44 (.06)*	-.31 (.07)*	.19 (.06)*	-.08 (.07)	.24 (.06)*	.06 (.07)	.52 (.06)*	-.43 (.10)*	
Gender female	-.32 (.06)*	.32 (.06)*	-.17 (.05)*	.42 (.05)*	.12 (.06)*	.34 (.05)*	-.21 (.06)*	.11 (.06)	-.40 (.07)*	.24 (.09)*	-.47 (.09)*

Note. ASC = academic self-concept; VAL = value belief; AC =Advanced course. Because the correlation matrix includes dichotomous variables, we employed the WLSMV estimator (i.e., polychoric correlations were used). Self-concepts, values, grades, and test-scores were measured in Grade 9. Course choice refers to the advanced course attendance of students in Grade 12 (with the choice already having been made after Grade 10). STEM intention was measured in Grade 12. Model fit indices of the CFA model: χ^2 (240) = 398.48, $p < .06$; CFI = .986, TLI = .981, RMSEA = .036).

- $p < .05$.

Table 3
Selected Hypotheses and Path Coefficients for Estimated Model

			Hypothesis		Model 1	Model 2	Model 3
			Number	Direction	β (*SE*)	β (*SE*)	β (*SE*)
Math grade	→	Math ASC	1.a	+	0.85 (0.09)*	0.91 (0.09)*	
German grade	→	Math ASC	1.b	-	-0.18 (0.06)*	-0.16 (0.06)*	
Math grade	→	German ASC	1.b	-	-0.26 (0.06)*	-0.26 (0.06)*	
German grade	→	German ASC	1.a	+	0.72 (0.08)*	0.75 (0.07)*	
Math test	→	Math ASC	1.a	+	0.36 (0.09)*		0.56 (0.09)*
German test	→	Math ASC	1.b	-	-0.18 (0.08)*		-0.17 (0.07)*
Math test	→	German ASC	1.b	-	-0.05 (0.12)		-0.01 (0.10)
German test	→	German ASC	1.a	+	0.12 (0.09)		0.26 (0.09)*
Math grade	→	Math VAL	1.a	+	0.58 (0.08)*	0.61 (0.09)*	
German grade	→	Math VAL	1.b	-	-0.14 (0.08)	-0.16 (0.07)*	
Math grade	→	German VAL	1.b	-	-0.28(0.06)*	-0.34 (0.07)*	
German grade	→	German VAL	1.a	+	0.53 (0.09)*	0.51 (0.08)*	
Math test	→	Math VAL	1.a	+	0.22 (0.08)*		0.40 (0.08)*
German test	→	Math VAL	1.b	-	-0.21 (0.09)*		-0.22 (0.07)*
Math test	→	German VAL	1.b	-	-0.23 (0.10)*		-0.22 (0.09)*
German test	→	German VAL	1.a	+	0.10 (0.11)		0.21 (0.10)*
Math grade	→	Math AC			-0.12 (0.22)	-0.05 (0.20)	
German grade	→	Math AC			0.08 (0.16)	0.16 (0.15)	
Math test	→	Math AC			0.40 (0.19)*		0.39 (0.18)*
German test	→	Math AC			0.03 (0.19)		0.03 (0.19)
Math ASC	→	Math AC	2a	+	0.77 (0.24)*	0.82 (0.22)*	0.91 (0.20)*
German ASC	→	Math AC	3b	-	-0.38 (0.15)*	-0.33 (0.15)*	-0.40 (0.16)*
Math VAL	→	Math AC	2a	+	0.26 (0.19)	0.25 (0.18)	0.29 (0.22)
German VAL	→	Math AC	3b	-	-0.23 (0.14)	-0.26 (0.15)*	-0.24 (0.15)
Math ASCxVAL	→	Math AC	2a	+	0.15 (0.11)	0.17 (0.09)*	0.17 (0.15)
Math grade	→	German AC			0.11 (0.14)	0.10 (0.14)	
German grade	→	German AC			-0.14 (0.11)	-0.17 (0.11)	
Math test	→	German AC			-0.02 (0.13)		-0.01 (0.13)
German test	→	German AC			-0.10 (0.12)		-0.12 (0.12)
Math ASC	→	German AC	3a	-	-0.14 (0.14)	-0.16 (0.16)	-0.16 (0.16)
German ASC	→	German AC	2b	+	0.42 (0.13)*	0.40 (0.14)*	0.42 (0.15)*
Math VAL	→	German AC	3a	-	-0.25 (0.16)	-0.24 (0.17)	-0.28 (0.18)
German VAL	→	German AC	2b	+	0.14 (0.17)	0.13 (0.17)	0.15 (0.18)
German ASCxVAL	→	German AC	2b	+	0.01 (0.09)	0.02 (0.09)	0.02 (0.11)
Math grade	→	STEM intention			0.03 (0.12)	0.04 (0.12)	
German grade	→	STEM intention			0.01 (0.12)	0.03 (0.12)	
Math test	→	STEM intention			0.03 (0.16)		0.03 (0.16)

German test	→	STEM intention			0.06 (0.11)		0.07 (0.11)
Math ASC	→	STEM intention	2c	+	-0.13 (0.14)	-0.13 (0.13)	-0.12 (0.13)
German ASC	→	STEM intention	3c	-	0.02 (0.12)	0.04 (0.12)	0.03 (0.12)
Math VAL	→	STEM intention	2c	+	0.44 (0.15)*	0.45 (0.15)*	0.46 (0.16)*
German VAL	→	STEM intention	3c	-	-0.21 (0.11)*	-0.22 (0.11)*	-0.21 (0.11)*
Math ASCxVAL	→	STEM intention	2c	+	0.00 (0.05)	0.00 (0.05)	0.05 (0.07)
math AC	→	STEM intention	4a	+	0.31 (0.22)	0.32 (0.21)	0.29 (0.21)
German AC	→	STEM intention	4b	-	-0.23 (0.32)	-0.22 (0.33)	-0.25 (0.32)
Gender female	→	Math grade			0.20 (0.09)*	0.20 (0.09)*	
	→	German grade			0.56 (0.08)*	0.56 (0.08)*	
	→	Math test			-0.34 (0.10)*		-0.34 (0.10)*
	→	German test			0.18 (0.10)		0.18 (0.10)
	→	Math ASC	5a	-	-0.58 (0.12)*	-0.74 (0.11)*	-0.28 (0.12)*
	→	German ASC	5b	+	0.20 (0.13)	0.21 (0.11)*	0.49 (0.12)*
	→	Math VAL	5a	-	-0.26 (0.14)*	-0.36 (0.14)*	-0.11 (0.11)
	→	German VAL	5b	+	0.43 (0.14)*	0.54 (0.13)*	0.62 (0.12)*
	→	math AC	5a	-	-0.02 (0.26)	-0.17 (0.24)	-0.06 (0.24)
	→	German AC	5b	+	-0.04 (0.19)	-0.03 (0.20)	-0.07 (0.21)
	→	STEM intention	5a	-	-0.63 (0.21)*	-0.63 (0.21)*	-0.60 (0.20)*

Note. ASC = academic self-concept; VAL = Value; AC =Advanced course. Grades were reverse-coded so that higher grades reflected higher achievement. Standard errors are shown in parentheses. The MLR Estimator with LINK = PROBIT was used for all models. All paths refer to nonstandardized coefficients. Grades and achievement tests were standardized prior to the analyses. Self-concepts, values, grades, and test-scores were measured in Grade 9. Course choice refers to the advanced course attendance of students in Grade 12 (with the choice already having been made after Grade 10). STEM intention was measured in Grade 12. Hypotheses 1c and 5c are not shown as they refer to more complex effect pattern rather than a single path. Syntax files for the models can be found in Supplement A.

- *$p < .05$.*

Table 4
Tests of indirect effects on educational choices.

Predictor	Outcome	Est. (95% CI)		
		Total indirect effect	Direct effect	Total effect
Gender female	STEM intention	-0.39 (-1.00, 0.07)*	-0.45 (-1.51, -0.25)*	-0.83 (-2.13, 0.65)*
Gender female	Math AC	-1.13 (-2.00, -0.78)*	0.13 (-0.52, 0.47)	-1.00 (-2.04, -0.74)*
Gender female	German AC	0.52 (0.27, 0.78)*	0.03 (-0.43, 0.31)	0.54 (0.13, 0.79)*
Gender female	Math ASC	-0.13 (-0.45, 0.32)	-1.18 (-1.82, -0.78)*	-1.31 (-1.96, -0.76)*
Gender female	German ASC	0.22 (0.15., 0.39)*	0.21 (0.03, 0.40)*	0.43 (0.29, 0.68)*
Gender female	Math VAL	-0.02 (-0.10, 0.07)	-0.16 (-0.34, -0.06)*	-0.18 (-0.36, -0.07)*
Gender female	German VAL	0.21 (0.12, 0.33)*	0.52 (0.23, 0.63)*	0.73 (0.43, 0.86)*
Math ASC	STEM intention	0.25 (0.01,1.58)*	-0.29 (-1.78. -0.02)*	-0.05 (-0.28, 0.06)
German ASC	STEM intention	-0.27 (-2.21, -0.01)*	0.47 (-0.11, 2.43)	0.20 (-0.31, 0.41)
Math VAL	STEM intention	0.35 (-0.10, 0.74)	0.53 (0.15, 1.84)*	0.89 (0.42, 2.15)*
German VAL	STEM intention	-0.19 (-0.54,0.13)	-0.40 (-0.83, 0.22)	-0.58 (-0.93, -0.02)*
Math grade	Math AC	1.05 (0.88, 2.00)*	-0.11 (-0.59, 0.15)	0.94 (0.75, 1.65)*
German grade	Math AC	-0.49 (-1.15, -0.40)*	-0.03 (-0.21, 0.49)	-0.51 (-0.94, -0.29)*
Math test	Math AC	0.29 (0.12, 0.52)*	0.27 (0.07, 0.66)*	0.56 (0.33, 1.02)*
German test	Math AC	-0.10 (-0.29, 0.03)	0.12 (-0.09, 0.39)	0.02 (-0.24, 0.29)
Math grade	German AC	-0.44 (-0.69, -0.29)*	0.14 (-0.13, 0.40)	-0.30 (-0.56, -0.14)*
German grade	German AC	0.40 (0.30, 0.68)*	-0.10 (-0.43, 0.09)	0.30 (0.11, 0.54)*
Math test	German AC	-0.10 (-0.16, 0.03)	-0.13 (-0.24, 0.16)	-0.23 (-0.30, 0.10)
German test	German AC	0.12 (-0.04, 0.12)	-0.03 (-0.28, 0.13)	0.09 (-0.24, 0.18)
Math grade	STEM intention	0.28 (-0.36, 0.47)	0.03 (-0.14, 1.04)	0.32 (0.15, 0.85)*
German grade	STEM intention	-0.24 (-0.39,0.37)	-0.01 (-0.74, 0.24)	-0.25 (-0.53, 0.04)
Math test	STEM intention	0.28 (0.03, 0.92)*	-0.19 (-0.61, 0.19)	0.09 (-0.06, 0.47)
German test	STEM intention	-0.02 (-0.14, 0.35)	0.06 (-0.52, 0.13)	0.04-0.35, 0.15)

Note. All effects are based on the specification in Table 3 (except for the interaction effects, see below); the direct effects thus represent marginal/conditional direct effects. The mean of the posterior distribution of each parameter is used as the point estimate. We used 100,000 iterations. Because the Bayesian estimator in *Mplus* cannot handle multiply imputed datasets, we used a different IRT-based indicator of person ability, the weighted likelihood estimate (Warm, 1989) for German and mathematics test-scores. Furthermore, the Bayesian estimator cannot be combined with a latent interaction factor. However, these interactions were not significant in the full model (see Table 3). All effects are marginal (see Table 3; thus, controlling for all other variables).

*$p < 0.05$ for positive estimates or $p > .95$ for negative estimates (Bayesian p-value, one-tailed; i.e., proportion of parameter estimates through iterations that are positive)

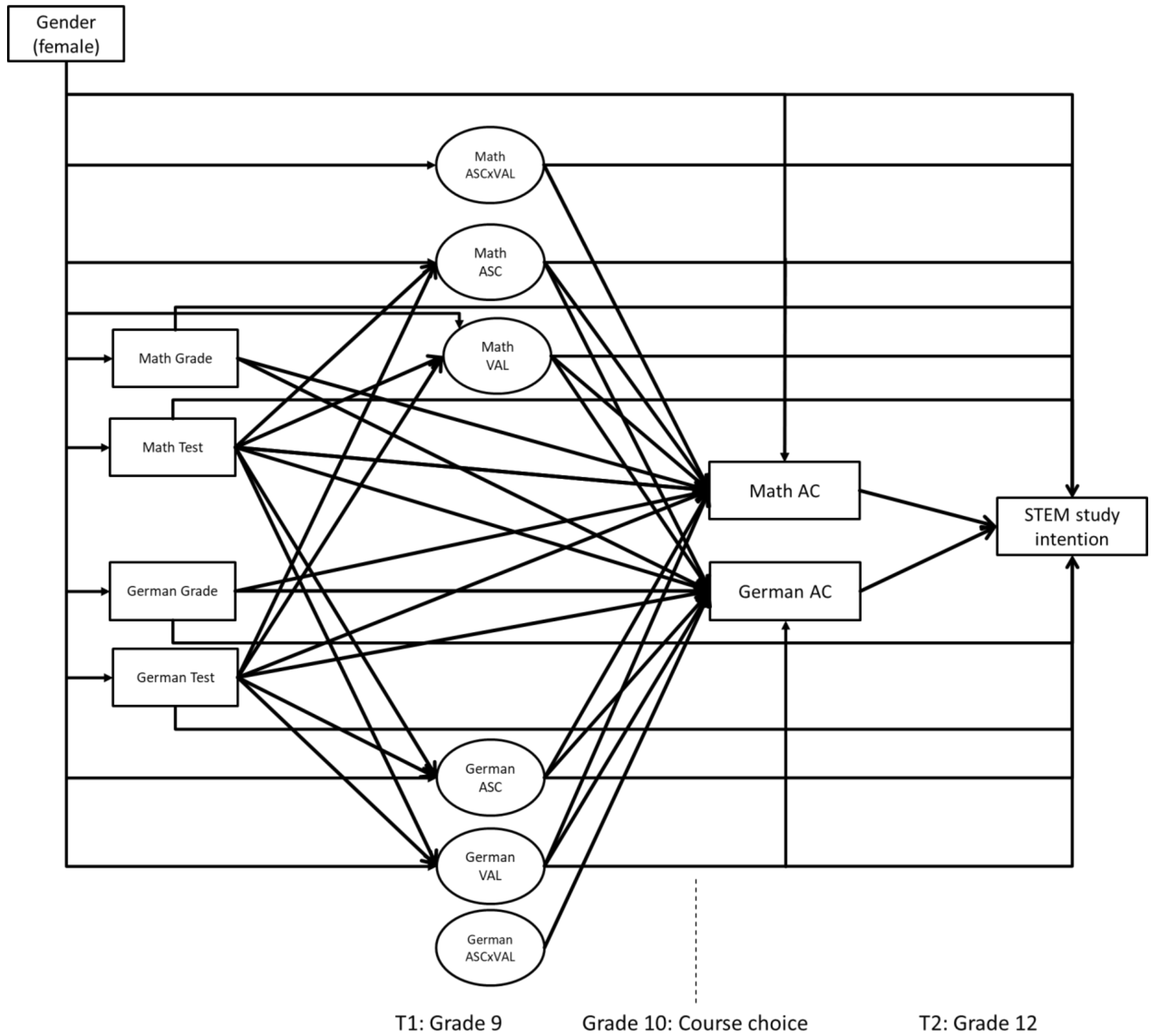

Figure 1. Overview of estimated path models. Full measurement models for the self-concept and value factors were estimated but omitted for clarity. Similarly, residual variances were estimated but omitted for clarity. ASC = academic self-concept, VAL = value, AC = advanced course choice. For the path coefficients for this model, see Table 3.

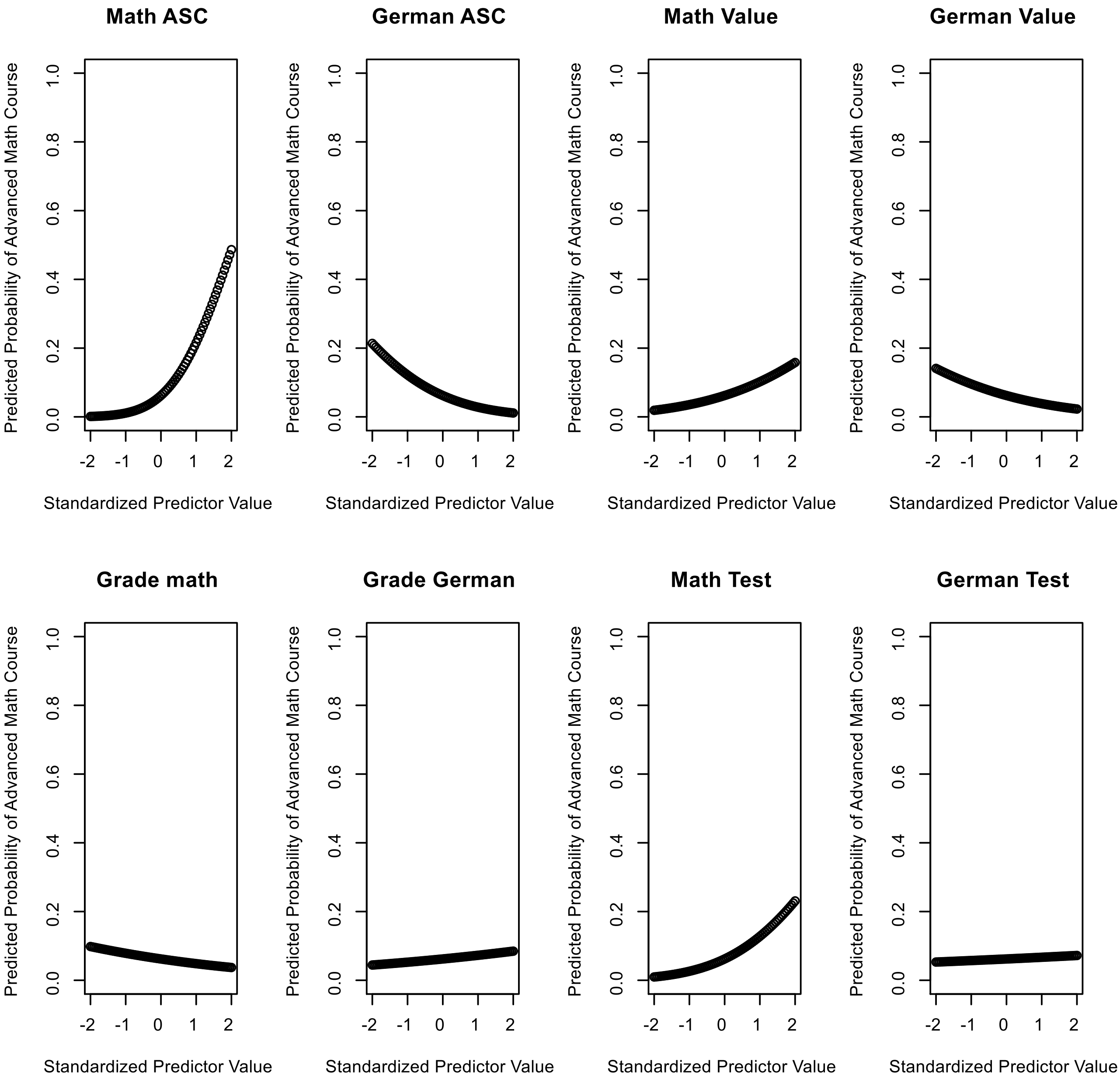

Figure 2. Predicted probabilities for taking an advanced math course for different standardized values of a given predictor based on the full model (Table 3). All other continuous predictors were set to the sample mean and the dichotomous predictors to the reference category.

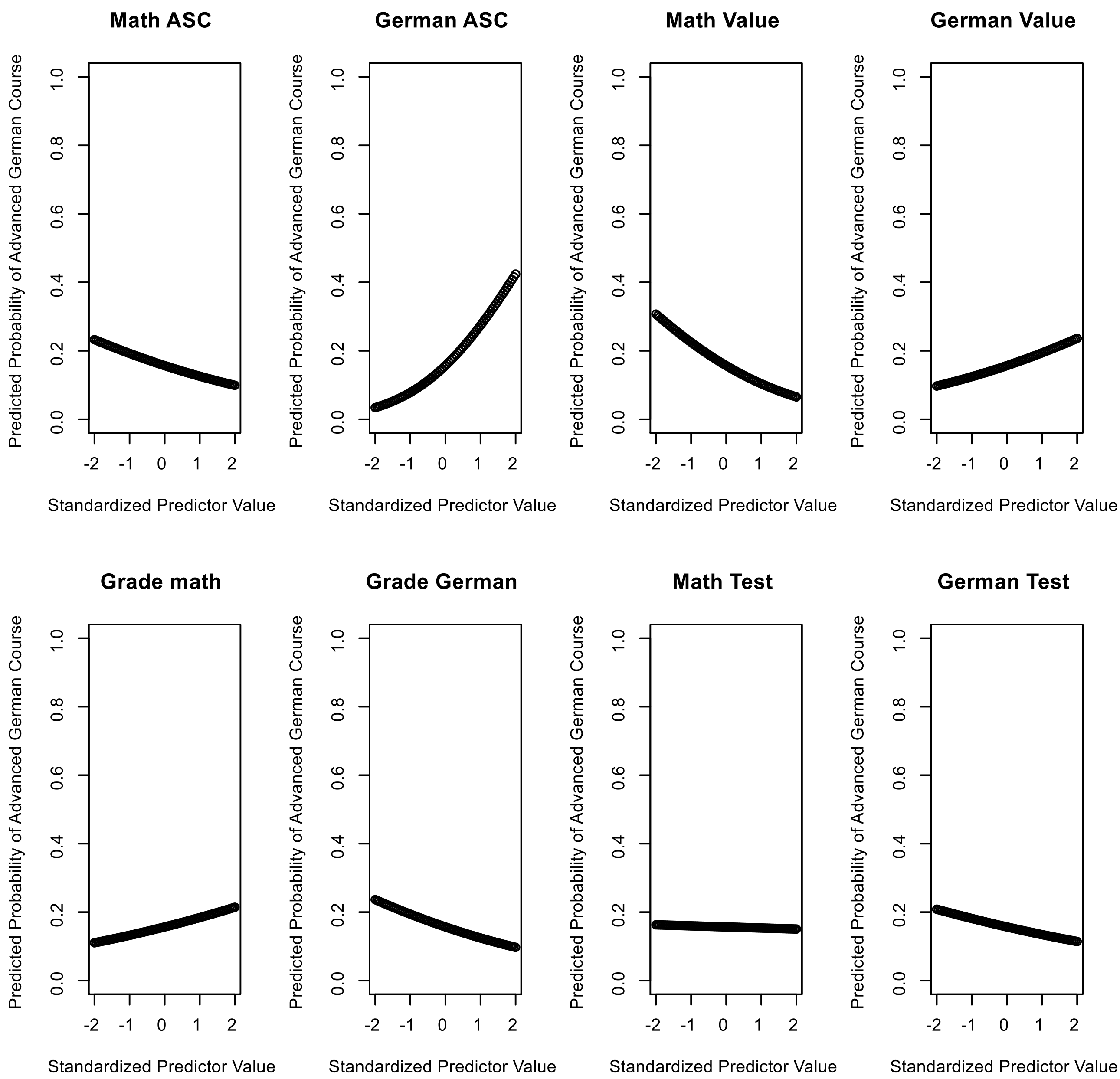

Figure 3. Predicted probabilities for taking an advanced German course for different standardized values of a given predictor based on the full model (Table 3). All other continuous predictors were set to the sample mean and the dichotomous predictors to the reference category.

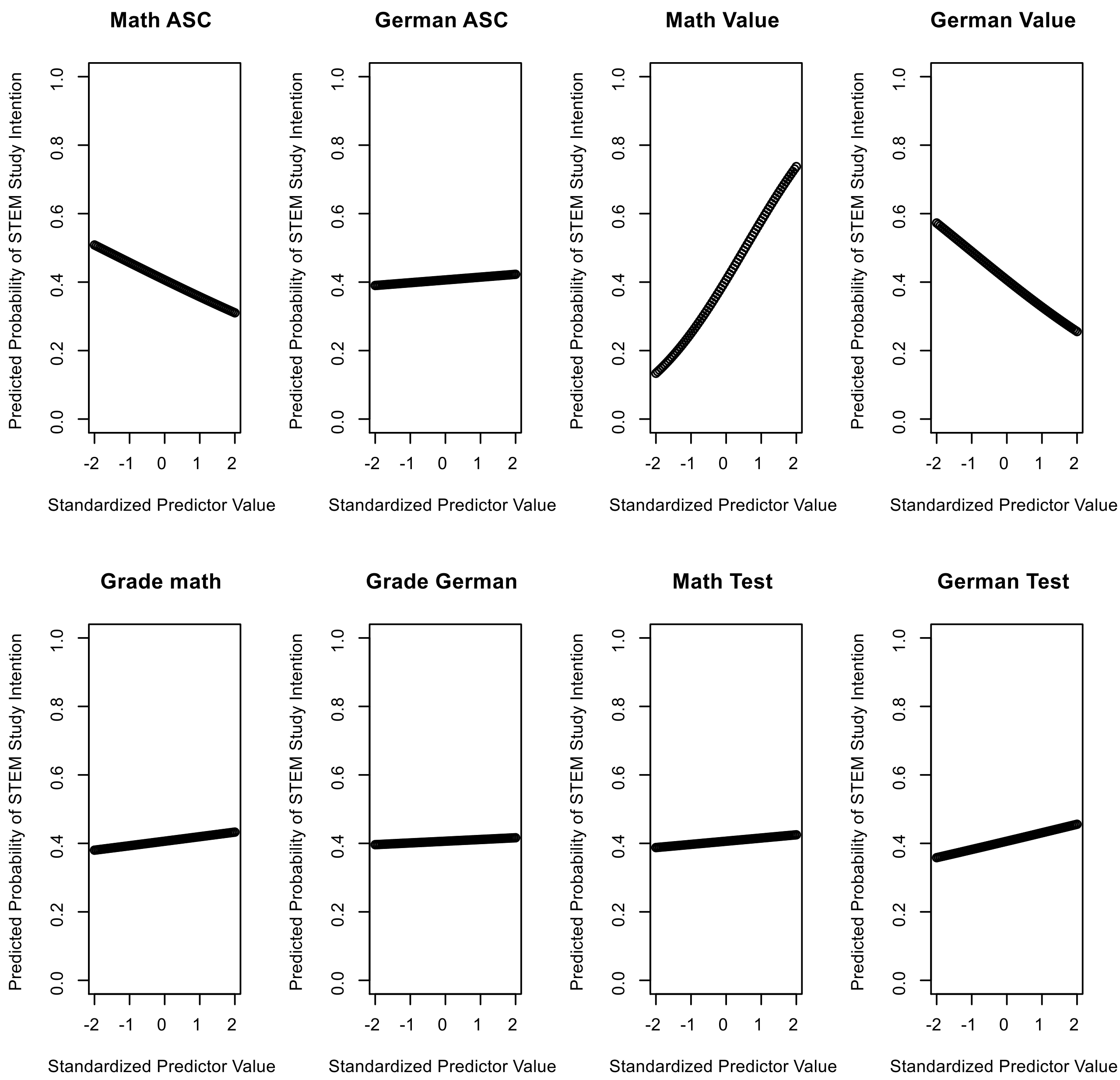

Figure 4. Predicted probabilities for the intention to study a STEM subject at university for different standardized values of a given predictor based on the full model (Table 3). All other continuous predictors were set to the sample mean and the dichotomous predictors to the reference category.