Extending Expectancy-Value Theory Predictions of Achievement and Aspirations in Science: Dimensional Comparison Processes and Expectancy-by-Value Interactions

J Guo, HW Marsh, PD Parker, AJS Morin, T Dicke

Learning and Instruction 49, 81-91

Abstract

Based on TIMSS data (18,047 Grade 8 students from the four OECD countries that collected data for multiple science domains), this study integrated dimensional comparison theory and expectancy-value theory and tested predictions about how self-concept and value are related to achievement and coursework aspirations across four science domains (physics, chemistry, earth science, and biology). First, strong support for social comparisons suggested that high achievement in a particular domain enhance students' motivation in the same domain, which in turn predicted domain-specific aspirations. Particularly, self-concept significantly interacted with value to predict aspirations. Second, in the processes underlying the formation of self-concept and intrinsic value, students tended to engage in negative dimensional comparisons between contrasting domains (physics vs. biology) but positive dimensional comparisons between assimilating domains (physics vs. chemistry). Similar dimensional comparison processes were evident for the effects of self-concept and intrinsic value on aspirations. The results generalized well across all countries.

Keywords: self-concept, expectancy-value, science subjects, coursework aspirations, latent interaction

Extending Expectancy-Value Theory Predictions of Achievement and Aspirations in Science: Dimensional Comparison Processes and Expectancy-by-Value Interactions

The issue of talented and capable students opting out of the STEM (i.e., science, technology, engineering, and mathematics) pipeline has been a topic of enduring interest in the science education community. Given that dropping out of science coursework at high school makes it very difficult to undertake STEM college majors and careers, growing attention in research on science motivation has focused on disentangling the relationship between students' motivational beliefs and achievement in science on one hand, and high-school science course taking, aspirations, and persistence on the other (e.g., Guo, Parker, Marsh, & Morin, 2015; Nagy et al., 2008; Parker et al., 2012).

These studies have demonstrated that motivation beliefs (e.g., academic self-concept and value beliefs) represent important determinants of achievement-related decisions in STEM subjects, net of individual's actual ability and achievement (Wang & Degol, 2013). However, much of this research has focused on motivational beliefs in general science, whereas science choices and aspirations are often measured in specific science domains. Indeed, the process of subject selection is inherently comparative. For example, let us consider the decision to major in physics at college. Students will be most likely to select this major only if they hold high confidence in their ability to do well in the course required by this major and place high value on majoring in physics by comparing the physics major to other majors including other science domains (see Eccles, 2009). Such intraindividual dimensional comparisons have been found to be useful for predicting academic choices. Nevertheless, existing research has focused almost exclusively on the dimensional comparison processes between math and verbal domains (e.g., Parker et al., 2012).

The aim of this study was to overcome the shortcomings of prior research, by testing the relations between academic achievement, motivational beliefs, and coursework aspirations taking into account several different science disciplines. In pursuing this overarching aim, we integrated and extended two major theoretical models of academic motivation (i.e., dimensional comparison theory [DCT], Möller & Marsh, 2013; expectancy-value theory [EVT], Eccles, 2009) in relation to four major science domains (physics, chemistry, biology, and earth science).

First, contrasting achievement and motivation, we tested how students' subject-specific self-concept and intrinsic and utility values in science were shaped by dimensional comparisons. Second, extending theoretical developments based on DCT, we explored how such dimensional comparison processes predicted coursework aspirations across different science domains. Third, extending recent developments based on EVT, we tested how academic self-concept interacted with value beliefs in predicting aspirations.

The present study drew on eight-grade students from the Trends in International Mathematics and Science Study (TIMSS 2007). TIMSS has been a major basis of international comparisons of countries in terms of educational motivation and achievement in the four major science domains. Thus, it presents an unprecedented opportunity for researchers to investigate students' motivational pathways to different STEM-related fields. This study was among the first to take advantage of the TIMSS data to address this substantive issue. In order to test the cross-national generalizability of our results, we rely on a convenience sample of all OECD countries who chose to conduct separate motivation assessments in physics, chemistry, biology and earth science, including the Czech Republic, Hungary, Slovenia, and Sweden (Olson, Martin, & Mullis, 2008). We note that the current approach, aiming to identify pan-human generalizations rather than country-specific idiosyncratic effects, is well-aligned with the approach typically taken in the study of similar educational phenomenon (e.g., the Internal-External frames of reference [I/E] model, the Big-Fish Little-Pond effect) using large international data sets (Marsh et al., 2014, 2015).

Focusing on motivational beliefs in general science or a single subject domain would result in a very limited perspective in explaining achievement-related behavior choices in STEM and may even be counterproductive in understanding coursework selection and aspirations in particular science disciplines (Eccles, 2009). By evaluating the influence of the intraindividual dimensional comparisons in relation to self-concept and value within science domains, this investigation may shed some light on how achievement and motivational beliefs might affect the decision students make to remain in or leave from the pathway toward different STEM-related fields.

1 Dimensional Comparison Processes

Academic self-concept, the self-evaluation of a student's ability in a given domain, has been assumed to be a multifaceted, hierarchical construct including a number of self-perceptions in different academic domains (Marsh, 2007). In order to evaluate their strengths and weaknesses, students compare and contrast their own performances across different school disciplines (Möller & Marsh, 2013). The I/E model were originally developed to explain the apparently paradoxical relations among domain-specific self-concepts and achievement: near zero-correlations between math and verbal self-concepts despite math and verbal achievement being moderately to strongly correlated (Marsh, 2007). The I/E model posits that students form their verbal and math self-concepts as a function of two underlying processes: social and dimensional comparison. Using an external frame of reference, students conduct social comparisons by comparing their self-perceived performance in a subject domain with that of their peers in the same school or classroom. For instance, if students have higher math achievement than do their classmates, their math self-concept is also likely to be higher. Thus, the social comparison processes lead to a positive prediction from achievement and self-concept within a subject domain. Employing a dimensional frame of reference, students conduct dimensional comparisons by comparing their performances in one particular subject domain against their performance in other subject domains. However, the dimensional comparison processes are ipsative, so that high levels of math ability should lead to lower verbal self-concept once the positive effect of verbal ability is controlled for.

Recently, the I/E model has been extended into DCT (Möller & Marsh, 2013) by incorporating a wider variety of subject domains. DCT postulates that academic self-concepts are formed by different dimensional comparisons. On the one hand, contrasting dimensional comparison processes predict that good performance in one domain leads to lower self-concept in other domains (i.e., contrast effects). On the other hand, assimilating dimensional comparison processes are characterized by good performance in one domain leading to higher self-concept in other domains (i.e., assimilation effects). Whether students engage in contrasting or assimilating dimensional comparisons is related to their beliefs as to whether two abilities are negatively or positively correlated (Möller et al., 2015). One of the critical assumptions of DCT is that perceived subject similarity corresponds to the verbal-mathematical continuum of core academic self-concept domains (Möller & Marsh, 2013). This assumption has been well supported in both empirical and experimental studies. For example, Haag and Götz (2012) demonstrated that subjects (far from each other on the continuum, e.g., math vs. German) with low self-concept correlations were perceived as rather dissimilar and that subjects (close to each other, e.g., math vs. physics) with high self-concept correlations are perceived as more similar. A recent empirical study (Helm et al., 2016) also confirmed this assumption and addressed that contrast effects were stronger when students focus on differences between two subject domains than when they focused on similarities. Thus, according to the verbal-mathematical continuum of academic self-concept, assimilation effects are assumed to occur between "near" domains, whereas contrast effects are assumed to occur between "far" domains.

In relation to science domains, physics and chemistry are assumed to be located closer to the math domain, whereas biology is assumed to be located in the middle of the continuum. More recently, Jansen et al. (2014) contrasted achievement and self-concept in physics, chemistry, and biology and found that associations of self-concept with achievement and grades were substantial in the same domains. For cross-subject relations, they revealed slightly negative contrast effects between biology and physics but assimilation effects between chemistry and physics (for similar results, also see Jansen et al., 2015). However, these two previous studies focus on German high school students, and the findings have yet to be replicated with other populations across different science curricula. Moreover, these studies have not included earth science and thus miss out on the opportunity to gain insight into dimensional comparison processes between four major science disciplines.

More recently, based on DCT, the Generalized I/E (GI/E) Model (Möller et al., 2015) has been developed by connecting dimensional comparison processes to broader cognitive, affective, and motivational consequences. Dimensional comparisons are assumed to serve as a critical source of information as to students' strength and weakness across different domains. These self-evaluations would help students to distinguish domains in which they can specialize, and for which they could develop particular interests, emotions, and preferences. Thus, dimensional comparisons are underlying mechanisms for the process of self-differentiation to serve motivational needs (Möller et al., 2015). In this regard, the GI/E model assumption has been tested with respect to emotions (Goetz, Frenzel, Hall, & Pekrun, 2008), intrinsic and utility values (Nagy et al., 2008; Schurtz et al., 2014), and perceptions of the learning environment (Arens & Möller, 2016). For example, Schurtz et al. (2014) found negative contrast effects from grades to intrinsic value between math and English, following the typical I/E pattern. However, the negative impact of dimensional comparisons on intrinsic value was totally mediated by self-concepts (also see Nagy et al., 2008).

However, these studies mainly drew on math and verbal domains. Arens and Möller (2016) argued that the scope of subject domains and outcome variables that are subject to dimension comparison processes should be even broader. There is to our knowledge no study examining the potential operation of dimensional comparisons in the formation of students' values in "near" domains on the continuum, such as between science subdisciplines.

2 Integrating Dimensional Comparison into EVT

Dimensional comparison processes posited in DCT and the GI/E model have been integrated into modern EVT, which has been widely used to explain students' academic choice behaviors (Eccles, 2009). EVT posits that a *relative* intraindividual's hierarchy of competence beliefs (e.g., academic self-concept) and task value are influenced as a function of previous achievement across subject domains (Eccles, 2009). More importantly, these relative motivational beliefs are postulated to play important roles to link between achievement and behavioral choices and aspirations in EVT. All such behaviors are also assumed to be associated with costs, as one choice often eliminates other options (an ipsative process), and thus trigger dimensional comparison of achievement and motivation (Eccles, 2009). Put simply, individual differences in relative self-concept and task value attached to a domain compared to other domains influence course enrollment (Nagy et al., 2008). In this regard, the individual hierarchy of self-concept and value across domains are not only the consequences of dimensional comparisons but also the antecedents of behavioral choices.

In this study, we focus on two of these components: intrinsic and utility values. Intrinsic value, referring to the extent to which the person gains enjoyment from performing an activity, has been found to be a stronger predictor of academic engagement, effort exertion, and coursework aspirations (e.g., Guo et al, 2015, 2016). Utility value refers to how useful a task is for facilitating an individual’s long-range goals and helping an individual obtain long-range external rewards. It has been found to more closed to educational and career aspirations, particularly during the post-high school transition (Wigfield, Tonks, & Klauda; 2016)

Although the notion of dimensional comparison processes has been well integrated into EVT, relatively little empirical work has applied such processes to predict achievement-related choices. Nagy et al. (2008; Parker et al., 2012) presented one of the few exceptions and provided an excellent example of these processes. By comparing their performance in math and English, high school students who had better performance in math tended to become more confident and interested in math but less in English. Subsequently, the positive motivation in math led these students to choose an advanced math course but opt out of an advanced English course. Again, these studies only focus on math and verbal domains, which leaves open the question as to whether students engage in assimilating dimension comparisons between similar domains (e.g., physics and chemistry) during the decision-making process. Thus, this study integrates EVT with new insights from DCT and draws on multiple, similar (science) domains to explore how dimensional comparison processes predict coursework aspirations.

3 Interaction Between Self-Concept and Task Values

In addition to having the first-order effects, competence beliefs and value beliefs are assumed to interact with each other in influencing achievement-related behaviors and choices in early EVT (Atkinson, 1957). The expectancy-by-value interaction suggests that if students do not have confidence in their abilities to succeed in a task, then even high value beliefs will not be sufficient to motivate students to pursue the task. However, this multiplicative relation, which was the central assumption of classic EVT, has not been widely studied in modern EVT. Nagengast et al. (2011) attributed this to weak statistical methodology in testing interaction effects and addressed that the expectancy-by-value interaction should be returned "to its rightful place at the heart of EVT" (p. 1064).

Recently empirical studies have successfully reintroduced examination of interaction effects between expectancy and value in predicting educational outcomes based on the newer approaches (e.g., the unconstrained approach; Nagengast et al., 2011). For example, Guo, Parker et al. (2015) found that the interactions between high school math self-concept and values significantly predicted math course selection, matriculation results, subsequent STEM major choices and entry into university. However, most of this research only considered a single domain (e.g., science), and the researchers did not test the domain specificity of the patterns of results across different science domains. As a consequence, their research did not explore the ipsative dimensional comparison processes; a matter that has been subsequently addressed with the extension to DCT and its integration into EVT.

4 The present investigation

Drawing on DCT and EVT, the present investigation aims to examine the distinctiveness of relationships between domain-specific achievement, motivational beliefs (self-concept, intrinsic value, and utility value), and coursework aspirations across four major science subjects (physics, chemistry, earth science, and biology). Importantly, we explore the roles of expectancy-by-value interactions with dimensional comparison processes in predicting aspirations. Hence, the present study is unique in that it takes multiple science disciplines into account and integrates DCT and EVT to provide a greater understanding of the motivational dynamics leading students to making academic choices within STEM-related fields. More specifically, self-concept, intrinsic value, and utility value along with achievements and aspirations in the four science domains are simultaneously included in the hypothesized model where all achievements are linked to the domain-specific motivational beliefs that in turn predict coursework aspirations (See Figure 1).

Hypotheses

4.1.1 Hypothesis 1: Relations between achievement and motivational beliefs

a. We predict matching paths from each of the four achievement domains to self-concept, intrinsic value, and utility value in the same domain to be significantly positive.

b. For physics, chemistry, and biology, according to the verbal-math continuum of self-concept (Marsh, 1990), we hypothesize non-matching paths (cross-paths) relating to “far” domains (e.g., physics achievement predicts biology self-concept) to be negative (contrast effects), whereas we hypothesize these cross-paths relating to “near” domains (e.g., physics achievement predicts chemistry self-concept) to be positive (assimilation effects).

However, earth science has not been positioned in this continuum. According to TIMSS, earth science is concerned with the study of earth and its place in the solar system and the universe, covering the fields of geology, astronomy, meteorology, hydrology and oceanography (Olson et al., 2008). For the four targeted countries, earth science is taught as a separate natural science subject along with biology, chemistry, and physics in Czech Republic and Hungary; although earth science is not taught separately, it is mainly represented in physics and chemistry and only included as a small part of the social sciences subject of geography in Slovenia and Sweden (Olson et al., 2008, see Table 1). Thus, we expect that earth science is more closely related to the mathematical side than the verbal side of the verbal-mathematical continuum. More precisely, we hypothesize that earth science is located in the middle of physics/chemistry and biology on the continuum, given that topics covered in the teaching and learning of earth science are largely intertwined with some concepts also covered in biology, physics, and chemistry. However, given the absence of empirical evidence for this, we still leave cross-paths involving earth science as a research question to be explored.

4.1.2 Hypothesis 2: Relations between motivational beliefs and coursework aspirations

a. We predict matching paths to be significantly positive from self-concept, intrinsic value, and utility value in each domain to coursework aspirations in the same domain, even after controlling for achievement. Based on previous research, in predicting coursework aspirations, we hypothesize matching path coefficients for intrinsic value to be stronger than those for utility value and self-concept.

b. We hypothesize cross-paths relating to "far" domain (e.g., biology self-concept predicts physics aspirations) to be negative, whereas these cross-paths relating to "near domain" (e.g., physics self-concept predicts chemistry aspirations) to be positive. Again, we leave the pattern of the predictions in relation to earth science as a research question.

c. Consistent with the recent re-introduction of expectancy-by-value interactions into EVT, we predict that latent interactions between self-concept and values (intrinsic value and utility value) predict aspirations beyond the first-order ("main") effects of these latent constructs.

4.1.3 Research question: Generalizability of results

Cross-cultural comparisons provide researchers with a heuristic basis to test the external validity and generalizability of their measures, theories, and models. Typically, there are two main approaches to cross-cultural comparisons: the etic and emic perspectives. The etic perspective refers to the cultural universals with an emphasis on cross-cultural similarities of theoretical predictions and replicability of results, whereas the emic perspective refers to phenomena specific to a particular culture with an emphasis the uniqueness of an individual case in its own terms. Marsh, Martin, & Hau (2006) addressed that one of the ongoing challenges in cross-cultural research in education is to untangle the potentially confounding effects of differences in participants representing different cultural groups and the appropriateness of psychological measure in different cultural settings. To the extent that a strong theoretical model generalizes well to heterogeneous samples drawn from a diverse set of countries, there is strong support for the external validity and the robustness of the interpretations based on the theory.

Indeed, there is a strong basis for the etic approach based on Möller et al. (2009) meta-analysis that found no significant differences across countries in support for the DCT predictions in relation to verbal and math self-concept. More recently, Marsh et al. (2015) provided a more critical evaluation of the cross-cultural generalizability of the I/E patterns. Their findings showed the strong support for the generalizability of the DCT predictions in relation to math and general science across 12 nations based on TIMSS2007 data. In this regard, one purpose of our study was to expand the scope of tests of the generalizability of the DCT predictions beyond previous studies that have been the primary basis of cross-cultural tests of the universality of support for DCT predictions.

Therefore, we leave as an open research question whether the hypothesized associations will generalize across the four OECD countries. Given that students were exposed in substantially different cultural and educational contexts across countries (See Table 1), it would provide a strong test of the external validity of our findings.

5 Method

Participants

Although standardized tests in four science domain-specific subjects (Physics, Chemistry, Earth Science, and Biology) are administered to eighth-grade students in all participating countries, TIMSS surveys in relation to the four subjects were only administered in countries teaching some or all of these subjects separately, rather than as a single, general subject (Olson et al., 2008). In TIMSS 2007 data, Czech Republic, Hungary, Slovenia, and Sweden are the *only* OECD countries in which students completed surveys in relation to these four science domains (Olson et al., 2008). Therefore, in the present study, the target population comprised eighth-grade students who participated in TIMSS 2007 from the four OECD countries described above. In total, we considered data from 18,047 students (51% boys) in 1,025 classes and 598 schools (see Appendix A).

Measure

Motivational factors. The measures of expectancy-value constructs were selected from the student-background questionnaire administered in TIMSS2007. All motivation items were coded on a four-Likert scale. For the present purposes, responses were reverse-scored, so that higher values represented more favorable responses and thus, higher levels of motivation.

A scale of students' Self-confidence in Learning Science that assesses how students think about their ability in specific domains was used to measure academic self-concept in TIMSS studies (Marsh et al., 2013, See Table 2). The students' Positive Affect Toward Science scale was applied to assess the affect experienced by students when participating in domain-related activities, in line with the notion of *intrinsic value* in the EVT. Likewise, the Students' Valuing Science scale was similar to utility value in the modern EVT, which assesses how well achievement in specific domains relates to current and future goals. These three latent constructs demonstrated satisfactory reliability across the four countries (see Appendix A).

Academic achievement. Participants' academic abilities of science are assessed though a range of questions in the four science subdomains. Two question formats were used in the TIMSS assessment – multiple-choice and written-response questions that involved a mixture of knowing, applying, and reasoning process (Olson et al., 2008).

Coursework aspirations. As there was only one item measuring students' achievement-related decisions in the TIMSS2007, following Marsh et al. (2013), this single item was used students' coursework aspirations in each subject area ("I would like to do more in Biology/Physics/Earth science/Chemistry in school."). The response scale ranged from 1, indicating that the participants "disagree a lot" to 4, indicating "agree a lot".

Data Analysis

In the present study, multi-group confirmatory factor analyses (CFAs) and structural equation models (SEMs) were conducted with Mplus 7.11 using the robust maximum likelihood estimator. The unconstrained approach (Nagengast et al., 2011) was utilized to model the latent interactions between self-concept and task value in predicting coursework aspirations. The classroom clustering and weighting variables were used to control for the clustering sample (see Appendix B and C). We used full information maximum likelihood (FIML) estimation to handle a relatively small amount of missing data on the remaining items (6.3% to 18.2% in Sweden and less than 2% for other countries).

Preliminary Analyses

Preliminary analyses described in details Appendix D demonstrated: (a) there was good support for the factor structures underlying the multiple domains of self-concept, intrinsic value, and utility value; (b) rigorous tests of factorial invariance showed that factor loadings, variances and covariances for motivational beliefs, achievement, and aspirations were invariant over the four OECD countries (Models MG1–MG4, See Table 3), and (c) there was good support for the convergent and discriminant validity of motivation beliefs in relation to achievement and aspirations, particularly for self-concept and intrinsic value, to a lesser extent, but also for utility value.

6 Results

Tests of Predictions Relating Achievement to Motivation Beliefs: Hypothesis 1

Matching paths. In this SEM model, we included one set of 16 (4 x 4; 1 matching path + 3 non-matching paths for each domain) paths from achievement in each science domain to each of the four self-concepts with two additional sets of 16 paths from achievement to each of the four intrinsic values and each of the four utility values (Models MG5–MG7, See Table 1). Of particular importance were the substantial path coefficients between paths from achievement to motivation constructs in matching domains compared to those in non-matching domains. To clarify these critical path coefficients, we computed summary statistics for matching paths, non-matching paths, and their difference (see Appendix E). As seen in Figures 2 based on Model MG7b where factor loadings and factor variances and covariances, and path coefficients were invariant across countries (see subsequent discussion), the matching paths from achievement to matching self-concept (Mean [*M*] = .19, SE = .01) and intrinsic value (*M* = .14, SE = .01) factors were positive across the four science disciplines. However, the matching paths for utility value were relatively small (*M* = .05, SE = .01).

Non-matching paths. The means across the 12 remaining non-matching path coefficients from achievement in each domain to non-matching motivational beliefs were substantially smaller than the corresponding matching coefficients (self-concept: Δ [mean of matching paths – mean of non-matching paths] *M* = .16, SE = .01; intrinsic value: Δ *M* = .16, SE = .01; utility value: Δ *M* = .07, SE = .01). More specifically, consistent with predictions from Hypothesis 1b, cross-paths between physics and biology were negative, whereas those between physics and chemistry were positive. We also found that cross-paths between chemistry and biology were slightly positive but significantly weaker than those between physics and chemistry (see Appendix E). Cross-paths between earth science and the other science domains were slightly positive or non-significant. It should be noted these patterns of results were only evident in relation to self-concept and intrinsic value.

Mediating role of self-concept. Following Nagy et al., (2008), we evaluated whether effects of achievements on task and intrinsic value could be explained by self-concept. In the mediation model (MG10) achievements in the four science domains predicted self-concepts, which in turn predicted intrinsic and utility values. In this model, the four domain-specific self-concepts and values along with achievements were also allowed to predict coursework aspirations. However, it is important to emphasize that the goodness of fit of this mediation model (MG10) is necessarily the same as the original non-mediation model (MG7b), as are the total (direct + indirect) effects of achievement; that is some of the effects interpreted as direct effects in MG7b are now are now interpreted as mediated effects in MG10, but the total effects are the same. The results revealed that all 32 direct paths from the four science achievements to each of the intrinsic and utility values were relatively small (from -.05 to .05; M = .01) in the mediation model. Subsequently, we evaluated a nested model where these 32 direct paths were constrained to be 0. There was a negligible decrease in model fit (Δ CFI = .002, Δ TLI = .001, Δ RMSEA = .001) when compared to the fully mediated model. Consistent with previous research these results can be interpreted to mean that the statistically significant total effects of achievements on intrinsic and task value are largely mediated by self-concept. Although the cross-sectional nature of our data dictate caution in the interpretation of the mediation model, the total effects in the mediation model (MG10) are the same as the direct effects in the original non-mediation model (MG7b). In this sense, the interpretations of total effects in the mediation model are the same as those of the direct effects in the non-mediation model.

Tests of Predictions Relating Motivational Beliefs to Aspirations: Hypothesis 2

Matching paths. We began with an evaluation of models without latent interactions. Consistent with predictions from Hypothesis 2a, matching paths from self-concept, intrinsic value and utility value in each domain to coursework aspirations, were substantially positive, controlling for achievement (see Figure 2). The mean across the four matching path coefficients for intrinsic value (*M* = .67, SE = .01) was substantially larger than that for self-concept (*M* = .10, SE = .01) and utility value (*M* = .06, SE = .01).

Non-matching paths. Non-matching paths (cross-path) from motivational beliefs to aspirations smaller than the corresponding matching paths (self-concept: D *M* = .09, SE = .02; intrinsic value: D *M* = .66, SE = .01; utility value: D *M* = .05, SE = .01). In line with predictions from Hypothesis 2b, cross-paths between physics and biology were significantly negative. Again, the pattern of results was found for self-concept and intrinsic value but not utility value. However, the majority of cross-paths involving self-concept, intrinsic value, and utility value were non-significant or slightly positive.

Latent interactions. We added two sets of domain-specific latent product variables to the Model MG7b: one based on product indicators for the self-concept and intrinsic value (MG8a-MG8b), and one based on those for self-concept and utility value (MG9a-MG9b). It should be noted that all path coefficients in the model with interactions are similar to those without interactions (see Appendix E). The mean of matching paths involving self-concept-by-intrinsic value and self-concept-by-utility value interactions were significantly positive (M = .12, SE = .01; M = .09, SE = .01, respectively). Given that the sizes of matching interaction path coefficients for different domains were similar, a simple-slopes plot was constructed, based on the mean of matching interaction path coefficients (see Figure 3). Tests of the simple slopes indicated that the slope for the effect of self-concept on aspirations for intrinsic and utility values of -1 SD below the mean (M = -.02, SE = .02, *p* = .211; M = -.01, SE = .01, *p* = .346, respectively) was non-significant. However, the slopes at average intrinsic and utility values became statistically significant (M = .10, SE = .01, *p* < .001; M = .08, SE = .01, *p* < .001, respectively), which was smaller than those for intrinsic and utility values of +1 SD above the mean (M = .22, SE = .02, *p* < .001; M = .17, SE = .02, *p* < .001, respectively). Figure 3 clearly shows the interactive relations of domain-specific self-concept and task value in predicting coursework aspirations: high self-concept only contributes to high aspirations when intrinsic and utility values are moderately elevated. However, when either utility value or intrinsic value are low, the contribution of self-concept in the prediction of aspirations is absent, which implies that high self-concept cannot compensate for low value (and vice versa). Supplemental analyses suggest that both types of domain-specific interactions (self-concept-by-intrinsic value and self-concept-by-utility value) make similar contributions to the prediction of aspirations when both product variables are considered simultaneously (see Appendix G).

Tests of Predictive Relations Over Countries

In order to test the generalizability of our results, we estimated a series of multiple-group SEMs testing whether path coefficients were invariant across the four countries (Models MG5–MG9b, see Table 3). We conducted pair comparisons for the models where the same measurement invariance was imposed (i.e., factor loadings, factor variances, and factor covariances) and the only difference was whether or not structural coefficients were freely estimated (e.g., MG7a vs. MG7b). Although the imposition of the additional constraints on structural coefficients resulted in some decreases in model fit, these decreases were negligible, and all models provided a satisfactory level of fit to the data. To more directly compare the similarity of country-specific path coefficients, we also calculated a profile similarity index (PSI). The PSI is an estimate of the correlations between path coefficients obtained from different countries. For all path coefficients based on Model 7a, the PSI indicated the very high level of similarity across the four countries (range from to .861 to .957, see Appendix H for country-specific path coefficients). Thus, there was strong support for the invariance of path coefficients over the four countries.

7 Discussion

In this study we adopted a multidimensional perspective on self-concept and intrinsic and utility values in science domains, and examined associations among achievement, motivational beliefs and coursework aspirations. Our findings suggest that outcomes in any one domain depend not only on accomplishments, self-concept beliefs, and value perceptions in that domain, but also on how these constructs compare to those in other, contrasting domains.

The Relations between Achievement and Motivational beliefs

Our findings supported DCT to confirm that students receive information from two main sources to form their self-concept: (a) they engage in social comparison with others as a way to judge their own abilities as evident by strong domain-specific relations between achievement and motivational beliefs; (b) students systematically evaluate their abilities by comparing difference subject domains (dimensional comparison processes). More importantly, our findings supported the crucial assumption of DCT that student tend to make both assimilating or contrasting dimensional comparisons, which is related to perceived subject similarity. Specifically, students are likely to engage in contrasting dimensional comparison between physics and biology which are separated by the greatest distance on the continuum of academic self-concepts (relative to other science domains). However, most previous support for such contrasting comparison is based on studies of math and verbal domains that are at opposite ends of the academic self-concept continuum. Simultaneously, students are likely to engage in assimilating dimensional comparison between physics and chemistry. This indicates that students apparently perceive physics and chemistry to be similar and complementary subjects, such that skills acquired in one subject will help success in the other subject, and achievement feedback in one subject may provide an additional source of positive information to help evaluate abilities in the other subject.

The assimilating dimensional comparisons are also evident between chemistry and biology, but they are significantly smaller than those between physics and chemistry. This result is in line with the verbal-math continuum, suggesting that chemistry and physics would be perceived as more similar to each other than chemistry and biology. However, these assimilation effects are not contradictory to the contrasting dimensional comparisons between physics and biology. Students who have high ability in chemistry tend to have high self-concept in both physics and biology, while highly able students in biology tend to have low self-concept in physics (and vice versa) given the same ability in chemistry.

With respect to earth science, the contrasting dimensional comparison processes apparently were not triggered in relation to other science domains. Instead, students are likely to engage in assimilating dimensional comparisons between earth science and other science domains in similar size. This indicated that students perceive earth science to be relatively similar to other science domains, implying that earth science would be located between physics/chemistry and biology in the verbal-math continuum. Note that this study is among the first to incorporate earth science and explore perceived similarity in relation to other domains. Thus, the results provide new theoretical and substantive insights into I/E model and DCT.

By integrating DCT into EVT, the results suggested that the two main sources involving achievement/ability comparison also significantly influence the development of students' intrinsic value. This finding suggests that when students perceive school subjects to be similar (e.g., physics and biology), intrinsic motivation in one is likely to generalize to the other, whereas when they perceive those subjects to be distinct (physics vs. biology), liking of one subject domain tends to wane if students have high achievement in the other domain. However, the pattern of results for utility value was somewhat weaker than that for intrinsic value. A theoretical reason may be that utility value is more related to an individual's personal and collective identities, whereas intrinsic value is more related to performance-based experiences. The formation of utility value may rely on other sources, such as cultural and parent subjective norms (Wigfield et al., 2016). Put simply, parents who value math are likely to communicate these beliefs to children as a way for children to understand that math is important and useful, which can influence students' own valuing of math. Another reason might be that students are not able to distinguish utility value in different science subjects at Grade 9, as evident by the low degree of domain specificity of utility value (see Appendix D). The domain specificity of the construct is one of the bases underlying dimensional comparison mechanisms. The pattern of relations between the motivational factors and achievement is largely a function of the domain-specific nature of this factor. Previous research has suggested that a lower degree of domain specificity for the motivational constructs is associated with weaker support for the I/E model (Marsh et al., 2013).

The follow-up analyses indicated that the influence of dimensional comparison on the development of students' task values was largely mediated by self-concept. This result reinforces the central role of self-concept in terms of DCT, but the cross-sectional nature of our data dictate caution in the interpretation of the results. Hence, pursuit of this issue is a potentially important direction for further research based on longitudinal data where stronger tests of the causal ordering implicit in the mediation model are possible.

The Relations between Motivational beliefs and Aspirations

Consistent with a prior prediction, this study found that self-concept and intrinsic and utility values are positively associated with coursework aspirations in the same domain. Importantly, this study is among the first to test latent expectancy-by-value interactions for multiple science domains within the same model. There is strong evidence of the high domain specificity of interactive relations in predicting coursework aspirations. The interactive roles of self-concept and value suggest that students with both high science self-concept and task value are more likely to aspire to engage in science. However, students with high self-concept are unlikely to desire to pursue science in the future, if they ascribe a low level of intrinsic value to science. Similarly, students who value math are also unlikely to desire to enter a scientific career, if their science self-concept remains low. Therefore, this study provides strong support for the theoretical claim that self-concept and value interact in predicting achievement-related outcomes.

Dimensional comparison processes involving self-concept and value. This study extends prior research by integrating EVT and DCT and exploring predictions from motivational beliefs to educational aspirations. Contrasting dimensional comparison between physics and biology is evident for self-concept and intrinsic value. This means that for example, students who have high self-concept and interest in physics but even higher self-concept and interest in biology are likely to have lower aspirations in physics compared to students who have the same level of self-concept and interest in physics but lower self-concept and interest in biology. Thus, aspirations in one science domain depend not only on abilities, self-concept, and intrinsic value in that domain, but also on relative abilities and motivation in other science domains. These findings shed further light on the important roles played by dimensional comparison processes in shaping academic pathways to different STEM fields, and underline the importance of differentiating motivational beliefs across science domains.

However, it should be noted that all cross-paths (dimensional comparisons) between achievement, motivational beliefs and coursework aspirations were relatively weak, particularly for the assimilation effects. These results are consistent with recent self-concept research on science domains (Jansen et al., 2014). This may be because the four science subjects considered here are all relatively similar, compared to the more obviously contrasted academic continuum, ranging from relatively pure verbal subjects to relatively pure mathematical subjects (Marsh, 1990). Nevertheless, mathematics and verbal skills are posited as the endpoints of the academic continuum were not considered in this study.

Generalizability of the results

How science subjects are taught in a given learning environment varies as a function of the country, state or school system, and this is particularly true for earth science. Based on TIMSS data, the present study evaluated the only four available participating OECD countries that assessed students' motivational beliefs in the four science domains. Despite substantial variations in the sociocultural and educational background (see Table 1), the pattern of results is invariant across the four countries, supporting the external validity of our findings. Particularly, the cross-cultural support for the generalizability of DCT predictions reflects a broader tendency of students to engage in both assimilating and contrasting dimensional comparisons to develop their self-concept, task values, and aspirations in science. Our results indicate that even students who excel at science, particularly in physics, might have high self-concept and values in both physics and chemistry; however, they might have an average or below average self-concept and values in biology, which may seem paradoxical in relation to their good achievement (better compared to other students but not compared to their own performance in other science domains). This indication is inconsistent with teachers' perception of formation of students' motivation. Previous studies have shown that teachers tend to believe that students who are capable in one academic domain tend to be seen as having high self-concept and values in all domains, while students who are not capable in one area are seen as having low self-concept and values in all domains (Marsh, 2007). Thus, the generalized pattern has fundamental implications about the way teachers give feedback to students in different academic domains (see further discussion below). However, it should be noted that generalizability over only four countries is not sufficient to provide strong support for cross-cultural generalizability. But it sheds light on the generalized motivational mechanism by the integration of EVT and DCT and offers a good starting point for further research.

Implications for Instructional Practices

With respect to instructional practices, the high domain specificity of self-concept and intrinsic value suggests that interventions targeting general academic, or even a general science self-concept and intrinsic value, may not be beneficial in promoting students' motivation in STEM areas. Rather, interventions targeting a specific academic self-concept domain, with the integration of self-enhancement (self-concept enhances ability) and skill development (ability improves self-concept) strategies, have been shown to be much more effective than those solely targeting a global or skill-based self-concept (Marsh, 2007). Interventions designed to increase students' perceptions of the relevance of academic subjects to their lives through teachers and parents have been found to be effective in triggering students' interest and to promote academic performance in STEM topics (Harackiewicz, Rozek, Hulleman, & Hyde, 2012).

Furthermore, we recommend that teachers should be aware not only of the dimensional comparison processes underlying the formation of students' self-concept and intrinsic value, but also of the comparison processes leading students to different levels of coursework engagement and aspirations. Particularly the contrasting comparisons between physics and biology may help to explain the gender imbalance in STEM careers with girls' underrepresentation in physics-related careers but slightly overrepresentation in biology-related careers (Wang & Degol, 2013). Understanding such comparison processes would also help teachers provide effective feedback to students. In particular, attributional feedback, goal feedback, and contingent praise, as forms of constructive feedback, have been identified as effective methods of boosting self-concept (O'Mara et al., 2006). Thus, our findings would help educational policymakers and practitioners to improve retention in STEM classes through high school and could be particularly beneficial in supporting girls to pursue physics-related careers.

In addition, the distinctiveness of the interactive relations between self-concept and value beliefs across science domains, suggests that interventions targeting the promotion of aspirations to STEM majors should seek to enhance both domain-specific self-concept and task value. This suggests that multicomponent interventions (Gläser-Zikuda, Fuß, Laukenmann, Metz & Randler, 2005) might be more effective in promoting students' motivation than those based on self-concept and value interventions considered separately.

Limitations and Directions for Future Research

Several limitations to this study, and some caveats, must be noted. First, in the present cross-sectional study, the issue of the temporal or causal ordering among achievement, motivational beliefs and aspirations could not be addressed on the basis of a single measurement point. Thus, a longitudinal replication would enable us to draw stronger conclusions about the directional influences of self-concept and value and the importance of their interactions.

Second, as our study is limited to the four OECD countries where science is taught as separate subjects, it is also important to replicate the results in settings where students are taught science as an interdisciplinary, unified subject. Relatedly, the strengths of DCT and EVT predictions in science is likely to vary as a function of age, as the further students go in school the more differentiated the coursework is likely to be. This is particularly the case as students move into higher education. Thus, research across different international samples covering multiple age groups, school subjects and schooling systems would be useful, to clarify the generalizability of our findings.

Third, our findings support the assumption that students make assimilating or contrasting dimensional comparisons are related to how they perceive similarity of two or more domains These results are also consistent with the experimental studies suggesting that lower perceived subject similarity would lead to stronger self-concept differences than did higher perceived similarity (Helm et al., 2016). However, such experimental study focusing on the effect of perceived subject similarity and dimensional comparisons on task value has been sparse. It, therefore, would be another avenue for future research.

Finally, given that the present investigation only focuses on two out of four major value components and single-item coursework aspirations, future research should consider psychometrically stronger, multi-item measures of the four value components and coursework aspirations.

References

Arens, A.K., & Möller, J. (2016). Dimensional comparisons in students' perceptions of the learning environment. *Learning and Instruction*, 42, 22-30

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological review, 64*, 359–372.

Eccles, J. S. (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*, 78–89.

Gläser-Zikuda, M., Fuß, S., Laukenmann, M., Metz, K., & Randler, C. (2005). Promoting students' emotions and achievement—Instructional design and evaluation of the ECOLE-approach. *Learning and Instruction*, 15, 481–495.

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.

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

Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation. *Journal of Educational Psychology*, *92*, 331-341.

Haag, L., & Götz, T. (2012). Mathe ist schwierig und Deutsch aktuell: Vergleichende Studie zur Charakterisierung von Schulfächern aus Schülersicht [Math is difficult and German up-to-date: Comparative study of characterizations of school subjects from students' points-of-view]. *Psychologie in Erziehung und Unterricht*, 59, 32–46.

Harackiewicz, J. M., Rozek, C. S., Hulleman, C. S., & Hyde, J. S. (2012). Helping Parents to Motivate Adolescents in Mathematics and Science: An Experimental Test of a Utility-Value Intervention. *Psychological Science*, 23, 899–906.

Helm, F., Mueller-Kalthoff, H., Nagy, N., & Mo ller, J. (2016). Dimensional Comparison Theory: Perceived Subject Similarity Impacts on Students Self-Concepts. *AERA Open*, 2(2), 1–9.

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.

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*.

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

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

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., & Abduljabbar, A. S. (2015). Dimensional Comparison Theory: Paradoxical relations between self-beliefs and achievements in multiple domains. *Learning and Instruction*, *35*, 16–32.

Marsh, H. W., Martin, A. J., & Hau, K-T. (2006). A multiple method perspective on self-concept research in educational psychology: A construct validity approach. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 441–456). Washington, DC: APA

Möller, J., Helm, F., Müller-Kalthoff, H., Nagy, N., & Marsh, H. W. (2015). Dimensional comparisons and their consequences for self-concept, motivation, and emotion. *International Encyclopedia of the Social & Behavioral Sciences*, 26, 430–436.

Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, *120*, 544–60.

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

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

Nagy, G., Garrett, J., Trautwein, U., Cortina, K., 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: APA.

Olson, J. F., Martin, M. O., & Mullis, I. V. S. (2008). *TIMSS 2007 technical report.* Chestnut Hill, MA: TIMSS & PIRLS.

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*, 1629–42.

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.

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*, 304–340.

Wigfield, A., Tonks, S., & Klauda, S. L. (2016). Expectancy-value theory. In K. R. Wentzel & D. B. Miele (Eds.), Handbook of motivation at school (pp. 55–74). New York: Routledge.

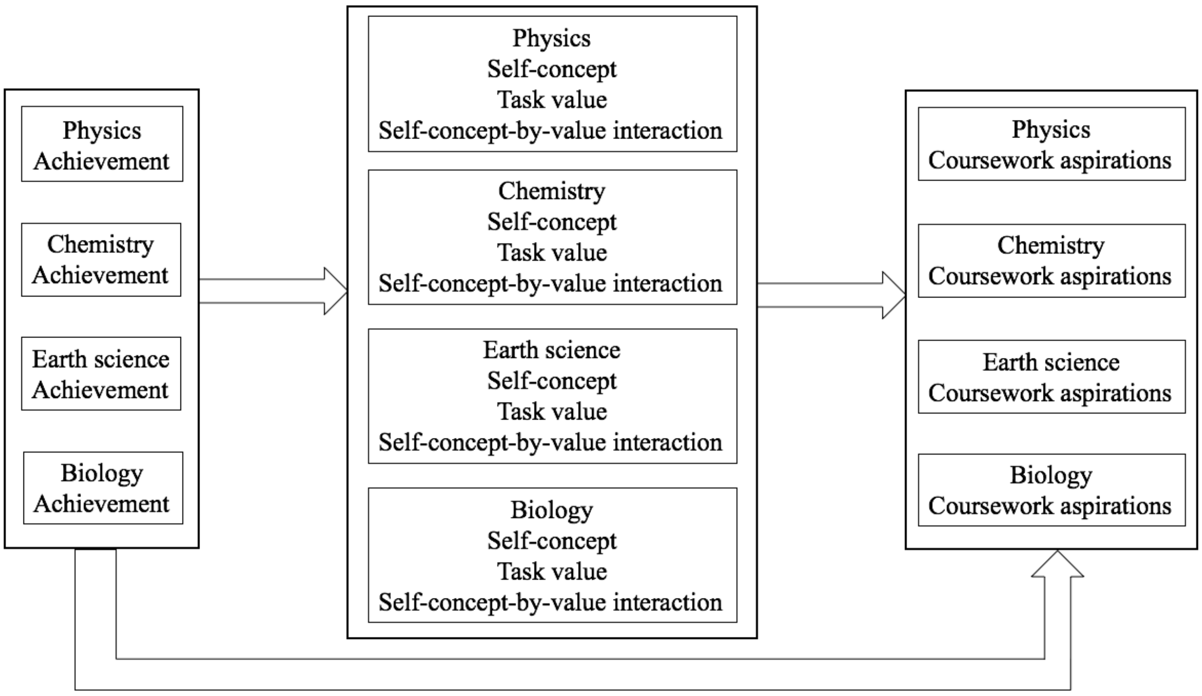

Figure 1. Hypothesized Model.

Note. Self-concept, intrinsic value, and utility value along with achievements and aspirations in the four science domains are simultaneously included in the hypothesized model. The model depicted is a “full-forward” structual equation model that is saturated, in the sense that the four achievements are allowed to predicted domain-specific motivational beliefs and all motivational beliefs along with the achievements are allowed to predict the four coursework aspirations.

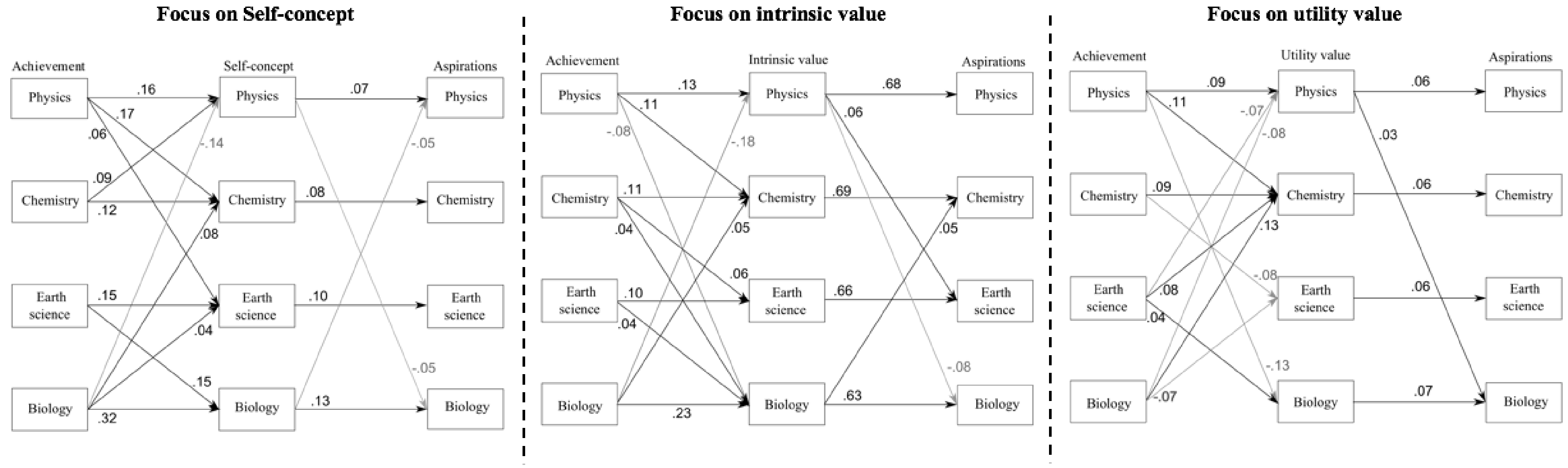

Figure 2. Structural path model of the relations between achievement, motivational beliefs (self-concept, intrinsic value, utility value), and coursework aspirations across the four science domains.

Note. The path coefficients reported in the figure are based on the hypothesized model excluding the self-concept-by-value interactions (Model MG7b), where self-concept, intrinsic value, and utility value along with achievements and aspirations in the four science domains are simultaneously included. It should be noted that all path coefficients in the model with interactions are similar to those without interactions (see Appendix E). Only statistically significant regression paths ($p < .05$) are presented. Negative, significant paths are shaded in gray.

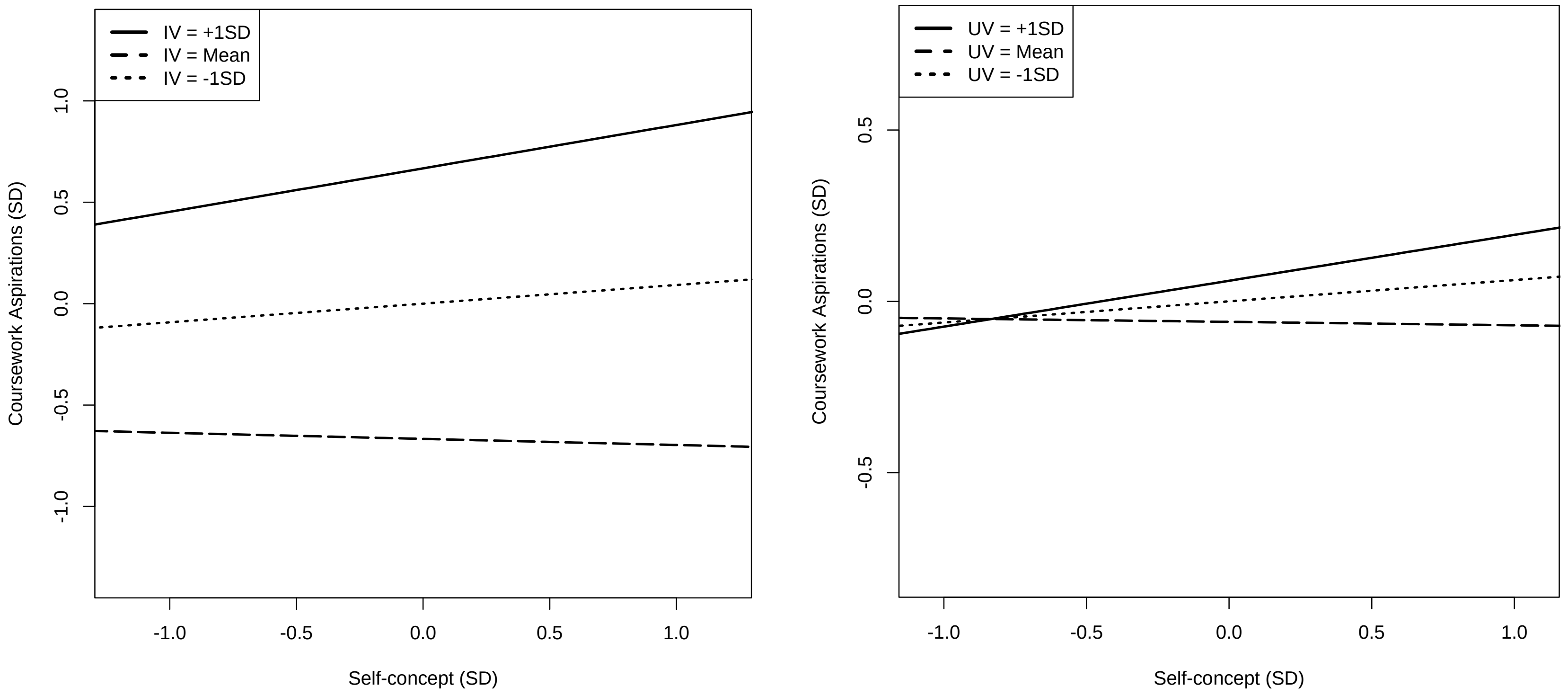

Figure 3. Simple-slopes depicting the effects of latent interactions (self-concept by intrinsic value and self-concept by utility value) on coursework aspirations. *Note.* IV = intrinsic value; UV = utility value.

Table 1.

Country Characteristics and Science Curriculum for the Four Targeted Countries.

Country characteristics					
	HDI	GNI per Capita	Public Expenditure on Education (% of GDP)	Net Enrollment Ratio in Education (secondary)	Life Expectancy at Birth (Year)
Czech Republic	.891	12790	4	95	77
Hungary	.874	10870	5	90	73
Slovenia	.917	18660	6	91	78
Sweden	.956	43530	7	99	81

Science Curriculum					
	National Science Curriculum	Science Curriculum Structure (Grade)	One curriculum for All students	National or Regional Assessments	Instruction for Earth Science in Lower Secondary Grade
Czech Republic	Yes	6-9	Yes	NO	separate subject
Hungary	Yes	5-6,7-8	Yes	Yes	separate subject
Slovenia	Yes	7-9	Yes[1]	NO	mainly in physics.
Sweden	Yes	6-9	Yes	Yes	mainly in physics and chemistry

Note. HDI = Human Development Index; GNI = Gross National Income; [1] One curriculum for all students, but different groups of students have different difficulty levels.

Table 2

A Priori Factor Structure Relating The TIMSS Motivation Items to Latent Factors

Items	Item wording	Factor loading			
		Physics	Chemistry	Earth science	Biology
Self-concept					
SCP1	I usually do well in Physics/Chemistry/Earth science/Biology	.81	.82	.81	.79
SCP2	I learn things quickly in Physics/Chemistry/Earth science/Biology	.82	.82	.80	.79
SCN1	Physics/Chemistry/Earth science/Biology is more difficult for me	.50	.52	.52	.54
SCN2	Physics/Chemistry/Earth science/Biology is not one of my strengths	.60	.63	.64	.65
Intrinsic value					
IVP1	I enjoy learning Physics/Chemistry/Earth science/Biology	.87	.88	.86	.84
IVP2	I like Physics/Chemistry/Earth science/Biology	.66	.70	.73	.74
IVN1	Physics/Chemistry/Earth science/Biology is boring	.88	.89	.90	.90
Utility value					
UVP1	I think learning Physics/Chemistry/Earth science/Biology will help me in my daily life	.66	.66	.53	.57
UVP2	I need Physics/Chemistry/Earth science/Biology to learn other school subjects	.66	.63	.59	.56
UVP3	I need to do well in Physics/Chemistry/Earth science/Biology to get into the university of my choice	.83	.83	.80	.81
UVP4	I need to do well in Physics/Chemistry/Earth science/Biology to get the job I want	.84	.84	.81	.81

Note. These results are based on Model MG4. Factor loadings were constrained to be equal across the four countries. The wording of the items was rigorously parallel for the corresponding science domain-specific scales. P = physics; C = chemistry; E = earth science; B = biology; SCP = self-concept (positive); SCN = self-concept negative; IVP = intrinsic value (positive); IVN = intrinsic value (negative); UVP = utility value (positive).

Table 3

Model Fit Statistics for the Multi-group CFA and SEM Models Used in the Present Study

Model	Description	χ^2	*df*	CFI	TLI	RMSEA
CFA						
MG1	SC + IV + UV + ACH + ASP, Correlated Uniqueness + configural	18038	3912	.964	.951	.028
MG2	MG1 + FL invariance	19416	4008	.961	.948	.029
MG3	MG2 + FV invariance	20138	4068	.959	.947	.030
MG4	MG3 + CV invariance	23961	4638	.951	.944	.030
SEM						
MG5a	MG2 + freely estimate PC	19416	4008	.961	.948	.029
MG5b	MG2 + PC invariance	21088	4344	.958	.948	.029
MG6a	MG3 + freely estimate PC	20146	4068	.959	.947	.030
MG6b	MG3 + PC invariance	21790	4404	.956	.947	.030
MG7a	MG4+ freely estimate PC	22327	4302	.954	.944	.030
MG7b	MG4+ PC invariance	23961	4638	.951	.944	.030
MG8a	MG6b + freely estimate FL, FV, CV, PC relating to SCxIV	29880	6934	.945	.936	.027
MG8b	MG6b + FL, FV, CV, PC invariance relating to SCxIV	31512	7228	.942	.936	.027
MG9a	MG6b + freely estimate FL, FV, CV, PC relating to SCxUV	25644	7740	.955	.947	.023
MG9b	MG6b + FL, FV, CV, PC invariance relating to SCxUV	28647	8184	.948	.942	.024
MG10	Mediation model	23961	4638	.951	.944	.030

Note. PC = path coefficients; SC = self-concept; IV= intrinsic value; UV = utility value; ASP = coursework aspirations; SCxIV = the product term of self-concept by intrinsic value interaction; SCxUV = the product term of self-concept by utility value interaction; FL = factor loading; FV = factor variances; CV = factor covariances.

Online Supplemental Materials for:

Extending Expectancy-Value Theory Predictions of Achievement and Aspirations in Science: Dimensional Comparison Processes and Expectancy-by-Value Interactions

Note: These appendices are intended to appear only on a website hot-linked to the article, and are not intended for the printed article.

Table of Contents

External Appendix A:

Sample Size and Reliabilities of The TIMSS Motivation Constructs

Table A1

Sample Size and Reliabilities of The TIMSS Motivation Constructs Based on Four Science Domains for Four OECD Countries

	Sample Size				Reliability Estimates												
Country	Student	Class	School	%boys	PSC	CSC	ESC	BSC	PIV	CIV	EIV	BIV	PUV	CUV	EUV	BUV	Mean
Czech Republic	4842	212	147	52%	.84	.85	.83	.82	.84	.86	.86	.85	.84	.86	.86	.85	.83
Hungary	4108	246	144	50%	.83	.82	.83	.82	.84	.85	.87	.88	.84	.85	.87	.88	.83
Slovenia	4029	260	148	50%	.77	.80	.79	.80	.83	.87	.87	.87	.83	.87	.87	.87	.82
Sweden	5068	307	159	52%	.79	.79	.79	.79	.87	.88	.88	.88	.87	.88	.88	.88	.84
Total	18047	1025	598	51%	.81	.82	.81	.81	.85	.87	.87	.87	.84	.84	.80	.79	.83

Note. The column headed Mean is the mean of the eight reliability estimates. The wording of the items was rigorously parallel for the corresponding science domain-specific scales. Reliability estimates are Cronbach’s alpha estimates. P = physics; C = chemistry; E = earth science; B = biology; SC = self-concept; IV = intrinsic value; UV = utility value.

Say

External Appendix B:

Unconstrained Approach, Standardization, and Annotated Mplus Syntax

Unconstrained Approach

In comparison to the traditional constrained approach (e.g., Jöreskog & Yang, 1996;) and the partially constrained approach (Wall & Amemiya, 2001), the unconstrained approach is relatively simple to implement in that most of the complicated constraints required in the original Kenny and Judd's approach are relaxed (Marsh, Wen & Hau, 2004). The unconstrained approach has shown good performance as the constrained approach when the underlying assumptions of the constrained approach are met in the simulation study, and much better performance when these assumptions are not met – which is generally the case (Marsh et al., 2004).

The SEM with two latent predictors and their interacting latent variable is typically specified as:

$$h = g_1 x_1 + g_2 x_2 + g_3 x_1 x_2 + z. \qquad (1)$$

where g_1, g_2 and g_3 are the partial regression coefficients of the latent predictor variables and their cross-product and z is the structural model residual. The latent predictors x_1 and x_2 as well as the latent outcome variable h are each inferred from at least two indicator as specified in the corresponding measurement models. x_1, x_2 and x_3 are allowed to be correlated with each other, but each is uncorrelated with measurement errors and the residual term z .

$$x_{ij} = l_{xi} x_i + d_{ij}, y_k = l_{yk} h + e_k, \qquad (2)$$

where x_{ij} is the jth indicator of the ith latent predictor variable x_i, l_{xi} is the corresponding factor loading and d_{ij} is the corresponding residual, y_k is the kth indicator of the latent outcome variable h, l_{yk} is the corresponding factor loading, and e_k is the corresponding residual.

Product-indicator approaches such as the unconstrained approach identify the latent cross-product x_1 x_2 by products of indicators of the latent predictor variables, according to the following measurement model

$$x_{1i} x_{2l} = l_{1i2l} x_1 x_2 + d_{1i2l}, \qquad (3)$$

where x_{1i} is the ith indictor of x_i and x_{2l} is the lth indicator of x_2, l_{1i2l} is the corresponding factor loading on the latent product variable and d_{1i2l} is the corresponding residual. The critical problem with the indicator approach is how to form the product indicator. All indicators of the latent variables are centred before the product indicators are compute (Marsh et al., 2004). According to the guiding principles proposed in Marsh et al. (2004), (a) all the multiple indicators of both latent predictors are needed to use, and (b) the same indicator should not be re-used the same indicator in forming the indicators for the latent product variable (also see Marsh, Hau, Wen, Nagengast, & Morin, 2013). Hence, each indicator in x_1 and x_2 should be used only once in the formation of the product indicators. In this study, product indicators are formed based on the reliabilities of the indicators of x_1 and x_2 (i.e., the best item in x_1 with the best item in x_2, for detailed discussion about construction of product indicators see Marsh et al., 2004; Marsh, Hau et al., 2013).

Standardization

First, all individual indicators (rating item, test scores and coursework aspirations) were standardized in relation to the total sample mean and standard deviation, as recommended by Marsh and his colleagues (Marsh et al., 2004; Marsh, Hau et al., 2013). Second, for total group analysis, product indicators for the latent interactions were formed using the match-pair strategy according to Marsh, Hau et al.(2013)'s guiding principles (also see Marsh et al., 2004 for more discuss about the product predictors selection procedure). For the multi-group analysis, the standardized indicators were centered (but not re-standardized the product term) within country-specific mean before forming the product indictors for the latent interaction variable (Nagengast et al., 2011). In order to obtain appropriate standardized results (Wen, Marsh, & Hau, 2010), for total group analysis all latent factors (including the latent product variables) were then standardize in relation to the total sample. For multi-group analyses, the critical assumption of test whether the pattern of results generalizes across groups is invariance of factor structure. To provide parameter estimates standardized to a common metric over the multiple groups, factor loadings and factor variances are needed to be invariant across the four countries. More specifically, we conducted a preliminary CFA model in which factor loadings and factor variances were constrained to be invariant over the multiple groups, and the metric was identified by fixing the factor variances of constructs to be 1.0 across the four groups, instead of fixing the first factor loading to 1.0. In subsequent SEMs these standardized factor loadings were used to define the latent factors, fixing the first factor loading for each factor to the value obtained in the CFA, in which the factor variances were fixed to be 1.0. In this way, all parameter estimates were estimated in relation to a standardized metric that was common across the four countries, providing appropriate standardized results (see Wen et al., 2010 for more details; also see below for the Mplus syntax). As showed in the main text, we also conducted a series of invariance tests with respect to factor covarances and path coefficient for multi-group measurement and structural models As the assumption of invariance was tenable, all results reported in this study were based on multi-group SEM with factor loading, path coefficients and factor variances and covariance invariances

The Annotated Mplus Syntax for Model

```
        DATA: FILE = "Multi_Sci_ExT_fix_MG.csv";
        VARIABLE:
        NAMES = IDCNTRY IDSCHOOL
        !domain-specific self-concept; P = physics; C = chemistry; E = earth science; B = biology;
        PSCP1 PSCP2 PSCN1 PSCN2 CSCP1 CSCP2 CSCN1 CSCN2
        ESCP1 ESCP2 ESCN1 ESCN2 BSCP1 BSCP2 BSCN1 BSCN2
        !domain-specific intrinsic value;
        PIVP1 PIVN1 PIVP CIVP1 CIVN1 CIVP2 EIVP1 EIVN1 EIVP2 BIVP1 BIVN1 BIVP2
        !domain-specific utility value;
        PVAL1 PVAL2 PVAL3 PVAL4 CVAL1 CVAL2 CVAL3 CVAL4
        EVAL1 EVAL2 EVAL3 EVAL4 BVAL1 BVAL2 BVAL3 BVAL4
        !domain-specific self-concept-by- intrinsic-value product terms
        ! three product terms for
        Pmsciv1 Pmsciv2 Pmsciv3 Cmsciv1 Cmsciv2 Cmsciv3
        Emsciv1 Emsciv2 Emsciv3 Bmsciv1 Bmsciv2 Bmsciv3
        !domain-specific achievement
        PACH CACH EACH BACH
        !domain-specific coursework aspirations
        PCW CCW ECW BCW
        HOUWGT CONTSCHL;
        MISSING=.;
        CLUSTER= CONTSCHL;
        ! cluster variable is the classroom;
        !We note that the classroom is the critical clustering variable for TIMSS data because class was the !sampling
unit used in the TIMSS sampling design, which was based on sampling all stu dents within intact !classes; most
schools are represented by a single class, and a given class might not be representative of the !school from which it
came.

        WEIGHT = HOUWGT;
        ! HOUWGT is the weighting variable in the TIMSS database; incorporates six components; ! three have to
!do with sampling of the school, class and student, and adjustment factors ! associated with non-!participation at the
level of the school, class and student.
```

```
  grouping is IDCNTRY (203=CZE 348=HUN 705=SVN 752=SWE); ! identify the four OECD countries

  ANALYSIS: ESTIMATOR = MLR; TYPE = COMPLEX;

    H1ITERATIONS = 20000;
    ITERATIONS = 100000;
processors =2;
define : CONTSCHL=(IDCNTRY*1000000)+IDCLASS; # define cluster variable
standardize PACH CACH EACH BACH PCW CCW ECW BCW;

    MODEL:
PSC by PSCP1@.818 PSCP2-PSCN2; CSC by CSCP1@.823 CSCP2-CSCN2;
ESC by ESCP1@.814 ESCP2-ESCN2; BSC by BSCP1@.791 BSCP2-BSCN2;

PIV by PIVP1@.867 PIVN1 PIVP2; CIV by CIVP1@.879 CIVN1 CIVP2;
EIV by EIVP1@.858 EIVN1 EIVP2;BIV by BIVP1@.841 BIVN1 BIVP2;

PVAL by PVAL1@.662 PVAL2-PVAL4; CVAL by CVAL1@.664 CVAL2-CVAL4;
EVAL by EVAL1@.531 EVAL2-EVAL4; BVAL by BVAL1@.566 BVAL2-BVAL4;

PmscXiv by Pmsciv1@.812 Pmsciv2 Pmsciv3 ;CmscXiv by Cmsciv1@.798 Cmsciv2 Cmsciv3;
EmscXiv by Emsciv1@.797 Emsciv2 Emsciv3 ;BmscXiv by Bmsciv1@.789 Bmsciv2 Bmsciv3;
!fixed factor loading of first indicator of each factor to provide common metric standardization;


!Correlated uniquenesses for negative worded items
  BSCN1 BSCN2 ESCN1 ESCN2 CSCN1 CSCN2 PSCN1 PSCN2  WITH
  BSCN1 BSCN2 ESCN1 ESCN2 CSCN1 CSCN2 PSCN1 PSCN2 ;

  EIVN1 BIVN1 CIVN1 PIVN1 WITH EIVN1 BIVN1 CIVN1 PIVN1 ;

  BSCN1 BSCN2 ESCN1 ESCN2 CSCN1 CSCN2 PSCN1 PSCN2  WITH
  EIVN1 BIVN1 CIVN1 PIVN1 ;

  ! Correlated uniquenesses for parallel items
  BSCP1 ESCP1 CSCP1 PSCP1 WITH BSCP1 ESCP1 CSCP1 PSCP1;
  BSCP2 ESCP2 CSCP2 PSCP2 WITH BSCP2 ESCP2 CSCP2 PSCP2;

  EVAL1 BVAL1 CVAL1 PVAL1 WITH  EVAL1 BVAL1 CVAL1 PVAL1 ;
  EVAL2 BVAL2 CVAL2 PVAL2 WITH  EVAL2 BVAL2 CVAL2 PVAL2 ;
  EVAL3 BVAL3 CVAL3 PVAL3 WITH  EVAL3 BVAL3 CVAL3 PVAL3 ;
  EVAL4 BVAL4 CVAL4 PVAL4 WITH  EVAL4 BVAL4 CVAL4 PVAL4 ;

  EIVP1 BIVP1 CIVP1 PIVP1 WITH  EIVP1 BIVP1 CIVP1 PIVP1 ;
  EIVP2 BIVP2 CIVP2 PIVP2 WITH  EIVP2 BIVP2 CIVP2 PIVP2 ;

  ! Correlated uniquenesses for parallel worded product terms
  Pmsciv1 Cmsciv1 Emsciv1 Bmsciv1 with Pmsciv1 Cmsciv1 Emsciv1 Bmsciv1;
  Pmsciv2 Cmsciv2 Emsciv2 Bmsciv2 with Pmsciv2 Cmsciv2 Emsciv2 Bmsciv2;
  Pmsciv3 Cmsciv3 Emsciv3 Bmsciv3 with Pmsciv3 Cmsciv3 Emsciv3 Bmsciv3;

! paths from achievement to self-concept
PSC on physi (pscVphy); PSC on chems (pscVche); PSC on earth (pscVear); PSC on biolo (pscVbio);
CSC on physi (cscVphy);CSC on chems (cscVche);CSC on earth (cscVear);CSC on biolo (cscVbio);
ESC on physi (escVphy);ESC on chems (escVche);ESC on earth (escVear);ESC on biolo (escVbio);
```

```
BSC on physi (bscVphy);BSC on chems (bscVche);BSC on earth (bscVear);BSC on biolo (bscVbio);
! paths from achievement to intrinsic value
PIV on physi (pivVphy);PIV on chems (pivVche);PIV on earth (pivVear);PIV on biolo (pivVbio);
CIV on physi (civVphy);CIV on chems (civVche);CIV on earth (civVear);CIV on biolo (civVbio);
EIV on physi (eivVphy);EIV on chems (eivVche);EIV on earth (eivVear);EIV on biolo (eivVbio);
BIV on physi (bivVphy);BIV on chems (bivVche);BIV on earth (bivVear);BIV on biolo (bivVbio);
! paths from achievement to utility value
PVAL on physi (pvaVphy);PVAL on chems (pvaVche);PVAL on earth (pvaVear);PVAL on biolo (pvaVbio);
CVAL on physi (cvaVphy);CVAL on chems (cvaVche);CVAL on earth (cvaVear);CVAL on biolo (cvaVbio);
EVAL on physi (evaVphy);EVAL on chems (evaVche);EVAL on earth (evaVear);EVAL on biolo (evaVbio);
BVAL on physi (bvaVphy);BVAL on chems (bvaVche);BVAL on earth (bvaVear);BVAL on biolo (bvaVbio);
PmscXiv-BmscXiv on physi-biolo; ! paths from achievement to product variables
phyCours- bioCours on PSC-BVAL; !paths from motivational factors to coursework aspirations
phyCours- bioCours on physi-biolo; ! product variables to coursework aspirations
phyCours- bioCours on physi-bioCours  ! achievement to coursework aspirations
! factor variances invariances
PSC-BmscXiv (fa1-fa16);  physi-bioCours (fa21-fa28); ! constrain factor variance to be equal
factor covariances invariances
PSC with CSC-BIV(p1-p11);CSC with ESC-BIV (c1-c10);ESC with BSC-BIV (e1-e9);
BSC with PVAL-BIV (b1-b8);PVAL with CVAL-BIV (pv1-pv7);CVAL with EVAL-BIV (cv1-cv6);
EVAL with BVAL-BIV (ev1-ev5);BVAL with PIV-BIV (bv1-bv4);PIV with CIV-BIV(pi1-pi3);
CIV with EIV BIV (ci1-ci2);EIV with BIV (ei1);

PSC with PmscXiv-BmscXiv (pscx1-pscx4);CSC with PmscXiv-BmscXiv (cscx1-cscx4);
ESC with PmscXiv-BmscXiv (escx1-escx4);BSC with PmscXiv-BmscXiv (bscx1-bscx4);
PIV with PmscXiv-BmscXiv (pivx1-pivx4);CIV with PmscXiv-BmscXiv (civx1-civx4);
EIV with PmscXiv-BmscXiv (eivx1-eivx4);BIV with PmscXiv-BmscXiv (bivx1-bivx4);
PVAL with PmscXiv-BmscXiv (pslx1-pslx4);CVAL with PmscXiv-BmscXiv (cslx1-cslx4);
EVAL with PmscXiv-BmscXiv (eslx1-eslx4);BVAL with PmscXiv-BmscXiv (bslx1-bslx4);

PmscXiv with CmscXiv EmscXiv BmscXiv (x23-x25);CmscXiv with EmscXiv BmscXiv (x26-x27);
EmscXiv with BmscXiv (x28);

physi with chems earth biolo (l1-l3);chems with earth biolo (l4-l5);earth with biolo (l6);

phyCours with cheCours earCours bioCours (l7-l9);cheCours with earCours bioCours (l10-l11);earCours with bioCours (l12);

Model HUN:
[PSCP1-Bmsciv3]; [PSC-BmscXiv@0]; !freely estimate items intercepts for each group
[PACH-BBORE];[physi-bioCours@0];
MODEL SVN:
[PSCP1-Bmsciv3]; [PSC-BmscXiv@0];
[PACH-BBORE];[physi-bioCours@0];
MODEL SWE:
[PSCP1-Bmsciv3]; [PSC-BmscXiv@0];
[PACH-BBORE];[physi-bioCours@0];

!!! create the summary for match and non-matching cross-paths from achievement to motivational factors (see Table E1-E3 in Appendix E in the supplemental materials)
MODEL CONSTRAINT:
!mean of 16 match and non-matching cross-paths involving self-concept
NEW(Mn_scVah,Mn_ivVah,Mn_vaVah,Mn_siVah);
Mn_scVah=(pscVphy+pscVche+pscVear+pscVbio+cscVphy+cscVche+cscVear+cscVbio+escVphy+
escVche+escVear+escVbio+bscVphy+bscVche+bscVear+bscVbio)/16;
```

```
        !mean of 16 match and non-matching cross-paths involving intrinsic value
        Mn_ivVah=(pivVphy+pivVche+pivVear+pivVbio+civVphy+civVche+civVear+civVbio+eivVphy+
        eivVche+eivVear+eivVbio+bivVphy+bivVche+bivVear+bivVbio)/16;
        !mean of 16 match and non-matching cross-paths involving utility value
        Mn_vaVah=(pvaVphy+pvaVche+pvaVear+pvaVbio+cvaVphy+cvaVche+cvaVear+cvaVbio+evaVphy+
        evaVche+evaVear+evaVbio+bvaVphy+bvaVche+bvaVear+bvaVbio)/16;

        NEW(Mm_scVah,Mm_ivVah,Mm_vaVah,Mm_siVah);
        Mm_scVah=(pscVphy+cscVche+escVear+bscVbio)/4; !mean of 4 matching cross-paths involving self-
concept
        Mm_ivVah=(pivVphy+civVche+eivVear+bivVbio)/4; !mean of 4 matching cross-paths involving intrinsic
value
        Mm_vaVah=(pvaVphy+cvaVche+evaVear+bvaVbio)/4; !mean of 4 matching cross-paths involving utility
value

        !NoMath
        NEW(No_scVah,No_ivVah, No_vaVah,No_siVah);
        No_scVah=(pscVche+pscVear+pscVbio+cscVphy+cscVear+cscVbio+escVphy+escVche+escVbio+bscVphy
+bscVche+bscVear)/12; !mean of 12 non-matching cross-paths involving self-concept
        No_ivVah=(pivVche+pivVear+pivVbio+civVphy+civVear+civVbio+eivVphy+eivVche+eivVbio+bivVph+bi
vVche+bivVear)/12; !mean of 12 non-matching cross-paths involving intrinsic value
        No_vaVah=(pvaVche+pvaVear+pvaVbio+cvaVphy+cvaVear+cvaVbio+evaVphy+evaVche+evaVbio+bvaVp
hy+bvaVche+bvaVear)/12; !mean of 12 non-matching cross-paths involving utility value
```

External Appendix C:
Weight and Goodness of Fit

Weighting

Consistent with its two-stage stratified sampling design, TIMSS provides the HOUWGT weighting variable that has six components, one each for school, class and student level, and one each for adjustment factors associated with non-participation at these three levels (See Marsh, Abduljabbar et al., 2013 for additional detail on the development of this weighting variable). HOUWGT is based on the actual number of students in each participating countries that is appropriate for correct computation of standard errors and tests of statistical significance. Thus, the HOUWGT weighting variable was taken into account in the data analysis.

Goodness of Fit

A number of traditional indices that are relatively independent of sample size were utilized to assess model fit (Hu & Bentler, 1999): the comparative fit index (CFI), the root-mean-square error of approximation (RMSEA) and the Tucker-Lewis Index (TLI). To explore how well the hypothesized relations generalize across the four OECD countries, we conducted multiple-group analyses (Bollen, 1989) and tested a series of increasingly stringent invariance constraints on the parameters of measurement and structural parts of the model, in which little or no change in goodness of fit supported invariance of the factor structure and parameter estimates (Millsap, 2011; see Appendix D in the supplemental materials for more details). We note that to compare differences in patterns of relations among multiple groups, it is only necessary to have factor loadings invariant for latent variable models (Millsap, 2011; Nagengast et al., 2011). Nevertheless, to facilitate interpretation of the parameter estimates in relation to a common metric over the multiple groups, we also tested invariance models of factor variances/covariances and path coefficients over the four countries (see Appendix C in the supplemental materials for the standardization procedure).

Values greater than .95 and .90 for CFI and TLI typically indicate excellent and acceptable levels of fit to the data. RMSEA values of less than .06 and .08 are considered to reflect good and acceptable levels of fit to the data. To explore how well the hypothesized relations generalize across the four OECD countries, we conducted multiple-group analyses (Bollen, 1989) and tested a series of increasingly stringent invariance constraints on the parameters of measurement and structural parts of the model, in which little or no change in goodness of fit supported invariance of the factor structure (Marsh, Hau et al., 2013). Chen (2007) have suggested that if the decrease in CFI is not more than .01 and the RMSEA increases by less than .015 for the more parsimonious model, then invariance assumptions are tenable. To facilitate interpretation of parameter estimates in relation to a common metric over the multiple groups, factor variances and covariances are also constrained to be invariant over the four countries in this study (see Appendix E in the supplemental materials for the standardization procedure). Other more stringent tests would have been necessary in order to support the test of latent mean differences over time or models based on the use of manifest, rather than latent, scale scores, which is not the case in the present study.

External Appendix D:
Preliminary Analyses Tests

Table D1

Model Fit Statistics for the CFA Models Used in the Present Study

Model	Description	χ^2	*df*	CFI	TLI	RMSEA
Total group (TG) analysis						
CFA						
TG1	SC + IV + UV	10757	722	.963	.952	.028
TG2	SC + IV + UV + SCxIV + SCxUV	16766	2090	.953	.942	.020
TG3	SC + IV + UV + SCxIV + SCxUV + ACH + ASP	19406	2506	.957	.946	.019
Second-order CFA model						
TG4	SO(SC + IV +UV)	40918	773	.852	.820	.054
Multi-group (MG) analysis						
CFA						
MG1	SC + IV + UV + ACH + ASP, CUs, Configural	18038	3912	.964	.951	.028
MG2	SC + IV + UV + ACH + ASP, CUs, IN = FL	19416	4008	.961	.948	.029
MG3	SC + IV + UV + ACH + ASP, CUs, IN = FL, FV	20138	4068	,959	.947	.030
MG4	SC + IV + UV + ACH + ASP, CUs, IN = FL, FV, CV	23961	4638	.951	.944	.030

Note. CFA = confirmatory factor analysis; SEM = Structural equation modelling; PC = path coefficients; SC = self-concept; IV= intrinsic value; UV = utility value; ASP = coursework aspirations; SCxIV = the product term of self-concept by intrinsic value interaction; SCxUV = the product term of self-concept by utility value interaction; IN = invariant; CUs = correlated uniquenesses; UCUs = uncorrelated uniquenesses; FL = factor loading; FV = factor variances; CV = factor covariances; INT= item intercepts; Unq = item uniquenesses; FMn = factor latent mean.

Factor Structure: Preliminary CFA

Total group CFA. In the preliminary analyses, we evaluated a series of CFAs of the factor structures underlying the multiple domains of self-concept, intrinsic value and utility value, and their relations to parallel measures of achievement and coursework aspirations. We began with an evaluation of the results based on the total group. A critical feature of the TIMSS data is that each motivation construct was measured by a mixture of positively and negatively worded items, with parallel wording across the four science domains. This requires the inclusion of a priori correlated uniquenesses, relating responses to negatively worded items and parallel worded items, to obtain unbiased parameter estimates (see Marsh, Abduljabbar et al., 2013, 2015). Following previous TIMSS research (Marsh, Abduljabbar et al., 2013), these a priori correlated uniquenesses were included in all CFA and SEM models. The goodness of fit for the CFA models with proper methodological control for item wordings was good (e.g., CFI &TLI > .942; see models TG1–TG3 in Table D1).

We also tested a second-order CFA model where global science self-concept, intrinsic value and utility value were formed by the four corresponding first-order constructs from each science domain. However, the second-order CFA model was highly unsatisfactory in terms of model fits (e.g., CFI & TLI < .852; Model TG4), thus providing support for the domain specificity and discriminant validity of these factors. This result indicates that it is important to distinguish the patterns of theoretical predictions in relation to each of the four science domains.

Tests of Invariance of Factorial Structure Over Countries: Multi-group CFA

A key interest of the present study is to evaluate the degree to which the results generalize across the four OECD countries included in our sample. We began with an evaluation of invariance of the factor structure over multiple groups (four OECD countries) based on CFAs. The fit indices for the baseline model with no invariance constraints were very good (e.g., CFI = .964, Model MG1 in Table 2). There was a negligible decrease in fit (Δ CFI = .003, Δ TLI = .003) for Model MG2, in which the factor loadings were constrained to be equal across groups, suggesting that the invariance of factor loadings was supported by the data. Similarly, adding equality constraints on the factor variances (MG3) and covariances (MG4) resulted in a satisfactory level of fit to the data, and only a negligible change in fit (Δ CFI = .008, Δ TLI = .003). These results support the generalizability of the factor structure of the five constructs across the four countries.

Domain specificity of Motivation Responses, Achievement, and Aspirations

We examined relations among the five constructs to evaluate the expected domain specificity of the motivation responses. Latent correlations among the 20 constructs (4 domains x 5 constructs) based on Model MG4 with invariant factor loadings, variances, and covariances for motivational beliefs, achievement, and aspirations over the four OECD countries, are presented in Table D2(below). The latent correlations among the four self-concept factors (r = .28 to .42) and among the four intrinsic value factors (r = .23 to .40) in different science domains were modest. These correlations were smaller than those among utility value factors (r = .46 to .65). Of particular relevance, correlations among the four coursework aspirations (r = .21 to .38) were much smaller than those among the four achievement scores (r = .77 to .81). In summary, there was good support for the high domain specificity of self-concept and intrinsic value, but the support for utility value was much weaker. Our findings also provided good support for the domain specificity of coursework aspirations but relatively weak support for the domain specificity of achievement scores.

Latent correlations among all the constructs (4 latent factors for each self-concept, intrinsic value and utility value, and 4 corresponding measure of achievement and coursework aspirations) are presented in Table 2 based on Model MG4. In support of the convergent validity of latent constructs, correlations between each self-concept and the matching intrinsic value were consistently substantial (rs vary from .74 - .79). Also, convergent validity correlations involving utility value were consistently moderate for matching domains of self-concept and intrinsic value (*rs* vary from .41 to .34 and from .45 to .51 respectively). Both self-concept and intrinsic value were highly correlated with the matching area of coursework aspirations (convergent validities, *rs* vary from .57 - .62 and from .75 to .78 respectively), whereas the correlations between each utility value and matching domains of aspirations were moderate (*rs* vary from .40 to .45). However, the convergent validity correlations involving self-concept and intrinsic value were somewhat weaker for the corresponding measure of achievement scores (*rs* vary from .24 to .32 and from .09 to .16 respectively). The sizes of convergent validities of utility value in relation to achievement scores were substantially small and even non-significant (*rs* vary from -.03 to .09). Achievement scores also have relatively weak convergent validity in relation to coursework aspirations (*rs* vary from .05 to .08).

In support of the discriminant validities of self-concept and intrinsic value, the convergent validities were substantially larger than correlations among self-concept factors (*rs* vary from .31 to .42) and among intrinsic value factors (*rs* vary from .21 to .40) and correlations between each self-concept factor and non-matching domains of intrinsic value (*rs* vary from .14 to .90). In support of the discriminant validities of self-concept and aspirations, the convergent validities were much larger than correlations between each self-concept and non-matching measure of coursework aspirations (*rs* vary from .09 to .22) and correlations among domain-specific aspirations (*rs* vary from .21 to .38). Similar patterns were also found for the convergent validities of intrinsic value in relation to aspirations. Although the convergent validities involving utility value were higher than correlations of utility value to non-matching domains of self-concept, intrinsic value and aspirations, these convergent validities were weaker than correlations among utility value factors (*rs* vary from .46 to .65). Likewise, whilst the convergent validities involving achievement were slightly larger than correlations of achievement to non-matching area of self-concept, intrinsic value and aspirations, they were much smaller than correlations among science domain-specific achievement scores (*rs* vary from .77 - .81).

In summary, consistent with previous research (e.g., Marsh, Abduljabbar et al., 2013), self-concept was more highly correlated with achievement scores, whereas intrinsic value was more highly correlated with coursework aspirations. The results provided good support for the convergent and discriminant validity of self-concept, intrinsic value and coursework aspirations in relation to each other. Whereas there was good support for the convergent validity of utility value, utility value only had limited discriminant validity in relation to self-concept, intrinsic value and coursework aspirations. Achievement scores had weak convergent validity and discriminant validity in relation to self-concept, intrinsic value and aspirations but not utility value.

Table D2

Latent Correlations Among Self-Concept, Intrinsic Value, Utility Value, Achievement Scores and Coursework Aspirations Based on Four Science Domains

	Science self-concept				Science intrinsic value				Science utility value				Science achievement				Science aspirations			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Science self-concept																				
1.PSC	–																			
2.CSC	.42	–																		
3.ESC	.34	.28	–																	
4.BSC	.31	.41	.35	–																
Science intrinsic value																				
5.PIV	.79	.27	.20	.18	–															
6.CIV	.27	.79	.15	.30	.38	–														
7.EIV	.20	.14	.77	.19	.28	.23	–													
8.BIV	.13	.25	.15	.74	.25	.40	.28	–												
Science utility value																				
9.PUV	.41	.23	.14	.14	.50	.28	.19	.18	–											
10.CUV	.19	.40	.11	.22	.25	.49	.16	.29	.65	–										
11.EUV	.13	.14	.34	.13	.18	.19	.45	.19	.50	.53	–									
12.BUV	.13	.21	.11	.39	.18	.30	.19	.51	.46	.63	.56	–								
Science achievement																				
13.PACH	.32	.25	.21	.20	.16	.12	.07	.02	.09	-.01	-.05	-.02	–							
14.CACH	.31	.25	.21	.26	.16	.13	.07	.08	.09	.02	-.02	.03	.77	–						
15.ECAH	.29	.23	.24	.25	.13	.09	.09	.06	.07	-.02	-.03	-.02	.77	.78	–					
16.BACH	.28	.25	.23	.30	.12	.12	.08	.11	.04	-.01	-.05	.00	.80	.80	.81	–				
Science coursework aspirations																				
17.PAPS	.57	.20	.12	.11	.75	.30	.21	.18	.44	.25	.18	.18	.08	.05	.04	.03	–			
18.CAPS	.21	.62	.09	.22	.30	.78	.16	.31	.26	.44	.19	.27	.07	.05	.04	.06	.38	–		
19.EAPS	.14	.09	.60	.13	.21	.16	.76	.19	.18	.15	.40	.17	.03	.02	.06	.04	.26	.21	–	
20.BAPS	.09	.19	.11	.57	.18	.31	.20	.75	.17	.27	.19	.45	.01	.03	.04	.06	.22	.35	.25	–

Note. P = physics; C = chemistry; E = earth science; B = biology; SC = self-concept; IV= intrinsic value; UV = utility value; Standardized errors for all correlation coefficients are approximately .01. All correlations greater than .023 or less than -.023 are statistically significant (*p* < .05); shaded correlations are convergent validity coefficients involving two constructs in matching domains.

External Appendix E:
Full Results for the path invariance Model (Model 7b)

Table E1

The Predictive Effects of Achievement on Self-Concept, Intrinsic Value and Utility Value Based on Four Science Domains

	Motivation outcome variables		
	Model MG7b		
Predictors	Self-concept	Intrinsic value	Utility value
	Physics		
Physics Ach	.16 (.02)*	.13 (.02)*	.09 (.02)*
Chemistry Ach	.17 (.02)*	.11 (.02)*	.11 (.02)*
Earth science Ach	.06 (.02)*	-.00 (.02)	.01 (.02)
Biology Ach	-.09 (.02)*	-.08 (.02)*	-.13 (.02)*
	Chemistry		
Physics Ach	.09 (.02)*	.03 (.02)	.00 (.02)
Chemistry Ach	.12 (.02)*	.11 (.02)*	.09 (.02)*
Earth science Ach	.01 (.02)	.06 (.02)*	-.08 (.02)*
Biology Ach	.07 (.02)*	.04 (.02)*	-.01 (.02)
	Earth science		
Physics Ach	.00 (.02)	-.03 (.02)	-.07 (.02)*
Chemistry Ach	.00 (.02)	-.03 (.02)	.08 (.02)*
Earth science Ach	.15 (.02)*	.10 (.02)*	.02 (.02)
Biology Ach	.11 (.02)*	.04 (.02)*	-.07 (.02)*
	Biology		
Physics Ach	-.14 (.02)*	-.18 (.02)*	-.08 (.02)*
Chemistry Ach	.08 (.02)*	.05 (.02)*	.13 (.02)*
Earth science Ach	.04 (.02)*	.03 (.02)	-.07 (.02)*
Biology Ach	.32 (.02)*	.23 (.02)*	.02 (.02)
Summary (Means across different sets of path coefficients based on 4 domains			
Mn Total	.07 (.00)*	.03 (.00)*	.00 (.00)
Mn Match	.19 (.01)*	.14 (.02)*	.05 (.01)*
Mn NoMatch	.03 (.01)*	-.02 (.00)*	-.02 (.00)*
Difference	.16 (.01)*	.16 (.01)*	.07 (.01)*

Note. SC = self-concept; IV= intrinsic value; UV = utility value; Ach = Achievement; shaded estimates are path coefficients from achievement to motivational constructs in the matching domain; * $p < .05$

Table E2

The Predictive Effects of Self-Concept, Intrinsic Value, Utility Value and Their Interactions on Coursework Aspiration Based on Four Science Domains

		Motivation predictors		
		Model MG7b		
		SC	IV	UV
Physics Coursework Aspirations	Physics	.07 (.03)*	.68 (.03)*	.06 (.01)*
	Chemistry	-.02 (.03)	.04 (.03)	.02 (.01)
	Earth science	.05 (.02)*	.06 (.02)*	.01 (.01)
	Biology	-.05 (.02)*	-.08 (.02)*	.03 (.01)*
Chemistry Coursework Aspirations	Physics	-.01 (.03)	.02 (.03)	.02 (.01)
	Chemistry	.08 (.03)*	.69 (.03)*	.06 (.01)*
	Earth science	.06 (.02)*	.02 (.02)	.01 (.01)
	Biology	.02 (.02)	-.02 (.02)	.01 (.01)
Earth science Coursework Aspirations	Physics	-.05 (.03)	.06 (.03)	.01 (.01)
	Chemistry	.03 (.02)	-.04 (.02)	.01 (.01)
	Earth science	.10 (.03)*	.66 (.02)*	.06 (.01)*
	Biology	-.03 (.02)	.01(.02)	.01 (.01)
Biology Coursework Aspirations	Physics	-.05 (.02)*	.01 (.03)	.02 (.01)
	Chemistry	-.02 (.03)	.05 (.02)*	-.01 (.01)
	Earth science	-.04 (.02)	.02 (.02)	.01 (.01)
	Biology	.13 (.02)*	.63 (.02)*	.07 (.01)*
Mn Total		.03 (.01)*	.17(.00)*	.03 (.00)*
Mn Match		.10(.01)*	.67 (.01)*	.06 (.01)*
Mn NoMatch		.01 (.01)	.01 (.01)	.02 (.00)*
Difference		.09 (.02)*	.66 (.01)*	.05 (.01)*

Note. SC = self-concept; IV= intrinsic value; UV = utility value; SCxIV = self-concept by intrinsic value interaction; SCxUV = self-concept by utility value interaction; shaded estimates are path coefficients from motivational constructs to coursework aspirations in the matching domain; * $p < .05$.

Table E3

The Predictive Effects of Self-Concept, Intrinsic Value, Utility Value and Their Interactions on Coursework Aspiration Based on Four Science Domains

		Motivation predictors							
		Model MG8b				Model MG9b			
Outcomes		SC	IV	UV	SCxIV	SC	IV	UV	SCxUV
Physics Coursework Aspirations	Physics	.08 (.03)*	.67 (.03)*	.07 (.01)*	.09 (.01)*	.08 (.03)*	.72 (.03)*	.06 (.01)*	.09 (.01)*
	Chemistry	-.02 (.03)	.05 (.03)	.02 (.01)	-.02 (.01)	-.02 (.02)	.03 (.02)	.02 (.01)	-.01 (.01)
	Earth science	.05 (.02)*	.06 (.02)*	.01 (.01)	-.02 (.01)*	.05 (.02)*	.04 (.02)	.01 (.01)	-.00 (.01)
	Biology	-.05 (.02)*	-.07 (.02)*	.03 (.01)*	-.01 (.01)	-.05 (.02)*	-.06 (.02)*	.02 (.01)	-.00 (.01)
Chemistry Coursework Aspirations	Physics	-.01 (.03)	.03 (.03)	.02 (.01)	-.04 (.01)*	.01 (.02)	-.01 (.02)	.02 (.01)	.01 (.010)
	Chemistry	.10 (.03)*	.66 (.03)*	.05 (.01)*	.12 (.01)*	.07 (.03)*	.71 (.03)*	.06 (.01)*	.08 (.01)*
	Earth science	.06 (.02)*	.03 (.02)	.01 (.01)	-.03 (.01)*	.06 (.02)*	.01 (.02)	.02 (.01)	-.02 (.009)
	Biology	.02 (.02)	-.03 (.02)	.01 (.01)	.01 (.01)	.01 (.02)	-.02 (.02)	-.01 (.01)	.05 (.010)
Earth science Coursework Aspirations	Physics	-.05 (.03)	.05 (.03)	.01 (.01)	-.02 (.01)*	-.03 (.02)	.02 (.02)	.02 (.01)	.01 (.01)
	Chemistry	.03 (.02)	-.04 (.02)	.01 (.01)	.01 (.01)	.02 (.02)	-.04 (.02)	-.01 (.01)	.02 (.01)
	Earth science	.13 (.03)*	.64 (.02)*	.05 (.01)*	.13 (.01)*	.08 (.02)*	.70 (.02)*	.05 (.01)*	.09(.01)*
	Biology	-.03 (.02)	.01(.02)	.01 (.01)	-.02 (.01)*	-.02 (.02)	-.00 (.02)	.01 (.01)	-.02 (.01)*
Biology Coursework Aspirations	Physics	-.05 (.02)*	.01 (.03)	.02 (.01)	-.03 (.01)*	-.05 (.02)*	-.02 (.02)	.02 (.01)	.00 (.01)
	Chemistry	-.02 (.03)	.05 (.02)*	-.01 (.01)	-.01 (.01)	-.01 (.02)	.05 (.02)*	-.01 (.01)	-.02 (.01)
	Earth science	-.04 (.02)	.02 (.02)	.01 (.01)	-.03 (.01)*	-.04 (.02)	.01 (.02)	.01 (.01)	-.01 (.01)
	Biology	.14 (.02)*	.62 (.02)*	.08 (.01)*	.15 (.01)*	.09 (.02)*	.64 (.02)*	.07 (.01)*	.09 (.01)*
Summary (Means across different sets of path coefficients based on 4 domains									
Mn Total		.03 (.01)*	.17(.00)*	.03 (.00)*	.02 (.00)*	.01 (.01)	.18 (.00)*	.02 (.00)*	.02 (.00)*
Mn Match		.10(.01)*	.64 (.01)*	.06 (.01)*	.12 (.01)*	.08 (.01)*	.70 (.01)*	.06 (.01)*	.09 (.01)*
Mn NoMatch		.01 (.01)	.01 (.01)	.02 (.00)*	-.02 (.00)*	-.01 (.01)	.00 (.00)	.01 (.00)*	-.00 (.00)
Difference		.09 (.02)*	.66 (.01)*	.05 (.01)*	.14 (.01)*	.08 (.02)*	.70 (.02)*	.06 (.01)*	.08 (.01)*

Note. SC = self-concept; IV= intrinsic value; UV = utility value; SCxIV = self-concept by intrinsic value interaction; SCxUV = self-concept by utility value interaction; shaded estimates are path coefficients from motivational constructs to coursework aspirations in the matching domain; * $p < .05$.

External Appendix F:
Supplemental Analyses for the mediating role of academic self-concept

Given that some previous studies suggested that the dimensional comparisons between achievements were likely to indirectly affect value beliefs via the mediating role of self-concept (e.g., Nagy et al., 2008), we further tested whether the effects of achievement on task value could be explained by the effects of self-concept. To this end, we evaluated a mediation model in which the achievements in the four science domains predicted self-concepts, which in turn predicted intrinsic and utility values. In this model, the four domain-specific self-concepts and values along with achievements were also allowed to predict coursework aspirations. This mediation model provided an identical fit with original hypothesized model (Model MG7b). The results revealed that all 32 direct paths from the four science achievements to each of the intrinsic value and each of utility values were relatively small (from -.05 to .05; M = .01). Subsequently, we evaluated a nested model where these 32 direct paths were constrained to be 0. There was a negligible decrease in model fit (D CFI = .002, D TLI = .001, D RMSEA = .001) when compared to the fully mediated model, suggesting that self-concept entirely mediated the effects of achievement on values.

External Appendix G:
Supplemental Analyses for Interaction Effect Between Self-Concept and Value

In the main text, latent interactions between self-concept and intrinsic value as well as between self-concept and utility value, when these two multiplicative terms (self-concept x intrinsic value and self-concept x utility value) are considered separately.

Subsequently, we included the two sets of latent interactions into the same model (i.e, Model MG10a – MG10c in Table G1). All first-order effects and interaction effects between self-concept and intrinsic value were significantly positive and similar in size with the pattern of results from Model MG8a-MG8b (See Table 2 in the main text) where only self-concept and intrinsic value interactions were included (see Table G2). However, the interactions between self-concept and utility value lost their predictive power on coursework aspirations. Given that correlations between matching domains of latent product variables were substantial (*r* = .58 to .69, Table G3), we argue that the parameters involving interaction effects in this model should be interpreted with caution. In Model MG11c, we constrained the paths leading from self-concept by intrinsic value interactions to aspirations and those from self-concept by utility value interactions to be equal. The model fits the data as well, and there was a very small decrement in CFI (D .001) and RMSEA (D .001) but no difference in TLI in comparison to Model MG10c. We also found a notable reduction in the size of the standard errors (from [.011 to .016] to [.004 to .006]) associated with the paths from all domain-specific interactions to aspirations. The results for this model show that all domain-specific interactions positively predicted matching measures of aspirations (*M* = .06, SE = .003). Thus, the results suggest that both types of domain-specific latent interaction (self-concept-by- intrinsic value, and self-concept-by- utility value) may make similar contributions to the prediction of coursework aspirations (Marsh, Dowson, Pietsch & Walker, 2004).

Table G1

Model Fit Statistics for CFA and SEM Models Used in The Present Study

Model	Description	χ^2	*df*	CFI	TLI	RMSEA
SEM						
MG10a	SC + IV + UV + ACH + ASP, CUs, INV = FL, FV, PC; Free = PT(scXiv, scXuv)	33843	10516	.946	.936	.022
MG10b	SC + IV + UV + PT(scXiv,scXuv) + ACH + ASP, CUs, INV = FL, FV, PC	34934	10792	.945	.935	.022
MG10c	SC + IV + UV + PT(scXiv,scXuv) + ACH + ASP, CUs, INV = FL, FV, CV, PC	38453	11398	.939	.931	.023
MG11a	SC + IV + UV + ACH + ASP, CUs, INV = FL, FV, PC; Free = PT(scXiv, scXuv), PC (scXiv = scXuv)	34039	10580	.946	.936	.022
MG11b	SC + IV + UV + PT(scXiv,scXuv) + ACH + ASP, CUs, INV = FL, FV, PC (scXiv = scXuv)	35081	10808	.944	.935	.022
MG11c	SC + IV + UV + PT(scXiv,scXuv) + ACH + ASP, CUs, INV = FL, FV, CV, PC (scXiv = scXuv)	38629	11414	.938	.931	.023

Note. SC = self-concept; IV= intrinsic value; UV = utility value; PT = product term; ASP = coursework aspirations; scXiv = the product term of self-concept by intrinsic value interaction; scXuv = the product term of self-concept and utility value interaction; INV = invariant; CUs = correlated uniquenesses; UCUs = uncorrelated uniquenesses; FL = factor loading; FV = factor variances; CV = factor covariances; Free = PT (scXiv): freely estimate factor loading, factor variances and covariance and path coefficients with respect to scXiv; Free = PT (scXuv): freely estimate factor loading, factor variances and covariance and path coefficients with respect to scXuv; PC (scXiv = scXuv): constrain the path coefficients from scXiv to ASP and from scXuv to ASP to be equal

Table G2

The Predictive Effects of Self-Concept, Intrinsic Value, Utility Value and Their Interactions on Coursework Aspiration Based on Four Science Domains (Standardized Path Coefficients as A Ratio of Standard Errors)

Outcome↓		Motivation Predictors								
		Model MG10c					Model MG11c			
		SC	IV	UV	SCxIV	SCxUV	SC	IV	UV	scXiv = scXuv
Physics Coursework Aspirations	Physics	.064/.028	.686/.030	.083/.016	.063/.015	.045/.013	.060/.027	.696/.029	.078/.016	.053/.005
	Chemistry	-.026/.025	.040/.026	.016/.013	-.021/.013	.004/.013	-.019/.025	.032/.025	.018/.013	-.008/.005
	Earth science	-.071/.024	.054/.023	.004/.011	-.025/.012	.008/.013	-.062/.023	.044/.022	.007/.011	-.009/.005
	Biology	-.047/.023	-.065/.022	.022/.012	.002/.013	-.008/.015	-.047/.023	-.061/.022	.022/.012	-.002/.005
Chemistry Coursework Aspirations	Physics	-.011/.027	.024/.027	.023/.015	-.048/.014	.024/.013	.017/.028	-.012/.027	.030/.015	-.008/.006
	Chemistry	.074/.027	.689/.027	.057/.013	.113/.013	.005/.014	.061/.028	.718/.028	.053/.013	.059/.006
	Earth science	-.061/.023	.019/.022	.017/.011	-.031/.013	.001/.013	-.061/.024	.019/.022	.017/.011	-.015/.005
	Biology	.018/.022	-.016/.021	-.002/.013	-.010/.013	.018/.014	.015/.022	-.013/.021	-.004/.013	.003/.005
Earth science Coursework Aspirations	Physics	-.067/.027	.024/.028	.019/.014	-.051/.013	.029/.013	-.035/.027	.024/.027	.027/.015	-.009/.006
	Chemistry	.028/.024	-.043/.024	.006/.014	-.011/.012	.020/.013	.029/.025	-.045/.025	.007/.014	.006/.005
	Earth science	.101/.025	.662/.024	.060/.011	.130/.014	-.001/.013	.067/.025	.703/.024	.048/.011	.063/.005
	Biology	-.033/.022	.005/.021	.011/.013	-.021/.014	-.005/.014	-.024/.022	-.001/.021	.012/.013	-.015/.004
Biology Coursework Aspirations	Physics	-.023/.027	.004/.028	.021/.016	-.036/.013	.017/.011	-.006/.026	-.019/.026	.020/.015	-.006/.005
	Chemistry	-.048/.024	.046/.024	-.001/.014	-.017/.012	.006/.012	-.054/.024	.052/.025	-.001/.014	-.010/.005
	Earth science	-.042/.023	.018/.022	.007/.012	-.030/.013	.007/.013	-.041/.023	.017/.021	.016/.012	-.014/.005
	Biology	.123/.023	.635/.022	.077/.012	.159/.015	-.019/.016	.110/.023	.644/.022	.063/.013	.073/.004
Summary (Means across different sets of path coefficients based on 4 domains)										
Mn Total		.003/.003	.174/.003	.027/.002	.010/.003	.010/.003	.003/.003	.175/.003	.027/.002	.010/.001
Mn Match		.090/.012	.668/.013	.069/.006	.116/.007	.008/.007	.075/.013	.690/.013	.061/.006	.062/.003
Mn NoMacth		-.032/.005	.011/.005	.013/.003	-.025/.003	.010/.003	-.024/.005	.003/.005	.015/.003	-.007/.001
Difference		.122/.016	.657/.017	.056/.007	.141/.008	-.002/.008	.099/.017	.687/.017	.046/.007	.069/.004

Note. SC = self-concept; IV= intrinsic value; UV = utility value; scXiv = self-concept by intrinsic value interaction; scXuv = self-concept by utility value interaction. Shaded estimates are path coefficients from motivational constructs to coursework aspirations in the matching domain

Table G3

Latent Correlation Among Product Variables Based on Four Science Domains

	Self-concept by intrinsic value				Self-concept by utility value			
	1	2	3	4	5	6	7	8
Science self-concept by intrinsic value								
1.PSCxIV	–							
2.CSCxIV	.38	–						
3.ESCxIV	.25	.21	–					
4.BSCxIV	.22	.33	.28	–				
Science self-concept by utility value								
5.PSCxUV	.58	.20	.15	.14	–			
6.CSCxUV	.24	.59	.13	.20	.35	–		
7.ESCxUV	.14	.10	.57	.14	.23	.21	–	
8.BSCxUV	.13	.15	.14	.69	.19	.26	.21	–

Note. P = physics; C = chemistry; E = earth science; B = biology; SC = self-concept; IV = intrinsic value; UV = utility value; scXiv = self-concept by intrinsic value interaction; scXuv = self-concept by utility value interaction.

External Appendix H:
Full Results for country-specific path coefficients (Model MG7a)

Table H1

The Predictive Effects of Achievement on Self-Concept, Intrinsic Value and Utility Value Based on Four Science Domains (Czech Republic/Hungary/Slovenia/ Sweden)

	Motivation outcome variables		
	Model MG6		
Predictors	Self-concept	Intrinsic value	Utility value
		Physics	
Physics Ach	.144/.170/.186/.236	.125/.097/.147/.082	.094/.078/.083/.106
Chemistry Ach	.218/.151/.224/.154	.145/.119/.141/.085	.159/.085/.151/.094
Earth science Ach	.069/.075/.080/.071	.016/-.001/-.013/.013	-.005/.024/-.024/.026
Biology Ach	-.118/-.101/-.078/-.105	-.066/-.058/-.091/-.119	-.128/-.115/-.131/-.196
		Chemistry	
Physics Ach	.112/.078/.086/.102	.034/.016/-.015/-.025	.016/-.014/.042/.041
Chemistry Ach	.121/.104/.142/.119	.131/.088/.089/.143	.084/.132/.089/.061
Earth science Ach	.008/.017/.010/.018	.049/.050/.062/.086	-.118/-.075/-.130/-.081
Biology Ach	.079/.083/.082/.065	.039/.066/.085/.044	.024/.057/.017/-.017
		Earth science	
Physics Ach	-.010/-.010/.020/.014	-.045/-.047/.019/.023	-.092/-.114/-.075/-.094
Chemistry Ach	-.013/.020/.013/-.016	-.056/.011/.005/-.038	.111/.125/.053/.074
Earth science Ach	.178/.187/.125/.114	.123/.148/.085/.096	-.026/.019/-.018/.035
Biology Ach	.094/.074/.147/.073	.067/.053/.054/.040	-.088/-.094/-.053/-.097
		Biology	
Physics Ach	-.122/-.114/-.067/-.082	-.212/-.164/-.156/-.197	-.062/-.112/-.077/-.069
Chemistry Ach	.093/.079/.069/.094	.091/.061/.053/.068	.101/.172/.111/.141
Earth science Ach	.037/.084/.044/.064	.017/.053/-.022/.047	-.106/-.071/-.092/-.062
Biology Ach	.267/.271/.348/.252	.232/.232/.280/.175	.032/.017/.042/.024

Note. SC = self-concept; IV= intrinsic value; UV = utility value; Ach = Achievement; shaded estimates are path coefficients from achievement to motivational constructs in the matching domain; * $p < .0$

Table H2

The Predictive Effects of Self-Concept, Intrinsic Value, Utility Value and Their Interactions on Coursework Aspiration Based on Four Science Domains (Czech Republic/Hungary/Slovenia/ Sweden)

		Motivation predictors		
		Model MG6b		
		SC	IV	UV
Physics Coursework Aspirations	Physics	.064/.066/.116/.084	.810/.657/.787/.737	.074/.064/.114/.087
	Chemistry	-.030/-.015/-.009/-.003	.019/.056/.033/.060	.017/-.018/-.013/.039
	Earth science	.075/.061/.069/.058	.085/.084/.046/.065	-.018/.034/-.001/.007
	Biology	-.077/-.048/-.087/-.053	-.112/-.670/-.112/-.073	.035/.064/.059/.050
Chemistry Coursework Aspirations	Physics	.017/.013/.013/-.019	.025/-.014/.038/.029	.032/.041/-.003/.045
	Chemistry	.072/.135/.128/.111	.775/.801/.676/.735	.054/.057/.071/.053
	Earth science	.059/.650/.083/.073	.036/.032/-.011/.038	.013/-.004/-.007/.020
	Biology	.036/.025/-.028/-.023	-.023/-.024/-.011/.004	.004/.034/-.013/.003
Earth science Coursework Aspirations	Physics	-.010/-.002/-.048/-.058	.041/.040/.022/.061	.024/-.004/.012/.028
	Chemistry	.016/.038/.022/.012	-.011/-.013/-.044/-.016	-.019/-.015/.025/.010
	Earth science	.063/.076/.113/.092	.672/.572/.692/.747	.057/.058/.103/.067
	Biology	-.025/-.035/-.041/-.043	.011/.010/.005/-.016	.006/.020/.014/-.006
Biology Coursework Aspirations	Physics	-.058/-.045/-.053/-.074	-.020/-.018/-.012/.016	.021/.003/.028/.038
	Chemistry	-.008/-.028/-.015/-.040	.047/.085/.053/.056	-.021/.011/-.008/-.002
	Earth science	-.052/-.018/-.015/-.010	.018/.007/-.020/-.037	.029/-.009/-.007/.033
	Biology	.116/.081/.117/.161	.627/.602/.710/.667	.126/.138/.090/.071

Note. SC = self-concept; IV= intrinsic value; UV = utility value; SCxIV = self-concept by intrinsic value interaction; SCxUV = self-concept by utility value interaction; shaded estimates are path coefficients from motivational constructs to coursework aspirations in the matching domain; * $p < .05$

References for Online Supplemental Materials

Bollen, K. A. (1989). Structural equations with latent variables. New York: Wiley.

Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. doi:10.1080/10705510701301834

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. doi:10.1080/10705519909540118

Jöreskog, K. G., & Yang, F. (1996). Nonlinear structural equation models: The Kenny–Judd model with interaction effects. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 57–87).Mahwah, NJ: Lawrence Erlbaum Associates.

Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J. S., Abdelfattah, F., Leung, K. C., … Parker, P. D. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology*, *105*(1), 108–128. doi:10.1037/a0029907

Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F., Nagengast, B., … Abu-Hilal, M. M. (2015). The Internal/External Frame of Reference Model of Self-Concept and Achievement Relations: Age-Cohort and Cross-Cultural Differences. *American Educational Research Journal*, *52*(1), 168–202. doi:10.3102/0002831214549453

Marsh, H. W., Dowson, M., Pietsch, J., & Walker, R. (2004). Why Multicollinearity Matters: A Reexamination of Relations Between Self-Efficacy, Self-Concept, and Achievement. *Journal of Educational Psychology*, *96*(3), 518–522. doi:10.1037/0022-0663.96.3.518

Marsh, H. W., Hau, K-T.,Wen, Z., Nagengast, B. & Morin, A. J. S. (2013). Statistical Models of Moderation. In Little, T. (Ed.), *Oxford Handbook of Quantitative Methods*. New York: Oxford University Press.

Marsh, H. W., Wen, Z., & Hau, K.-T. (2004). Structural equation models of latent interactions: evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, *9*(3), 275–300. doi:10.1037/1082-989X.9.3.275

Mullis, I. V. S., Martin, M. O., Olson, J. F., Berger, D. R., & Stanco, G. M. (Eds.). (2008). TIMSS 2007 encyclopedia: A guide to mathematics and science education around the world (Vols. 1 and 2). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

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

Wen, Z., Marsh, H. W., & Hau, K.-T. (2010). Structural Equation Models of Latent Interactions: An Appropriate Standardized Solution and Its Scale-Free Properties. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(1), 1–22.