Marsh, Lüdtke, Nagengast, Trautwein, Abduljabbar, Abdelfattah, & Jansen. (2015). Dimensional Comparison Theory: Paradoxical relations between self-beliefs and achievements in multiple domains. *Learning and Instruction, 35,* 16-32.

Abstract

The internal/external frame of reference (I/E) model posits paradoxical relations between achievement and self-concept in mathematics and verbal domains, in which achievement in each domain has a positive effect on self-concept in the matching domain (e.g., mathematics achievement on mathematics self-concept) but a negative (contrastive) effect on self-concept in the non-matching domain (e.g., mathematics achievement on verbal self-concept). Extending the I/E model, Dimensional Comparison Theory (DCT) posits that self-evaluations are based on dimensional comparisons (e.g., how my accomplishments in one domain compare with my accomplishments in another domain) as well as the more traditional social and temporal comparisons, and on other sources of information about one's accomplishments. Extending the traditional tests of the I/E model, DCT predicts strong contrast effects only for contrasting domains that are at the opposite ends of the theoretical continuum of academic self-concept (far comparisons: e.g., the negative effect of math achievement on verbal self-concept), but much weaker negative contrast or even positive assimilation effects for complementary domains that are close to each other (near domains: e.g., positive effects of math achievement on physics self-concept; positive effects of native language on foreign language self-concept). Here we illustrate new predictions, theoretical insights, and methodology associated with DCT based on multiple academic domains (native language, foreign language, history, biology, physics and math), showing significant contrast effects for far comparisons and significantly less contrast or assimilation effects for near domains.

Keywords

Internal/external frame of reference effect Assimilation and contrast effects Social and dimensional comparison processes Self-concept vs. self-efficacy Multilevel structural equation models

Self-concept, one of the oldest constructs in psychology, is a major focus in many disciplines, an important mediating factor that facilitates the attainment of other desirable outcomes, and is central to the positive psychology revolution, which focuses on how healthy, normal, and exceptional individuals can get the most from life (e.g., Diener, 2000, Marsh and Craven, 2006, Seligman and Csikszentmihalyi, 2000). Particularly in education settings, a positive academic self-concept (ASC) is both a highly desirable goal and a means of facilitating subsequent academic achievement, academic accomplishments, and educational choice behaviors, including subject choice, coursework selection, academic persistence, and long-term educational attainment (e.g., Chen et al., 2013, Guay et al., 2004, Marsh, 1991, Marsh, 2007, Marsh and Craven, 2006, Pinxten et al., 2010). Thus, Marsh and Yeung (1997; also see Parker et al., 2012) found that ASCs predicted future coursework selection better than corresponding measures of academic achievement, whilst Marsh and O'Mara, 2008, Marsh and O'Mara, 2010 found that ASCs predicted long-term educational attainment five years after high school graduation better than school grades, IQ, standardized tests, or socio-economic status.

In the last 30 years self-concept research has experienced dramatic growth in the quality of measurement instruments, theoretical models, quantitative methodology, and research design. The cornerstone of this resurgence is the classic review article by Shavelson, Hubner, and Stanton (1976), who posited self-concept as a multidimensional, hierarchical construct, where different facets of academic self-concept are substantially correlated and form a single higher-order academic self-concept factor. This is consistent with the positive relations routinely observed among achievements in different school subjects (Marsh, 2007). However, subsequent research revealed that mathematics self-concept and verbal self-concept in particular were nearly uncorrelated; this led to the Marsh/Shavelson revision (Marsh & Shavelson, 1985). Marsh and Shavelson posited two higher-order academic self-concept factors (mathematics/academic and verbal/academic), a continuum of core academic self-concept factors ranging from verbal self-concept at one end to math self-concept at the other end, and an ordering of academic self-concepts in other domains along this continuum. This perspective has resulted in increased attention to frame of reference effects based on dimensional comparisons, which are the focus of the present investigation.

Since at least the time of William James (1890/1963), psychologists have realized that self-concepts are based in part on evaluations of objective achievements in relation to frames of reference. Depending on the frame of reference, or the standard of comparison against which individuals evaluate themselves, the same objective accomplishment can lead to quite different self-concepts. Self-concepts have important consequences for future choices, for behavior and performance. In the broader psychological literature, the two most frequently posited frames of reference are social and temporal comparisons (Albert, 1977, Festinger, 1954, Möller, 2005, Möller et al., 2009, Möller et al., 2011); self-concepts are based in part on how our current accomplishments compare with past performances (temporal comparisons) and on how they compare with the accomplishments of others within the current context (social comparisons; e.g., classmates in one's school or class). However, as emphasized by Möller and Marsh (2013, p. 1) in their empirical and theoretical foundation for dimensional comparison theory (DCT): “In addition to social comparisons (Festinger, 1954) and temporal comparisons (Albert, 1977), dimensional comparisons are presented as a largely neglected but influential process in self-evaluation.”

Möller and Marsh (2013) stress that the internal/external frame of reference (I/E) model that forms the theoretical and empirical basis of the DCT has been studied widely only in educational settings, and mostly in relation to the math and verbal domains. DCT (Marsh et al., 2014, Möller and Marsh, 2013) incorporates an extensive body of educational research based on the I/E model, placing dimensional comparisons into a broader theoretical foundation, in relation to more general psychological models of self-evaluation, person perception, frames of reference, and social comparison. Hence, the overarching purpose of the present investigation is to provide new empirical support for the extension of the I/E model into DCT.

1. The internal/external frame of reference (I/E) model and Dimensional Comparison Theory (DCT)

1.1. The internal/external frame of reference (I/E): theoretical basis and empirical support

Initially, the I/E model was developed to explain why math self-concept and verbal self-concept are nearly uncorrelated, even though achievement in the same areas are strongly correlated (typically 0.5 to 0.8, depending on how achievement is measured; for further discussion, see Marsh, 1986, Marsh, 2007). This finding led to the Marsh/Shavelson revision of the original Shavelson model. In addition to this theoretical advance, it also had practical implications for the processes underlying the formation of academic self-concept in different domains and how they are related to accomplishments in these domains.

The I/E model posits that academic self-concept (ASC) in a particular school subject is formed in relation to two frames of reference: a) an external (social comparison) reference, in which students contrast their perceived performances in a particular school subject with the perceived performances of their peers in the same school subject, and b) an internal (dimensional or ipsative comparison) reference in which students contrast their own performances in one particular school subject against their performances in different school subjects. Tests of the I/E model typically are conducted by regressing math self-concept and verbal self-concept on math and verbal achievements (see Fig. 1A). The external comparison process predicts that good math skills lead to higher math self-concepts and that good verbal skills lead to higher verbal self-concepts. According to the internal dimensional comparison process, however, good math skills lead to lower verbal self-concepts, once the positive effects of good verbal skills are controlled: The better I am at mathematics, the poorer I am at verbal subjects, relative to my good math skills. Similarly, better verbal skills lead to lower math self-concepts once the positive effects of good math skills are controlled. In models used to test these predictions (see Fig. 1A), the horizontal (matching) paths leading from math achievement to math self-concept and from verbal achievement to verbal self-concept are predicted to be substantially positive. However, the cross-paths leading from math achievement to verbal self-concept and from verbal achievement to math self-concept (the gray lines in Fig. 1A) are predicted to be negative.

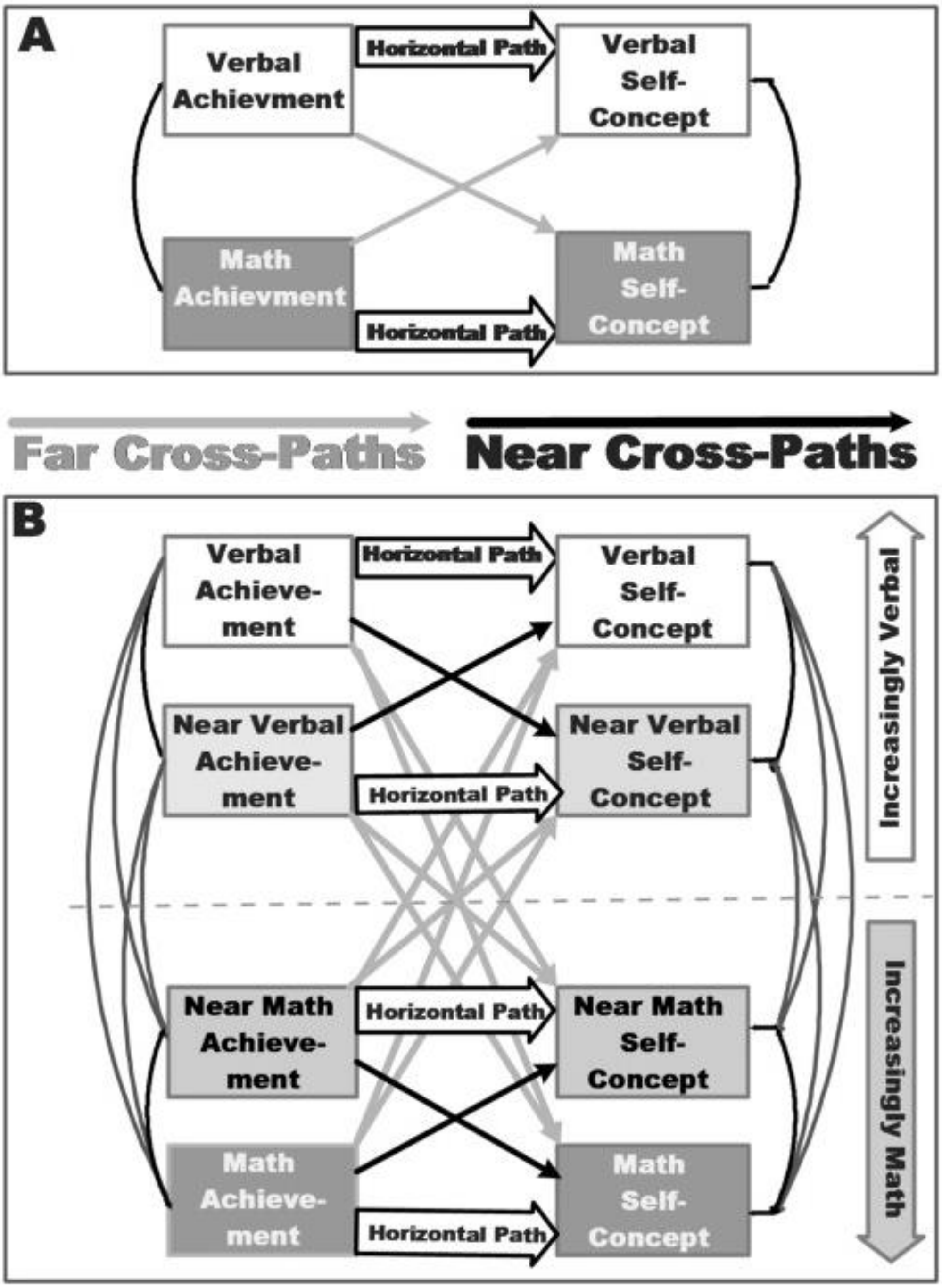

Fig. 1. (A). The 'Classic' Internal/External Frame of Reference (I/E) Model relating verbal and math achievement to verbal and math self-concept. According to predictions from the I/E model, the horizontal paths from achievement to self-concept in the matching domains (content area) are predicted to be substantial and positive, whereas the cross-paths from achievement in one domain area to self-concept in a non-matching domain are predicted to be negative (contrast). (B). Extending the I/E model to include the verbal domain, one or more verbally-related domains, the math domain, and one or more math-related domains. Far cross-paths (any of the math domains to any of the verbal domains) are again predicted to be

negative. However, the near cross-paths (relating different verbal domains to each other, or different math domains to each other) are predicted to be significantly less negative, non-significant, or even positive (assimilation).

In reviewing the literature, Möller et al. (2009; also see Marsh, 2007, Möller and Marsh, 2013) note that evidence in favor of this model for math and verbal domains comes from diverse sources and experimental designs. Marsh (1986)first proposed the I/E model based on results from 13 studies of the relations between math and verbal achievements and ASCs in the corresponding domains. Although there were consistently high correlations between math and verbal achievements (.42–.94), there were weak or even negative correlations between math self-concepts and verbal self-concepts (−.10 to .19). In the I/E path diagram (Fig. 1A), the (horizontal) paths from math achievement to math self-concept and from verbal achievement to verbal self-concept were substantial and positive. However, the cross-paths (Fig. 1A) from math achievement to verbal self-concept and from verbal achievement to math self-concept were significant and negative. In a large cross-cultural study, Marsh and Hau (2004; also see Marsh, Hau, Artelt, Baumert, & Peschar, 2006) demonstrated that support for these predictions generalized over large, nationally representative samples of 15-year-olds from 26 countries. In a meta-analysis of 69 datasets Möller et al. (2009) reported that math and verbal achievements were highly correlated (.67), but that self-concepts were nearly uncorrelated (.10). The horizontal paths from achievement to ASC in the matching domains were positive (.61 for math, .49 for verbal), but cross-paths were negative from math achievement to verbal self-concept (−.21) and verbal achievement to math self-concept (−.27). Thus, strong support for the classical I/E predictions in relation to the math and verbal domains, generalizes over different measures of achievement, over self-concept instruments, and over age, gender, and country.

Möller and colleagues (e.g., Möller, 2005, Möller and Köller, 2001a, Möller and Köller, 2001b, Möller and Savyon, 2003, Pohlmann and Möller, 2009) provided stronger tests of causal mechanisms posited in the I/E model, based on true random control trials. They experimentally manipulated the external (social) comparison process, based on performance feedback relative to other students, and the dimensional comparison process, based on feedback relative to performance by the same student on achievement from two subject-specific tasks. They showed how experimentally manipulating feedback on achievement in one subject area has an inverse effect on self-concept in a different school subject. Möller and Köller (2001a) concluded that "as shown experimentally for the first time, dimensional comparison information can have inverse effects on task-related cognitions in other domains" (p. 833). Extending empirical tests of the I/E model, they showed support both for an external comparison process (based on experimentally manipulated feedback about achievement compared to other students) and an internal comparison process (based on experimentally manipulated feedback about one's own performance in one area relative to performance in another area). In addition, Möller and Husemann (2006)conducted two introspective diary studies, which demonstrated that students spontaneously carry out dimensional comparisons in everyday life, with contrast effects from one domain to self-evaluations and emotions in the other.

The I/E model has also been heuristic in relation to other major theoretical models in psychology. For example, Pekrun, 2006, Goetz et al., 2008, Goetz et al., 2006) demonstrated that I/E predictions based on self-concept responses also generalize to emotional responses, and incorporated the I/E model into his control-value theory of achievement emotions. Similarly, Eccles and colleagues (Eccles et al., 2004, Nagy et al., 2008, Parker et al., 2012) integrated support for I/E predictions into expectancy-value theory for the prediction of gender differences in academic and career choice. However, again, these studies focused on math and verbal domains, which have been the focus of most I/E studies, rather than on a more diverse range of domains, as incorporated into the theoretical DCT extension of the I/E model.

Although some studies have explored other academic domains (e.g., Bong, 1998, Chiu, 2008, Chiu, 2012, Dickhäuser, 2003, Marsh et al., 2001, Möller et al., 2006b, Nagy et al., 2006, Xu et al., 2013, Yeung et al., 2001), these have not been situated into an over-arching theoretical framework such as that posited in DCT. Thus, Möller et al. (2009) called for an extension of the I/E model to other academic domains, asking, for example, whether: students see physics and mathematics as sufficiently distinct that better performances in one would lead to poorer self-concepts in the other (a contrast effect like that posited in the I/E model based on the mathematics and verbal domains), or would the two be seen as sufficiently similar so that better performance in one would lead to better self-concepts in the other (an assimilation effect)? (p. 1159).

2. Dimensional Comparison Theory (DCT): extending the I/E model to incorporate assimilation and contrast effects based on near and far comparisons

This article provides the first large-scale empirical test of the recently published theoretical account of DCT (Möller & Marsh, 2013; also see Marsh et al., 2013), which places the I/E model in a much broader, more general theoretical framework, incorporating a diverse range of academic domains in addition to the math and verbal domains in the classic I/E model.

The theoretical and empirical underpinning of the extension of the I/E model into DCT is based on an a priori verbal-to-math continuum of academic domains (Fig. 2). Theoretically, the continuum follows from the Marsh/Shavelson revision (Marsh, 1990a, Marsh, 1990b, Marsh and Shavelson, 1985) of the original Shavelson et al. (1976) multidimensional, hierarchical model of self-concept (see Marsh, 2007). Empirical support for the continuum is based on a higher-order factor analysis in which there are two higher-order factors (math and verbal self-concept) defined by lower-order self-concept factors, such that the similarity between different academic domains is operationalized as correlations among the different domains (Marsh, Byrne, & Shavelson, 1988).

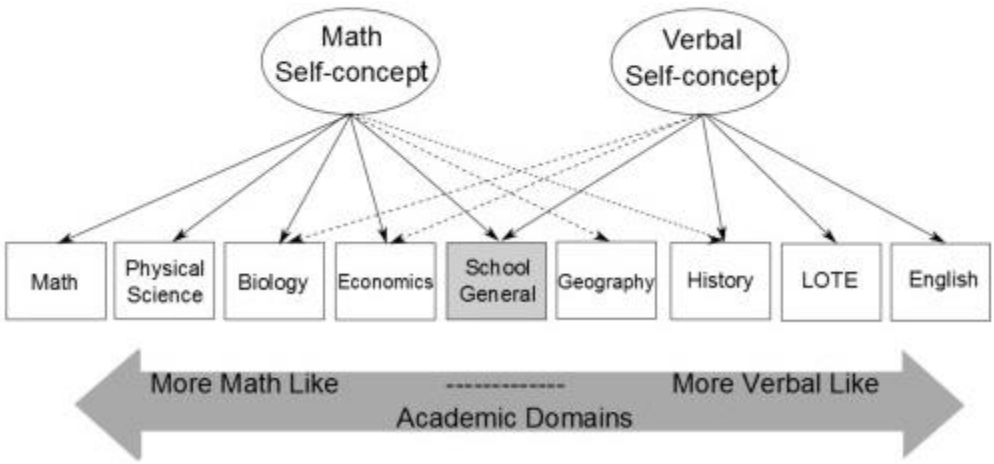

Fig. 2. The theoretical model of the structure of academic self-concept developed as part of the Marsh/Shavelson revision (Marsh & Shavelson, 1985) of the Shavelson et al. academic self-concept model (Shavelson et al., 1976), which includes a wider variety of specific

academic facets. In the context of that research, “English” refers to native language verbal domain, whereas LOTE (Language Other Than English) refers to non-native language verbal domains. Adapted with permission from 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.

The critical theoretical distinction between the I/E model and DCT is illustrated in Fig. 1. The classic I/E model (Fig. 1A) is based on only math and verbal domains, so that there are only “horizontal” paths between matching domains and “cross” paths between non-matching domains. The important contribution of DCT is to expand the scope of the theoretical perspective to include a wider variety of domains, in which the cross-paths involve “near” and “far” comparisons in relation to how similar or dissimilar different school subjects are to each other. A key feature of the expansion of the I/E model into DCT is to highlight this distinction between near and far comparisons, to test the a priori prediction that contrast effects based on “near” comparisons will be substantially less negative and may even be positive rather than negative (i.e., assimilation effects). Hence, similar subjects (e.g., native and foreign languages, or math and physics) might be seen as complementary, rather than contrasting, such that achievement in one domain contributes positively to self-concept in a complementary, near domain. The present investigation provides the first strong test of the extension of the I/E model into DCT, stronger in terms of the DCT theoretical basis, the range of different school subjects considered, and the statistical methods used.

3. A priori predictions: new tests of DCT

Because of the wide variety of school subjects, and given the range of methodological innovations introduced here, Studies 1 and 2 provide the strongest available tests of the newly formulated DCT and its extension of the classic I/E model. More specifically, here we expand on this previous research to test more explicitly the theoretical prediction that contrast effects (the negative cross-paths in Fig. 1) are maximized for “far” domains that are very distant from each other on the academic continuum (e.g., verbal and math; Fig. 2). However, for “near” comparisons, based on academic domains that are close to each other on the academic self-concept continuum (e.g., native language and foreign language; see Fig. 2), the cross-paths are predicted to be substantially less negative (weaker contrast) or even positive (assimilation). Based on this theoretical extension of the traditional I/E model, in the present investigation we hypothesize that:

•Matching paths from achievement to self-concept in the same domain (e.g., math achievement to math self-concept) will be substantial and positive;

For models with only math and verbal self-concepts, the I/E predictions will be replicated—substantially positive paths for matching domains, negative (contrast) cross-paths for non-matching domains (i.e., math achievement → verbal self-concept, verbal achievement → math self-concept)

In the extended DCT model, which includes a more diverse selection of school subjects (five in Study 1, six in Study 2), cross-paths relating “far” domains (math with English and German) will be significantly negative (contrast effect)

In the extended DCT model, cross-paths relating “near” domains (e.g., English and German) will be significantly less negative than paths based on far comparisons. We leave as a research question whether near paths will remain significantly negative (contrast, but less so than for far paths), non-significant, or even positive (assimilation effects).

4. Study 1: new tests of DCT predictions

Based on the continuum in Fig. 2, we grouped the subjects into two categories—verbal subjects (German, English, and second foreign language) and math/science subjects (mathematics and biology). Hence, paths leading from each achievement to a self-concept can be classified as **matching** when the achievement and self-concept domains are the same (i.e., math achievement to math self-concept), **near** when the path from achievement in one domain to a different (non-matching) ASC is in the same category (e.g., German achievement to English self-concept; math achievement to biology self-concept), or **far**, for paths from achievement in one domain to non-matching ASCs in a different category (e.g., German achievement to math self-concept; math achievement to German self-concept).

4.1. Methods

4.1.1. Participants

The sample consisted of 1140 German students (54.1% female) in 51 classrooms in 6 public, high ability track high schools (the German *Gymnasiums*). The students were enrolled in Grade 7s (59%) and 10 (41%), and ages varied from 12 to 18 years (M = 14.24, SD = 1.6). Participation was voluntary and required consent from parents. The participation rate was above 90% in all classes. The study took place in the classroom near the end of the school year and lasted 45 min.

4.1.2. Measures

Academic self-concept was assessed with three items (e.g., “I am good in English [mathematics/German/Second Foreign Language/biology]”). Strictly parallel wording was used for all domain-specific items. That is, the items for English, mathematics, German, second foreign language, and biology were exactly the same, except for the name of the subject. A 4-point Likert-type scale (1 = *completely disagree*, 4 = *completely agree*) was used for all items. Coefficient alpha estimates of internal consistency were satisfactory for all five academic domains ($.79 < \alpha < .85$). The grades awarded on the midterm report card were used as an indicator of prior achievement each of the school subjects. Grades were coded such that high scores indicate positive learning outcomes. In addition, four background/demographic variables were collected to serve as control variables: home language (German, other); Gender, student age, year in school (Year 7 or Year 10).

4.1.3. Innovative statistical analyses

Analyses conducted with Mplus 7.11 (Muthén & Muthén, 2008–2013) consisted of CFA and SEMs based on the Mplus robust maximum likelihood estimator (MLR), with standard errors and tests of fit that were robust in relation to non-normality of observations, and a design-based correction to control for the non-independence of observations (Muthén & Muthén, 2008–2013). Relatively small amounts of missing data (average coverage = 94%) were handled with the full information maximum likelihood (FIML) procedure implemented in Mplus. Goodness of fit was evaluated with traditional fit indices that are relatively independent of sample size (e.g., Hu and Bentler, 1999, Marsh et al., 1988a): Root Mean Square Error of Approximation (RMSEA), the Tucker–Lewis Index (TLI), and the Comparative Fit Index (CFI)—as well as a critical evaluation of parameter estimates in relation to a priori predictions.

Innovative tests of the I/E and DCT predictions were based on stronger statistical methodology than has been used in previous I/E research, and that also have broader application in educational and social science research more generally. In the evaluation of predictions for near, far and matching domains, we relied heavily on the flexibility of the "model constraint" function in Mplus, and the resulting tests of statistical significance based on these model constraints. In particular, we used these constraints to compare the mean of matching, near, and far paths across the five school subjects in relation to a priori predictions. In each of the models considered here, we included all variables, noting that a model with all variables simply correlated is equivalent (in terms of df and goodness of fit) to a model in which some of the correlations are represented as path coefficients. Thus for example, in a preliminary model based on the classic I/E model (Fig. 2), relations among German and math constructs were presented by appropriate path coefficients, whereas all other relations were represented as correlations. In the subsequent extended DCT model, involving all five domains, additional correlations were represented by appropriate paths in the SEM. Similarly, the set of control covariates was included in all models. However, for models that did not control for these effects, relations between covariates and other constructs were simply represented as correlations, whereas models designed to control for these covariates included these relations as paths in the SEM. Using this approach, all the different models incorporated the same variables and were equivalent in terms of df and goodness of fit. This strategy had important advantages for the comparison of models based on different sets of variables that potentially confound aspects of the measurement and structural models. This evolving methodology—combining the flexibility typically associated with analyses of manifest variables with the rigor of latent-variable models—is apparently a new contribution of the present investigation (see Supplemental Materials).

5. Results: study 1

5.1. Preliminary analyses: support for construct validity of the measures

In our preliminary analyses (presented more fully in the Supplemental Materials), we evaluated a series of CFAs to test the a priori structure underlying the multiple domains of ASC and how they relate to parallel measures of achievement and control variables. A critical feature of these data is that five sets of three self-concept items with parallel wording were used to measure each self-concept factor. Consistent with expectations, the CFA model with no control for parallel worded items was highly unsatisfactory in terms of goodness of fit, but when item method factors were included in the model, to account for parallel item wording effects, the fit was much better (Supplemental Materials Table 1). Finally, in model M3, we found strong support for the invariance of factor loadings of parallel items across the five domains. Although support for the invariance of factor loadings across the five self-concept domains is not essential for the present purposes, it facilitates interpretation of results in a substantially more parsimonious model (i.e., with fewer estimated parameters). Correlations among all the constructs (five self-concept domains, parallel measures of achievement, and control covariates) are presented in Table 1. Of particular relevance are the substantial correlations between each latent self-concept factor and the corresponding measure of achievement. In support of the convergent validity of the self-concept factors, correlations between each ASC factor and the matching area of achievement were consistently substantial (convergent validities, *r*s varying from .545 to .695; Table 1). In support of their discriminant validity, correlations between each ASC factor and non-matching areas of achievement were substantially smaller (rs vary from .096 to .322; Table 1) than the convergent validities. Consistent with previous research, correlations among the ASC factors (*r*s varying from .021 to .344) were systematically smaller than correlations among the five corresponding achievement measures (*r*s varying from .423 to .566). Consistent with the ordering of the self-concept factors along a verbal-math continuum, the correlation between math and German self-concept was close to zero (r = .021), and was the smallest of any correlation among these factors. Indeed, the three smallest correlations among ASC factors were for relations of math with German (.021), English (.085), and a second (non-native) foreign language (.041). However, German (native language) self-concept was more highly correlated with the second foreign language (r = .344) than with English, and the correlation between the two foreign languages was smaller (r = .255). Also, biology was as highly correlated with German (r = .258) as it was with mathematics (r = .257).

Although this matter is of tangential interest to the present investigation, control covariates were not substantially correlated with either self-concept factors (rs vary from −.089 to .212) or achievement scores (−.217 to .107). The pattern of gender differences is largely consistent with previous research (Crain, 1996, Eccles et al., 1999, Eccles et al., 1993, Marsh, 1989), in that boys had lower achievement scores than girls for all but maths (where differences were close to zero), but higher self-concept scores in math and biology; girls had higher self-concepts in German and English.

In summary, the factor structure underlying the self-concept responses is well-defined and is consistent with a priori expectations. The latent variable approach used here is critical in terms of taking into account the complex measurement structure. Correlations between self-concept and achievement scores provide strong support for the convergent and discriminant validity of the responses. The final measurement model (CFA-M3 in Supplemental Materials Table 1) is the basis of the subsequent structural equation models (SEMs), presented in the next section.

6. The classic I/E model and its extension based on DCT: assimilation and contrast effects

The SEM used here is equivalent (in terms of df and goodness-of-fit) to the CFA measurement model (see Supplemental Materials). The only difference is that relations between some of the constructs (particularly relations between achievement and self-concept) were replaced with path coefficients in order to test a priori predictions based on the I/E model and its extension in DCT. We begin by briefly testing the replicability of the classic I/E (Fig. 1A) based only on paths for the critical verbal (German-native language) and math domains; this is a logical prerequisite to testing the more complicated extension including all five domains, and also provides an advanced organizer to the subsequent analyses.

6.1. Classic I/E model

In tests of the “classic” I/E model (Fig. 1A) we considered two models (Table 2A), with and without controlling the effects of the four control variables (see earlier discussion of Table 1). For the present purposes we classified the critical path coefficients as matching (relating achievement to the matching area of self-concept; the “horizontal” paths in Fig. 1) and “far” comparisons (German achievement to math self-concept, Math achievement to German self-concept; “cross-paths” in Fig. 1A).

There is strong support for the replicability of the I/E model, in that: matching coefficients were consistently substantial and positive (Mn = .656, SE = .031 Table 2, no control); coefficients representing the far comparisons were consistently negative (Mn = −.162, SE = .023); and the difference between matching and far comparisons is highly significant (Mn Diff = .818, SE = .036). When controls for the covariates were added, the matching coefficients became slightly less positive (Mns of .649 vs .656), and the far coefficients became slightly less negative (Mns of −.135 vs −.162), but the pattern of results and statistical significance was unchanged.

In summary, the results support the replicability of the large volume of research based on the I/E model and its robustness in relation to the control variables considered here. We note that in the language of DCT, the negative effects for the far comparisons were contrast effects—higher achievement scores in one domain result in lower self-concept in the contrasting self-concept domain. Extending this model in the next section, we ask if the size or even direction of these contrast effects were consistent for "near" comparisons between more closely related domains that might be seen as complementary.

6.2. DCT extension to include five domains

In the extended SEM model we included (25 = 5 × 5) paths leading from each of the five ACH domains to each of the five latent ASC factors (see Table 2). These 25 path coefficients were classified a priori as matching, near and far comparisons, based on the theoretical ordering of ASC factors (Fig. 2). For the present purposes, path coefficients involving any pair of verbal-related constructs (i.e., domains near the verbal end of the continuum; German, English, and second foreign language in this study) or between the two more math/science constructs (i.e., domains near the math end of the continuum; biology and math in this study) constructs were considered to be "near" comparisons, whilst comparisons between achievement from one group to self-concept in the other group were considered "far" comparisons (see Fig. 1B). In juxtaposition to the classic I/E model, the critical new feature is the distinction between near and far comparisons.

The mean across the 25 path coefficients (Mn = .097, SE = .011, Table 2, no control) represents substantially positive coefficients for the five matching comparisons (Mn = .700, SE = .026) and modestly negative coefficients for the remaining 20 non-matching comparisons (Mn = -.054, SE = .010). For the present purposes, the critical feature of these non-matching coefficients is that the mean of the 8 near coefficients is not significantly different from zero (Mn = −.014, SE = .017), whilst the mean of the 12 "far" coefficients is significantly negative (Mn = −.080, SE = .010); the difference between the two is highly significant (Mn = .066, SE = .017). Again, the results were altered somewhat when controls for the set of four covariates were included (Mns = .689 for matching, −.002 for near, −.073 for far comparisons) but the pattern of results remains unchanged.

It is also interesting to compare the results for classic I/E (Table 2) and extended DCT (Table 2) models. In particular, the path coefficients for the two domains common to each model (German and math) were very similar. Hence, the larger mean matching paths in the extended model are due to large matching paths particularly for English (.787) and for second foreign language (.733). Similarly, the less negative mean far comparison in the extended model is due to the inclusion of new domains. In particular, the posited negative effect of biology achievement was non-significant for German self-concept (.013, SE = .044) and English self-concept (−.010, SE = .037), but was significant for self-concept in the second foreign language (−.107, SE = .040). Similarly, the effects of German and English achievement were not significant for biology self-concept (−.024, SE = .036; -.046, SE = .037 respectively), whilst the effect of second foreign language achievement on biology was significantly negative (−.117, SE = .045). This is also consistent with the earlier discussion of correlations among self-concept factors (Table 1), showing that biology self-concept is as highly related to German self-concept as it is to math self-concept (rs = .258 and .257, respectively), and that biology self-concept is more highly related to German and English achievements (rs = .243 and .242, respectively) than to math achievement (.194). These observations suggest, perhaps, that the classification of biology as a more math/science-like domain might be unwarranted—or at least that biology is no more complementary in relation to math than it is to verbal subjects (see subsequent, related discussion of Study 2, which includes physics as well as math and biology).

In summary, consistent with a priori predictions for the extended DCT model, the contrast effects that are such a salient feature of the classic I/E model, were largely limited to far comparisons in the extended DCT model. We note, however, that there were no instances of statistically significant assimilation effects (i.e., positive effects of near comparisons, the effect of achievement on a complementary self-concept domain).

7. Study 2: inclusion of more domains for ACH, ASC and self-efficacy responses

The purpose of Study 2 was to replicate and extend support for theoretical predictions based on DCT. The extension involved two important features. First, the set of academic domains was extended to include six subjects, including three that are classified as more math/science like (math, physics, biology) and three that are classified as more verbal like (German, English, History). The inclusion of physics provided a potentially important addition in that the results of Study 1 suggest that biology is seen as being as similar to verbal domains as it is to math domains, whilst previous research (e.g., Goetz et al., 2010, Möller et al., 2009) suggests that math and physics are more likely to be complementary. Also, the addition of history is a potentially interesting addition in that, whilst it is classified as more verbal-like, it is clearly neither a purely verbal subject nor a math/science subject, but falls near the center of the math-verbal continuum. Nevertheless, a priori predictions and their tests closely parallel those in Study 1.

The second component expands the extension of the I/E model into DCT by juxtaposing frame of reference effects for academic self-efficacy and self-concept responses. Although this distinction has been a critical issue for the theoretical basis of the I/E model (e.g., Marsh et al., 1991, Möller et al., 2009, Skaalvik and Skaalvik, 2002; see also, Marsh et al., in press), it has not been tested in relation to DCT. We begin with a review of theoretical and empirical research into the similarities and differences between self-concept and self-efficacy, and how these relate to the I/E model and its extension into DCT.

7.1. Self-concept vs self-efficacy: similarities and differences in relation to frame of reference effects

In the social sciences, jingle-jangle fallacies (Marsh, 1994, Marsh et al., 2003) are common. Jingle–jingle fallacies occur when similar names are given to psychological constructs that, in reality, differ from each other; or when psychological constructs that are in reality similar are given different names. This problem is potentially worrisome for the plethora of self-belief constructs. Self-beliefs predict important outcomes across a range of life domains (Bandura, 1986, Bandura, 2006a, Bandura, 2006b, Marsh, 2007), but particularly in educational research, the focus has been primarily on how academic achievement is related to academic self-concept and self-efficacy.

Similarities. As emphasized by many researchers (e.g., Bong & Skaalvik, 2003) both self-concept and self-efficacy are similar in their emphasis on perceived competence, in having a multidimensional and hierarchical structure and in prediction of future performance and choice. Historically it has been argued that self-concept is a global measure, whilst self-efficacy is more domain specific. However, particularly in educational settings, current approaches to self-concept typically focus on its domain specificity, whilst some researchers argue for the usefulness of global measures of self-efficacy (Chen et al., 2001, Jerusalem and Schwarzer, 1992, Schwarzer and Jerusalem, 1995, Skinner et al., 1988; also see Bandura, 2006a, Bandura, 2006b). Importantly, research clearly suggests that domain specific measures of self-efficacy and self-concept are more strongly correlated with criterion variables than are domain general versions of the same measure (Pajares & Schunk, 2005). Indeed, the link between general self-beliefs and achievement has been shown to be weak (e.g., Marsh & O'Mara, 2008).

7.1.1. A critical difference: criterion references vs normative (dimensional and social) frames of reference

Depending on how self-concept and self-efficacy constructs are measured, there are important differences between these constructs, in relation to the criteria used to evaluate success. For the present purposes, the critical distinction is that self-efficacy measures assess self-perceptions of capabilities in relation to clearly explicit, or at least implicit, criteria of success that are incorporated into the actual wording of the self-efficacy items. Thus, self-efficacy items typically are criterion-referenced and assess students' confidence in being able to solve specific problems similar to the actual test items (Marsh et al., 1997, Pajares and Schunk, 2005), or self-perceptions of their capability to successfully undertake the actions required to complete a specific task (e.g., "I would be able to calculate the area of a room in square meters"; Ferla, Valcke, & Cai, 2009).

Self-concept, on the other hand, is evaluative, relating to judgments about whether one's behavior matches one's own standards of worthiness and competence (e.g., "I am good at mathematics"; Marsh et al., 1991; also see Bong & Skaalvik, 2003). Because self-concept items typically do not contain a specific criterion-reference, students have to rely on other frames of reference, like the social and dimensional comparison processes posited in the I/E model. On this theoretical basis, Marsh (1993) argued that frame-of-reference effects are directly implicated in ASC responses, as individuals use the social and dimensional comparison processes to establish frames of reference for evaluating their own performances. Making a similar point, Bong and Skaalvik (2003) argued that "assessing one's capability in academic self-concept relies heavily on social comparison information," whereas "self-efficacy items solicit goal referenced evaluations and do not directly ask students to compare their abilities to those of others," and "provide respondents with a specific description of required performance as a referent against which to appraise their competence" (p. 9). Similarly, Bandura (1986) emphasized that self-esteem and self-concept— but not self-efficacy— are partly determined by "how well one's behavior matches personal standards of worthiness" (p. 410). However, Marsh et al. (1991) argued that the self-perceived worthiness of performance expectations in relation to personal and external standards is critical in motivating and sustaining performance. Ignoring the self-evaluative component of self-beliefs is likely to greatly reduce the power of self-beliefs in motivating students and predicting future choice behaviors and educational outcomes.

The relevance of this criterion vs. normative distinction has been highlighted in a series of studies testing the I/E model (Marsh et al., 1991; also see Skaalvik and Rankin, 1995, Skaalvik and Skaalvik, 2002) that adhered very closely to highly prescriptive approaches to the construction of self-efficacy items. Thus, students were presented with actual items from an achievement test and asked how likely it was they would be able to solve this item if it were presented in a test. They were then presented with a test based on the same or similar items, to assess self-efficacy (see Marsh et al., 1991, Skaalvik and Rankin, 1990, Marsh et al., 1997). Not surprisingly, particularly when both self-efficacy and test items were identical, self-efficacy responses were highly correlated with test scores. Of particular relevance is an early test of the I/E model by Skaalvik and Rankin (1990), who were not able to replicate support for the I/E predictions. However, Marsh et al. (1991) demonstrated that the failure to replicate was based on self-efficacy rather than on

self-concept responses. They showed, in a study based on students who completed both instruments, that I/E predictions were supported for ASC responses, but not for self-efficacy—a finding subsequently replicated by Skaalvik and Rankin (1995).
In exploring the implications of their results, Marsh et al. suggested that test-specific self-efficacy ratings severely truncated the operation of frame-of-reference effects and of evaluative responses to self-perceptions. For example, in judging the likelihood of correctly answering a test item on a math test, students have little need to compare their performance in one subject with their performance in other subjects such as English (internal, dimensional comparison) or with the performances of other students in mathematics (external, social comparison). Furthermore, a response of 80% likelihood of successfully answering a particular math problem does not indicate whether the student considers this to be a good or bad outcome in relation to internal or external frames of reference; this potentially undermines its usefulness in predicting future behavior. Based on these findings, Marsh et al. hypothesized that self-concept responses should be effective at predicting future choice behavior, and proposed the selection of specific school subjects as a particularly relevant area in which to evaluate these predictions. In support of these suggestions, Skaalvik and Rankin (1995) found that math and verbal self-concepts influenced corresponding measures of intrinsic motivation, effort, and anxiety, whereas math and verbal self-efficacy responses did not.

7.1.2. Differences between self-efficacy and self-concept responses: discriminant validity

Research has also revealed an additional critical difference between self-concept and self-efficacy constructs that is central to the I/E model and its extension to DCT. The initial basis of the I/E model was to posit internal, dimensional comparison processes to explain why math and verbal self-concepts were nearly uncorrelated, even though math and verbal achievements were highly correlated. However, because dimensional comparison processes are not implicated in self-efficacy measures, correlations between appropriately constructed math and verbal self-efficacy factors are predicted to be substantially correlated, similarly to the substantial correlations between math and verbal achievements. There now exists good support for these predictions (e.g., Marsh et al., 1991, Skaalvik and Rankin, 1990, Skaalvik and Rankin, 1995; also see Bong, 1998).

In their meta-analysis of I/E studies, Möller et al. (2009) found that the correlation between math and verbal self-efficacy measures (r = .50) was much higher than the correlation between self-concept measures (between r = −.09 and .17), and nearly as high as the corresponding correlation between math and verbal achievement correlation (r = .70). Commenting on ongoing debates about how self-concept and self-efficacy are impacted by different frames of reference, they concluded that:

> the I/E model is valid for self-concept but not for self-efficacy beliefs. Our results show that, like achievements, self-efficacy beliefs in math and verbal domains show significant positive correlations. In other words, dimensional comparisons do not seem to play an important role in the development of self-efficacy beliefs and, logically, the I/E model does not fit the data (p. 1158).

Valentine, DuBois, and Cooper (2004) provided a large scale meta-analysis of the effects of self-efficacy and self-concept on academic achievement in a matching domain—corresponding to the horizontal paths in the I/E model (Fig. 2). Importantly, they compared how achievement is related to self-efficacy and self-concept, but found no statistically significant difference between relations based on these two self-belief constructs. However, particularly pertinent to the present investigation, most of the studies in both these meta-analyses were based on either self-efficacy or on self-concept, but did not directly compare results based on both constructs in the same study. Consequently, it is not clear whether these differences are due to confounding with the many other differences between studies based on each of these constructs.

7.1.3. Self-concept and self-efficacy responses: jingle-jangle fallacies

Marsh et al. (2003) argued that support for predictions based on a priori distinctions between self-concept and self-efficacy responses is highly dependent upon the precise nature of the self-efficacy items. In particular, many so-called self-efficacy items are nearly indistinguishable from traditional self-concept items, in relation to the critical distinctions emphasized here (Marsh, Trautwein, Lüdtke, & Köller, 2008). Where explicit criteria of success are not provided as part of the item, this forces students to use additional standards to answer the item, such as the social and dimensional comparison processes that are the basis of DCT. Thus, some measures purporting to assess self-efficacy are based on stimuli likely to invoke social and dimensional comparison processes (e.g., "I'm certain I can do an excellent job on assignments and tests," where the term *excellent*might imply a comparison with the work of other students or, perhaps, the student's own performance in other school subjects). Indeed, particularly in relation to more global measures of self-efficacy, it is unclear whether the inclusion of explicit criteria of success is possible or even appropriate, further blurring the distinction between self-concept and self-efficacy. In summary, consistent with jingle-jangle fallacies, the empirically demonstrated distinction between self-concept and self-efficacy responses is likely to depend on the nature and wording of the items, rather than on the label assigned to the construct.

8. The present investigation: A priori predictions and new tests of DCT (study 2)

In Study 2 we evaluated a priori predictions and new tests of DCT in relation to a broader range of academic concept domains than have been previously considered (even in Study 1; see earlier discussion), and in relation to measures of homework self-efficacy. Although these self-efficacy items focused on self-perceived capability to successfully complete homework in each school subject domain, rather than on achievement per se, the construction of these items qualifies as self-efficacy items. For example, consider the item "If it is difficult to solve homework in physics, I know where I have to read to find the right solution." The criterion behavior is knowing where to read in order to successfully complete a homework assignment in physics. Importantly, there is no explicit or even implicit message that students have to invoke comparisons with other students (social comparison) or with their own accomplishments in other school subjects (dimensional comparison) to respond to this item.

Following from prior research on the juxtaposition of self-concept and self-efficacy responses in relation to the I/E model, the extension of the I/E model in DCT, and the nature of the stimuli used here, we offer tentative a priori predictions for self-efficacy responses (noting again that a priori predictions based on ASC responses are essentially the same as in Study 1—see earlier discussion).

- Correlations between achievement and homework self-efficacy in the same domain (e.g., math achievement to math homework self-efficacy) will be substantial and positive. Nevertheless, because the self-efficacy responses are in relation to capabilities of successful completion of homework in each domain, rather than underlying ability in each domain, we speculate that these correlations will be smaller for self-efficacy than for self-concept responses;

- Although the factor structures based on both self-concept and self-efficacy responses are predicted to be well defined, it is anticipated that the correlations among the self-concept factors (due in part to the internal comparison process) will be substantially smaller than for the self-efficacy responses. Thus, self-concept responses will be better differentiated than the corresponding self-efficacy responses in relation to correlations among domains within each construct and in relation to achievement.
- For models with only math and verbal domains, support for the generalizability of the I/E predictions would require substantially positive paths from achievement to self-concept for matching domains and negative (contrast) cross-paths for non-matching domains. Nevertheless, both the positive matching and the negative cross-paths are expected to be substantially smaller for self-efficacy responses. Indeed, Marsh et al.'s (1991) results suggest that cross-paths for self-efficacy might show little or no contextual effects at all.
- When additional school subjects are added to the model, support for the generalizability of the I/E predictions would still require cross-paths relating achievement to self-concept to be significantly negative (contrast effect) for far domains (math with English and German). However, the new DCT predictions would require cross-paths relating achievement to self-concept to achievement in near domains (e.g., math with physics) be significantly less neative (less contrast) or even positive (assimilation). Because there is neither a theoretical nor an empirical basis for predictions based on near paths in self-efficacy responses, we leave this as a research question.

8.1. Methods (study 2)

Sample. The sample, 511 students (53.0% female, Mn age = 14.7 years) from Grades 8 (51.5%) and 9 (48.5%), are described in more detail by Trautwein and Lüdtke (2009; see also Trautwein & Lüdtke, 2007). Within each class, students were randomly assigned to participate in either the present study or a second, unrelated study; an average of 12.17 students per class participated in the present study. Students were broadly representative of the German school population in terms of home background. Although participation was voluntary and required consent from parents, the participation rate was above 90% in all classes. As is typical of German schools, the students in each class was the same across the different subjects, most of which were taught by different teachers.

All students were enrolled in one of the three major school tracks in Germany: 47.0% in the academic-track *Gymnasium*, 36.8% in the intermediate-track *Realschule*, and 16.2% in the lower track *Hauptschule*. However, teachers are assigned to schools by the state government, and a statewide mandatory curriculum has been implemented in all schools, so that it is similar across different school types. The federal states further regulated the maximum length, educational goals, and scope of homework assignments in all school types. These regulations define the primary aim of homework assignments to be the improvement of skills, and a secondary goal to be the development of appropriate study skills and attitudes.

As in Study 1, Study 2 took place in the classroom near the end of the school year.

8.1.1. Measures

Achievement measures used as an indicator of prior achievement in each of the six domains were based on grades awarded on the midterm report card. Grades were coded, such that high scores indicate positive learning outcomes. Test materials consisted of a 7-

min standardized test of basic cognitive abilities (administered at the beginning of the session) and a questionnaire adapted from previous research (see Trautwein, Lüdtke, Schnyder, & Niggli, 2006) that was administered near the end of the school year. Strictly parallel wording was used for all domain-specific items: that is, the items for German, English, history, biology, mathematics, and physics were exactly the same, except for the name of the subject. A 4-point Likert-type scale (1 = *completely disagree,* 4 = *completely agree*) was used for all multi-item constructs.

A total of seven self-report items was used to assess self-concept and homework self-efficacy in each of the six domains. Based on a priori design and preliminary analyses, two of the items were found to be reasonably pure measures of self-concept in each domain: [(1) I am good in (subject); (2) No one can do everything. For (subject), I am just not gifted (reverse scored)].

Similarly, three items were found to be reasonably pure measures of homework self-efficacy: (1) If it is difficult to solve homework in (subject), I know where I have to read to find the right solution; (2) If I want, I always find a way to solve the homework in (subject); (3) If I make an effort, I can solve all the homework in (subject).

However, one additional homework item [With homework in (subject) I often feel completely lost. reverse scored] cross-loaded on both self-concept and homework self-efficacy scales for each of the six domains. This finding is reasonable, in that this item is more a homework self-concept item than a self-efficacy item, and is qualitatively different from the other self-efficacy items. Particularly as each of the self-concept scales consisted of only two items, whereas scales should ideally consist of at least three items, it was decided to retain this last item as an indicator both of self-concept and of self-efficacy (i.e., to allow it to cross-load on both factors for each of the six domains). In the final model (see subsequent discussion of M3 in Table 4, in which factor loadings are invariant across each of the six domains), this item loaded .45 on the self-concept factors and .24 on the self-efficacy factors. In summary, students scoring high on the self-concept items had positive self-beliefs about their ability in each subject, whilst those scoring high on homework self-efficacy had positive self-beliefs about their capability to successfully complete assigned homework tasks, even if these tasks were complicated.

In addition, background/demographic variables were collected, to serve as control variables: Gender, student age, year in school (Year 8 or Year 9), school-track (three tracks were represented by two dichotomous dummy variables), and cognitive abilities. The cognitive abilities measure was based on the Figure Analogies subscale from the Cognitive Abilities Test 4–12_R (Heller & Perleth, 2000), a German version of the CogAT by Thorndike and Hagen (1993). The internal consistency (Kuder–Richardson Formula 20) of the cognitive ability test was .92. Because the Figural Analogies is considered to be a test of reasoning that is relatively free of environmental effects, it is clearly different from our other measures of achievement based on school marks.

8.1.2. Analysis

Data analysis in Study 2 closely followed that of Study 1, based on Mplus's robust maximum likelihood estimator, a design-based correction to control for the hierarchical clustering of students, the “model constraint” function, FIML for missing data, and models similar to those in Study 1.

9. Results: study 2

9.1. Preliminary analyses: support for construct validity of measures

Following the logic of Study 1, we evaluated the a priori structure underlying the six multiple domains of ASC and homework self-efficacy, and how they relate to parallel measures of achievement and control variables. Critical features were again the parallel wording used to measure each self-concept and self-efficacy factor; the likely item-specific method effects as well as possible invariance across the self-concept and self-efficacy factors. Consistent with expectations, the CFA model with no control for parallel worded items resulted in a poor fit (based on traditional goodness of fit indices—Comparative fit index; Tucker–Lewis Index; Root Mean Square Error of Approximation), but when item method factors were included to control for the parallel wording of items, the fit was much better (see Supplemental Materials for Study 2, Table 3). In the final model, we found strong support for the invariance of factor loadings of parallel items across the six domains, for both self-concept and self-efficacy factors (see Supplemental Materials).

9.1.1. Control Variables

Again, the correlations involving control covariates (cognitive abilities, Gender, Year in school, academic track) are somewhat tangential to the present investigation, but of potential interest in their own right. Although covariates are not substantially correlated with any of the self-concept, self-efficacy, and achievement factors (*r*s varying from −.286 to .370), many of the largest relations involve gender: across self-concept, self-efficacy and achievement, girls had consistently higher scores for German and lower scores for math and particularly physics. Girls also had better achievement and self-concepts in German and biology, but these differences were not evident in homework self-efficacy factors.
There is also an interesting juxtaposition of cognitive abilities and academic track. Not surprisingly, compared to middle track students, high track students had substantially higher cognitive abilities (r = .349), whilst low-track students had lower cognitive abilities (r = −.296). However, each of these three variables had surprisingly modest correlations with the self-concept, self-efficacy, and achievement factors (rs = −.129 to .239). Indeed, the strongest relations tended to be between homework self-efficacy and track variables—positive for high track and negative for low track. The modest correlations involving cognitive abilities apparently reflected frame of reference effects in which students formed self-beliefs in relation to classmates in their same track, whilst teachers tended to grade on a curve, such that differences in grade distributions do not differ substantially as a function of track.

Latent correlations (Table 3) among all the constructs (6 self-concept factors, 6 homework self-efficacy factors, 6 corresponding measures of achievement, and control covariates) are based on the final CFA model (presented in the Supplemental Materials). Of particular relevance are the substantial correlations for matching domains of self-concept, self-efficacy, and the corresponding measure of achievement. In support of the convergent validity of the self-concept factors, convergent validity correlations were consistently substantial for matching domains of self-efficacy (*r*s varying from .502 to .699; Table 3) and for matching domains of achievement (*r*s varying from .517 to .636; Table 3). In support of their discriminant validity, these convergent validities were consistently much larger than correlations among the self-concept factors (*r*s varying from −.134 to .299; Table 3) and correlations between self-concept and non-matching achievements (*r*s varying from −.141 to .285; Table 4). Although there was also support for the convergent validity of homework self-efficacy in relation to achievement (*r*s varying from .250 to .316; Table 3), these convergent validities were still substantially higher than correlations of self-efficacy to non-matching areas of achievement (*r*s varying from −.153 to .204; Table 3). In summary, the results provide good support for the convergent and discriminant validity of self-concept and self-efficacy factors in relation to each other and to achievement.

9.1.2. Construct Validity of Domain-Specific Self-Belief Constructs

Consistent with the ordering of the self-concept factors along a verbal-math continuum, the four correlations relating math and physics self-concept to German and English self-concepts were all negative (*r*s varying from −.223 to −.096), whilst the most positive correlations were between math and physics (.451) and between German and English (.319). Correlations involving history and biology (*r*s varying from .063 to .264) consistently fell between these two sets of correlations. However, biology self-concept was more highly correlated with German, English and History (*r*s varying from .190 to .264) than with physics (.160) and math (.076). History self-concept was as highly correlated with physics (.250) and biology (.198) as with English (.068) or German (.208). Although correlations among achievement scores, and particularly among self-efficacy factors, were systematically larger (i.e., less distinct) than for self-concept factors, the pattern of results was similar for the two sets of constructs (i.e., low correlations relating math and physics to German and English, high correlations between math and physics, and between German and English). For both achievement and self-efficacy, biology tended to be correlated as high, or higher, with German, English and History, as it was with math and physics, whereas achievement in history was more correlated with biology, physics and math than with German or English.

9.1.3. Summary

In summary, the factor structure underlying the self-concept and homework self-efficacy responses is well-defined and consistent with a priori expectations. The latent variable approach used here was critical in terms of taking into account the complex measurement structure and the cross-loadings of the item, reflecting both self-concept and self-efficacy. Correlations between self-concept, self-efficacy, and achievement scores provide strong support for the convergent and discriminant validity of the responses to all three sets of constructs. However, in relation to the underlying verbal-math continuum (Fig. 2), these correlations suggest that biology might not be more math-like than verbal-like, and that history might not be more math/science-like than verbal-like. We return to this issue later. The final measurement model (described in detail in the Supplemental Materials) is the basis of the subsequent structural equation models (SEMs) presented in the next section.

9.2. Classic I/E model and its extension based on DCT: assimilation and contrast effects

Again, we note that the SEM used here is equivalent (in terms of df and goodness-of-fit) to the final CFA measurement model, except that relations between some of the constructs (particularly those relating achievement to self-concept and self-efficacy) are replaced with path coefficients in order to test a priori predictions based on the I/E model and its extension in DCT.

9.2.1. Classic I/E model

Again, we begin by briefly testing the replicability of the classic I/E model (Fig. 1A) based on only the verbal (German) and math domains. We considered four separate models which test predictions based on self-concept and homework self-efficacy responses, with and without controlling the effects of the five control variables (see earlier discussion of Table 3). For ASC responses, there is strong support for I/E predictions, in that matching coefficients are consistently substantial and positive (Mn = .614, SE = .041 Table 4A, no control), whereas far comparisons are consistently negative (Mn = −.143, SE = .036). When controls for the covariates are added, the matching coefficients become slightly less positive (Mns of .567 vs .614), and the far coefficients become slightly less negative (Mns of −.123 vs −.143), but the pattern of results and statistical significance is unchanged.

For homework self-efficacy responses, there is also reasonable support for I/E predictions, in that matching coefficients are consistently positive (Mn = .319, SE = .040 Table 4, no control) but far comparisons are negative (Mn = −.088, SE = .037). When controls for the covariates are added, the matching coefficients become less positive (Mns of .284 vs .319), and the far coefficients become slightly less negative (Mns of −.078 vs −.088), but the pattern of results and statistical significance is unchanged. However, whilst the effect of German achievement on math self-efficacy is consistently negative, the effect of math achievement on German self-efficacy is consistently non-significant.

In summary, there is good support for the I/E predictions in relation to ASC and their robustness in relation to the control variables considered here. However, the difference between the matching (horizontal) and far cross-paths is much larger for self-concept than for self-efficacy responses (.614 vs. −.143 for self-concept; .319 vs. −.088 for self-efficacy). Thus, support for I/E predictions in relation to homework self-efficacy is weaker, both in terms of the sizes of the coefficients and even the direction of the effects based on far comparisons. Extending this model in the next section, we ask whether the size or even direction of these contrast effects is consistent for "near" comparisons between more closely related domains that might be seen as complementary.

9.2.2. DCT extension to include six domains

In the extended SEM model we include 36 (6 × 6) paths, leading from each of the six achievement domains to each of the six latent self-concept factors, and another 36 paths leading from achievement domains to the six homework self-efficacy factors (see Table 4). Again, these are considered both with and without control for the set of five covariates (see earlier discussion). For each set of 36 paths, coefficients are classified a priori as matching, near and far comparisons, based on the theoretical ordering of ASC factors (Fig. 2). For the present purposes, path coefficients involving any pair of the verbal-like constructs (German, English, history) or between any pair of the math/science-like (biology, physics, and math) constructs are considered to be "near" comparisons, whilst comparisons between achievement from one group to self-concept or self-efficacy in the other group are considered as "far" comparisons.

The mean across the 36 path coefficients for self-concept (Mn = .087, SE = .007, Table 4, no control) and for self-efficacy (Mn = .045, SE = .010, Table 4, no control) represents substantially positive coefficients for the six matching comparisons (self-concept: Mn = .587, SE = .028; self-efficacy: Mn = .313, SE = .023) and substantially smaller coefficients for the remaining 30 non-matching comparisons (self-concept: Mn = -.014, SE = .007; self-efficacy: Mn = -.008, SE = .007). For the present purposes, the critical feature of these non-matching coefficients is that the mean of the 12 near coefficients is significantly positive for both self-concept and self-efficacy responses (self-concept: Mn = .048, SE = .017; self-efficacy: Mn = .029, SE = .014). In contrast, the mean of the 18 "far" coefficients is significantly negative for both self-concept (Mn = −.055, SE = .010) and self-efficacy (Mn = −.035, SE = .016). For both self-concept and self-efficacy responses, the difference between the near and far coefficients is significant. However, the difference between the matching (horizontal) and far cross-paths is much larger for self-concept than for self-efficacy responses (.587 vs. −.055 for self-concept; .313 vs. −.035 for self-efficacy). Again, the results are altered somewhat when controls for the set of five covariates are included, in that the means of the matching coefficients are smaller. However, the difference between near and far coefficients is nearly the same.

Inspection of individual paths shows that statistically significant paths for non-matching domains primarily involve the four domains that are most extreme on the underlying verbal-math continuum: the two verbal scales (German and English), math, and physics. In order to evaluate this possibility, we computed summary statistics based on just these four scales. For both self-concept and homework self-efficacy responses, contrast effects based on far comparisons were more negative (−.116 vs. −.055 for self-concept; −.070 vs. −.035 for self-efficacy), whilst assimilation effects based on near comparisons were more positive (.086 vs. .048 for self-concept, .037 vs. .024 for self-efficacy).

10. Discussion

Self-evaluations are formed in relation to social and temporal comparisons—the better I perform relative to others and relative to my past accomplishments, the more positive my self-evaluation. As emphasized by the I/E model and its extension into DCT, internal dimensional comparisons among different domains provide an additional basis of comparison, with important implications for theory, achievement, individual choice behavior, for policy and for practice. Indeed, a critical source of self-knowledge is knowing one's relative strengths and weaknesses in different domains as a basis of academic choices, accurate self-evaluation, self-improvement, self-maintenance, and self-enhancement. The classic I/E model focuses primarily on how math and verbal achievements are related to math self-concept and verbal self-concept (Fig. 1A). As shown by an impressive array of correlational, longitudinal, cross-cultural, experimental, and qualitative studies, better math achievement leads to lower verbal self-concept and the better verbal achievement leads to lower math self-concept—the critical, seemingly paradoxical prediction of the I/E model. Here we explore new methodological, theoretical, and substantive insights into the extension of the I/E model into DCT.

10.1. Dimensional Comparison Theory (DCT): near versus far comparisons

A central feature of DCT and its extension of the I/E model is the distinction between near and far comparisons. Here we integrate theoretical perspectives based on the Marsh/Shavelson structural model of self-concept (Fig. 2) with new insights from DCT. The extensive research literature based on the I/E model demonstrates that better achievement in one domain has a negative effect on ASCs in a far domain. In the broader social comparison literature, these negative effects are referred to as contrast effects (see Möller & Marsh, 2013). DCT posits that the negative (contrast) effects will diminish and may even become positive (assimilation) effects when dimensional comparisons are based on complementary domains that are close together on the academic continuum. Here we operationalized this extension and developed new statistical tests to evaluate these predictions. This article is the first test of DCT theoretical predictions based on a comprehensive range of academic domains. Both Studies 1 and 2 provide clear support for the critical prediction that paths for achievement to ASC based on near comparisons will be less negative than those based on far comparisons.

A key feature of DCT is the theoretical continuum of academic subjects along the math-to-verbal continuum (Fig. 2). In this respect it integrates well-established theoretical and empirical research based on the Marsh/Shavelson multidimensional, hierarchical model of self-concept with the social comparison research pursued in many disciplines. Nevertheless, there is a need for further research as to whether pair-wise differences between the perceived complementarity/antagonism of a whole range of academic domains can be adequately represented by this continuum as posited in DCT (e.g., more research like Goetz et al., 2010, who quantified differences between different school subjects based on students' responses academic emotions associated their different school subjects).

10.2. The juxtaposition between ASC & self-efficacy

Early I/E research provided dramatic support for the theoretical distinction between appropriately constructed self-efficacy and self-concept measures. In particular, self-efficacy responses were based on a criterion reference that was explicit in the construction of the item, whilst self-concept responses invoked social and dimensional comparisons to establish a frame of reference. Critically, there was little or no support for I/E predictions based on self-efficacy responses, whilst support for the predictions based on self-concept responses continued to grow. Also of relevance, there was a near-zero correlation between math and verbal ASC responses, but a substantial positive correlation between math and verbal self-efficacy responses. However, this clear distinction between self-efficacy and ASC responses was based on a narrowly defined self-efficacy construct in which the standard of comparison was made explicit in the construction of items, whilst many measures that are labeled self-efficacy do not meet these standards of instrument construction.

In Study 2, we extended DCT to consider homework self-efficacy responses. Consistent with expectations, correlations among the self-efficacy factors were much larger than correlations among the self-concept factors. Nevertheless, there was good support for the convergent and discriminant validity of self-efficacy responses in relation to both achievement and self-concept. Although there was clear support for some a priori predictions, there were also some problems. In particular, there was some—albeit weak—support for contrast effects based on self-efficacy responses, although this was weaker than for ASC responses. This leaves the question as to why there apparently were dimensional comparison effects for self-efficacy responses here, whilst there were none in the classic studies by Marsh et al. (1991) and by Skaalvik and Rankin (1995).

There are at least two obvious differences. First, in the present study self-efficacy responses were in relation to completion of homework, rather than to achievement per se. Second, the self-efficacy responses in the earlier studies were in relation to a highly prescriptive construction of self-efficacy items that was likely to truncate dimensional comparison effects even more than in the present investigation. We suspect that the construction of items is the critical issue, and note that the typical self-efficacy items do not meet these rigid standards. Hence, we suggest that the results here are likely to be more representative of self-efficacy research in general than those in studies by Marsh et al. and by Skaalvik and Rankin (1995). Nevertheless, clearly this is an area for further research.

10.3. Limitations and directions for further research

The contributions of the present investigation are primarily theoretical (testing new theoretical predictions based on DCT, testing theoretical distinctions between self-concept and self-efficacy) and methodological (stronger designs and better statistical methodology for testing these hypotheses). However, support for the DCT, like most I/E research more generally, is largely based on cross-sectional, correlational studies, so that causal predictions should be offered tentatively and results must be interpreted cautiously. Fortunately, there is now a growing body of I/E studies, using various combinations of qualitative introspection studies, longitudinal, quasi-experimental, and true experimental designs with random assignment of students, which support the I/E model (see earlier discussion). However, there is a need for further research along these lines to support the causal interpretation of new theoretical predictions based on DCT—particularly in relation to the distinction between matching, near, and far comparisons, but also in relations to the conditions under which near comparisons are likely to result in assimilation rather than contrast.

Further research is also needed to establish conditions under which assimilation might be expected, particularly given that there was support for assimilation in Study 2 but not Study 1. However, we do note that our a priori predictions were that negative cross-paths would be significantly less negative (less contrast) or even positive (assimilation), and support for these predictions was highly consistent across both studies. Furthermore, in the research literature on social comparison effects, support for assimilation effects is also inconsistent (Diener and Fujita, 1997, Marsh et al., 2008a, Marsh et al., 2010). Clearly, DCT predicts that assimilation is more likely for maximally near comparisons (e.g., math and physics) than far comparisons (e.g., math and verbal domains) and more likely for near comparisons that are based on two domains very close to each other on the academic continuum (e.g., math and physics compared to math and biology). In addition to these DCT we speculate that contextual or situations features might influence the results. Thus, if math and physics are taught as a unified subject, reinforcing their mutual dependence, then assimilation might be more likely. Also, if we asked students to directly rate the perceived similarity of different academic domains, assimilation might be more likely for students who perceived a given pair of domains to be more similar. These ratings of perceived similarity could be used to validate the continuum but also provide more nuanced tests of DCT predictions such that the perceived "nearness" or "farness" of different domains might influence the size of the observed relations over and above the more objective positioning of constructs on the continuum. Because DCT is relatively new, particularly these speculations about assimilation offer an exciting direction for further research.

References

Albert, S. (1977). Temporal comparison theory. Psychological Review, 84, 485e503. http://dx.doi.org/10.1037/0033-295X.84.6.485.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (2006a). Going global with social cognitive theory: from prospect to Pay dirt. In S. I. Donaldson, D. E. Berger, & K. Pezdek (Eds.), Applied psychology: New frontiers and rewarding careers (pp. 53e79). Mahwah, NJ: Lawrence Erlbaum.

Bandura, A. (2006b). Toward a psychology of human agency. Perspectives on Psychological Science, 1, 164e180. http://dx.doi.org/10.1111/j.1745-6916.2006.00011.x.

Bong, M. (1998). Tests of the internal/external frames of reference model with subject-specific academic self-efficacy and frame-specific academic self concepts. Journal of Educational Psychology, 90, 102e110. http://dx.doi.org/10.1037/0022-0663.90.1.102.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: how different are they really? Educational Psychology Review, 15(1), 1e40. http://dx.doi.org/10.1023/A:1021302408382.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general Self-Efficacy Scale. Organizational Research Methods, 4, 62e83. http://dx.doi.org/10.1177/109442810141004.

Chen, S.-K., Yeh, Y.-C., Hwang, F.-M., & Lin, S. S. J. (2013). The relationship between academic self-concept and achievement: a multicohortemultioccasion study. Learning and Individual Differences, 23, 172e178. http://dx.doi.org/10.1016/j.lindif.2012.07.021.

Chiu, M. S. (2008). Achievements and self-concepts in a comparison of mathematics and science: exploring the internal/external frame of reference model across 28 countries. Educational Research and Evaluation, 14, 235e254. http://dx.doi.org/10.1080/13803610802048858.

Chiu, M.-S. (2012). The internal/external frame of reference model, big-fish-littlepond effect, and combined model for mathematics and science. Journal of Educational Psychology, 104, 87e107. http://dx.doi.org/10.1037/a0025734.

Crain, R. M. (1996). The influence of age, race, and gender on child and adolescent multidimensional self-concept. In B. A. Bracken (Ed.), Handbook of self-concept: Developmental, social, and clinical considerations (pp. 395e420). Oxford, England: John Wiley.

Dickhauser, O. (2003). Überprüfung des erweiterten modells des internal/external frame of reference [Test of the extension of the internal/external frame of reference model]. Zeitschrift für Entwicklungspsychologie und Padagogische Psychologie, 35, 200e207. http://dx.doi.org/10.1026//0049-8637.35.4.200.

Diener, E. (2000). Subjective well-being The science of happiness and a proposal for a national index. American Psychologist, 55(1), 34e43. http://dx.doi.org/10.1037/0003-066X.55.1.34.

Diener, E., & Fujita, F. (1997). Social comparison and subjective well-being. In B. P. Buunk, & F. X. Gibbons (Eds.), Health, coping, and well-being: Perspectives from social comparison theory (pp. 329e358). Mahwah, NJ: Erlbaum.

Eccles, J., Barber, B., Jozefowicz, D., Malenchuk, O., & Vida, M. (1999). Self-evaluations of competence, task values, and self-esteem. In N. G. Johnson,

M. C. Roberts, & J. Worell (Eds.), Beyond appearance: A new look at adolescent girls (pp. 53e83). http://dx.doi.org/10.1037/10325-002.

Eccles, J. S., Vida, M., & Barber, B. L. (2004). The relation of early adolescent plans, and both academic ability and task value beliefs to subsequent college enrollment. Journal of Early Adolescence, 24, 63e77. http://dx.doi.org/10.1177/0272431603260919.

Eccles, J. S., Wigfield, A., Harold, R., & Blumenfeld, P. (1993). Age and gender differences in children's self and task perceptions during elementary school. Child Development, 64, 830e847. http://dx.doi.org/10.2307/1131221.

Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic selfconcept: reconsidering structural relationships. Learning and Individual Differences, 19(4), 499e505. http://dx.doi.org/10.1016/j.lindif.2009.05.004.

Festinger, L. (1954). A theory of social comparison processes. Human Relations, 7, 117e140. http://dx.doi.org/10.1177/001872675400700202.

Goetz, T., Cronjaeger, H., Frenzel, A. C., Lüdtke, O., & Hall, N. C. (2010). Academic self concept and emotion relations: domain specificity and age effects. Contemporary Educational Psychology, 35, 44e58.http://dx.doi.org/10.1016/j.cedpsych.2009.10.001.

Goetz, T., Frenzel, C. A., 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, 9e33. http://dx.doi.org/ 10.1016/j.cedpsych.2006.12.002.

Goetz, T., Frenzel, A. C., Pekrun, R., & Hall, N. C. (2006). The domain specificity of academic emotional experiences. The Journal of Experimental Education, 75, 5e29. http://dx.doi.org/10.1037/0022-0663.99.4.715.

Guay, F., Larose, S., & Boivin, M. (2004). Academic self-concept and educational attainment level: a ten-year longitudinal study. Self and Identity, 3(1), 53e68. http://dx.doi.org/10.1080/13576500342000040.

Heller, K. A., & Perleth, Ch. (2000). Kognitiver Fahigkeitstest für 4.-12. Klassen, Revision (KFT 4-12þ R) [Cognitive Ability Test, revised version (KFT 4-12 þ R)]. Gottingen, Germany: Hogrefe € .

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1e55. http://dx.doi.org/10.1080/10705519909540118.

James, W. (1890/1963). The principles of psychology. New York: Holt, Rinehart & Winston.

Jerusalem, M., & Schwarzer, R. (1992). Self-efficacy as a resource factor in stress appraisal processes. In R. Schwarzer (Ed.), Self-efficacy: Thought control of action (pp. 195e213). Washington, DC: Hemisphere.

Marsh, H. W. (1986). Verbal and math self-concepts: an internal/external frame of reference model. American Educational Research Journal, 23, 129e149. http://dx.doi.org/10.3102/00028312023001129.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: preadolescence to early-adulthood. Journal of Educational Psychology, 81, 417e430. http://dx.doi.org/10.1037/0022-0663.81.3.417.

Marsh, H. W. (1990a). A multidimensional, hierarchical model of self-concept: theoretical and empirical justification. Educational Psychology Review, 2, 77e172. http://dx.doi.org/10.1007/BF01322177.

Marsh, H. W. (1990b). The influences of internal and external frames of reference on the formation of math and English self-concepts. Journal of Educational Psychology, 82, 107e116. http://dx.doi.org/10.1037/0022-0663.82.1.107.

Marsh, H. W. (1991). The failure of high ability high schools to deliver academic benefits: the importance of academic self-concept and educational aspirations. American Educational Research Journal, 28, 445e480. http://dx.doi.org/10.2307/1162948.

Marsh, H. W. (1993). Academic self-concept: theory measurement and research. In J. Suls (Ed.), Psychological perspectives on the self, 4 pp. 59e98). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Marsh, H. W. (1994). Sport motivation orientations: beware of the jingle-jangle fallacies. Journal of Sport and Exercise Psychology, 16, 365e380.

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., et al. (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, 108e128. http://dx.doi.org/10.1037/a0029907.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of fit indices in confirmatory factor analysis: the effect of sample size. Psychological Bulletin, 102, 391e410. http://dx.doi.org/10.1037/0033-2909.103.3.391.

Marsh, H. W., Byrne, B. M., & Shavelson, R. (1988). A multifaceted academic selfconcept:

Its hierarchical structure and its relation to academic achievement. Journal of Educational Psychology, 80, 366e380. http://dx.doi.org/10.1037/0022-0663.80.3.366.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspectives. Perspectives on Psychological Science, 1, 133e163. http://dx.doi.org/10.1111/j.1745-6916.2006.00010.x.

Marsh, H. W., Craven, R. G., Hinkley, J. W., & Debus, R. L. (2003). Evaluation of the Big-Two-Factor Theory of academic motivation orientations: an evaluation of jingle-jangle fallacies. Multivariate Behavioral Research, 38, 189e224. http:// dx.doi.org/10.1207/S15327906MBR3802_3.

Marsh, H. W., & Hau, K. T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: cross-cultural generalizability of the internal-external frame of reference predictions across 26 countries. Journal of Educational Psychology, 96, 56e67. http://dx.doi.org/10.1037/0022-0663.96.1.56.

Marsh, H. W., Hau, K.-T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: cross-cultural, psychometric comparisons across 25 countries. International Journal of Testing, 6, 311e360. http://dx.doi.org/10.1207/ s15327574ijt0604_1.

Marsh, H. W., Kong, C. K., & Hau, K. (2001). Extension of the internal/external frame of reference model of self-concept formation: importance of native and non-native languages for Chinese students. Journal of Educational Psychology, 93, 543e553. http://dx.doi.org/10.1037/0022-0663.93.3.543.

Marsh, H. W., Moller, J., Parker, P., Xu, M. K., Benjamin, N., & Pekrun, R. (2014). Internal/ External frame of reference model. In J. D. Wright (Ed.), The International Encyclopedia of the social and behavioral sciences (2nd ed.). Oxford, UK: Elsevier (in press).

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self concept, self-esteem, achievement, and attainment over seven adolescent years: unidimensional and multidimensional perspectives of self-concept. Personality and Social Psychology Bulletin, 34, 542e552. http://dx.doi.org/ 10.1177/0146167207312313.

Marsh, H. W., & O'Mara, A. (2010). Long-term total negative effects of school average ability on diverse educational outcomes: direct and indirect effects of the big-fish-little-pond effect. German Journal of Educational Psychology, 24, 51e72. http://dx.doi.org/10.1024/1010-0652.a000004.

Marsh, H. W., Roche, L. A., Pajares, F., & Miller, D. (1997). Item-specific efficacy judgments in mathematical problem solving: the downside of standing too close to trees in a forest. Contemporary Educational Psychology, 22(3), 363e377. http://dx.doi.org/10.1006/ceps.1997.0942.

Marsh, H. W., Seaton, M., Kuyper, H., Dumas, F., Huguet, P., et al. (2010). Phantom behavioral assimilation effects: systematic biases in social comparison choice studies. Journal of Personality, 78, 671e710. http://dx.doi.org/10.1111/j.1467-6494.2010.00630.x.

Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., et al. (2008). The big-fish-little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. Educational Psychology Review, 20, 319e350. http://dx.doi.org/10.1007/s10648-008-9075-6.

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. Educational Psychologist, 20(3), 107e123. http://dx.doi.org/10.1207/s15326985ep2003_1.

Marsh, H. W., Trautwein, U., Lüdtke, O., & Koller, O. (2008). Social comparison and € big-fish-little-pond effects on self-concept and other self-belief constructs: role of generalized and specific others. Journal of Educational Psychology, 100, 510e524. http://dx.doi.org/10.1037/0022-0663.100.3.510.

Marsh, H. W., Walker, R., & Debus, R. (1991). Subject-specific components of academic self-concept and self-efficacy. Contemporary Educational Psychology, 16, 331e345. http://dx.doi.org/10.1016/0361-476X(91)90013-B.

Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: the effects of academic self-concept and achievement. American Educational Research Journal, 34, 691e720. http://dx.doi.org/10.3102/00028312034004691.

Moller, J. (2005). Paradoxical effects of praise and criticism: social, dimensional, and € temporal comparisons. British Journal of Educational Psychology, 75, 275e295. http://dx.doi.org/10.1348/000709904X24744

Moller, J., & Husemann, N. (2006). Internal comparisons in everyday life. € Journal of Educational Psychology, 98, 342e353. http://dx.doi.org/10.1037/0022-0663.98.2.342.

Moller, J., & Koller, O. (2001a). Dimensional comparisons: an experimental approach to the internal/external frame of reference model. Journal of Educational Psychology, 93, 826e835. http://dx.doi.org/10.1037/0022-0663.93.4.826.

Moller, J., & Koller, O. (2001b). Frame of reference effects following the € announcement of exam results. Contemporary Educational Psychology, 26, 277e287. http://dx.doi.org/10.1006/ceps.2000.1055.

Moller, J., & Marsh, H. W. (2013). Dimensional comparison theory. € Psychological Review, 120, 544e560. http://dx.doi.org/10.1037/a0032459.

Moller, J., Pohlmann, B., K € oller, 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, 1129e1167. http://dx.doi.org/10.3102/0034654309337522.

Moller, J., Retelsdorf, J., K € oller, O., & Marsh, H. W. (2011). The reciprocal I/E model: € an integration of models of relations between academic achievement and selfconcept. American Educational Research Journal, 48, 1315e1346. http:// dx.doi.org/10.3102/0002831211419649.

Moller, J., & Savyon, K. (2003). Not very smart, thus moral: dimensional comparisons between academic self-concept and honesty. Social Psychology of Education, 6(2), 95e106. http://dx.doi.org/10.1023/A:1023247910033.

Moller, J., Streblow, L., Pohlmann, B., & Koller, O. (2006). An extension to the internal/external frame of reference model to two verbal and numerical domains. European Journal of Psychology of Education, 21(4), 467e487. http://dx.doi.org/ 10.1007/BF03173515.

Muthen, L. K., & Muth en, B. O. (2008 e2013). Mplus user's guide. Los Angeles, CA: Muthen & Muthen.

Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., & Eccles, J. S. (2008). Gendered high school course selection as a precursor of gendered careers: the mediating role of self-concept and intrinsic value. In H. M. G. Watt, & J. S. Eccles (Eds.), Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences (pp. 115e143). http://dx.doi.org/10.1037/11706-004.

Nagy, G., Trautwein, U., Baumert, J., Koller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: effects of academic self-concept and intrinsic value. Educational Research and Evaluation, 12, 323e345. http://

dx.doi.org/10.1080/13803610600765687.
Pajares, F., & Schunk, D. H. (2005). Self-efficacy and self-concept beliefs: jointly contributing to the quality of human life. In H. W. Marsh, R. G. Craven, & D. M. McInerney (Eds.), International advances in self research (Vol. 2, pp. 95e123). Greenwich, CT: Information Age.
Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., & Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: a multicontext study. Developmental Psychology, 48(6), 1629e1642. http://dx.doi.org/10.1037/a0029167.
Pekrun, R. (2006). The control value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. Educational Psychology Review, 18, 315e341. http://dx.doi.org/10.1007/s10648-006-9029-9.
Pinxten, M., De Fraine, B., Van Damme, J., & D'Haenens, E. (2010). Causal ordering of academic self-concept and achievement: effects of type of achievement measure. British Journal of Educational Psychology, 80(4), 689e709. http://dx.doi.org/10.1348/000709910X493071.
Pohlmann, B., & Moller, J. (2009). On the bene € fit of dimensional comparisons. Journal of Educational Psychology, 101, 248e258. http://dx.doi.org/10.1037/a0013151.
Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), Causal and control beliefs Measures in health psychology: A user's portfolio (pp. 35e37). Windsor, UK: NFER-NELSON.
Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: an introduction. American Psychologist, 55, 5e14. http://dx.doi.org/10.1037/0003-066X.55.1.5.
Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: validation of construct interpretations. Review of Educational Research, 46, 407e444. http://www.jstor.org/stable/1170010.
Skaalvik, E. M., & Rankin, R. J. (1990). Math, verbal, and general academic selfconcept: the internal/external frame of reference model and gender differences in self-concept structure. Journal of Educational Psychology, 82(3), 546e554. http://dx.doi.org/10.1037/0022-0663.82.3.546.
Skaalvik, E. M., & Rankin, R. J. (1995). A test of the internal/external frame of

reference model at different levels of math and verbal self-perception. American Educational Research Journal, 32(1), 161e184. http://dx.doi.org/10.2307/1163217.

Skaalvik, E. M., & Skaalvik, S. (2002). Internal and external frames of reference for academic self-concept. Educational Psychologist, 37, 233e244. http://dx.doi.org/ 10.1207/S15326985EP3704_3.

Skinner, E. A., Chapman, M., & Baltes, P. (1988). Control, means-ends, and agency beliefs: a new conceptualization and its measurement during childhood. Journal of Personality and Social Psychology, 54, 117e133. http://dx.doi.org/10.1037/ 0022-3514.54.1.117.

Thorndike, R. L., & Hagen, E. (1993). The cognitive abilities test: Form 5. Itasca, IL: Riverside Publishing.

Trautwein, U., & Lüdtke, O. (2007). Students' self-reported effort and time on homework in six school subjects: between-students differences and withinstudent variation. Journal of Educational Psychology, 99(2), 432e444. http:// dx.doi.org/10.1037/0022-0663.99.2.432.

Trautwein, U., & Lüdtke, O. (2009). Predicting homework motivation and homework effort in six school subjects: the role of person and family characteristics, classroom factors, and school track. Learning and Instruction, 19(3), 243e258. http://dx.doi.org/10.1016/j.learninstruc.2008.05.001.

Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: support for a domain-specific, multilevel homework model. Journal of Educational Psychology, 98(2), 438e456. http://dx.doi.org/10.1037/0022-0663.98.2.438.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: a meta-analytic review. Educational Psychologist, 39(2), 111e133. http://dx.doi.org/10.1207/s15326985ep3902_3.

Xu, M. K., Marsh, H. W., Hau, K.-T., Ho, I. T., Morin, A. J. S., & Abduljabbar, A. S. (2013). The internal/external frame of reference of academic self-concept: extension to a foreign language and the role of language of instruction. Journal of Educational Psychology, 105(2), 489e503. http://dx.doi.org/10.1037/a0031333.

Yeung, A. S., Lee, J. C., & Wong, H. (2001, April). Testing Marsh's (1986) frame of reference model of self-concept with bilingual students. Paper presented at the AERA Annual Meeting, New Orleans, LA.