Arens, A. Katrin; Becker, Michael; Möller, Jens

Social and dimensional comparisons in math and verbal test anxiety. Within- and cross-domain relations with achievement and the mediating role of academic self-concept

Social and Dimensional Comparisons in Math and Verbal Test Anxiety: Within- and

Cross-domain Relations with Achievement and the Mediating Role of Academic Self-concept

A. Katrin Arens

German Institute for International Educational Research, Germany

Michael Becker

German Institute for International Educational Research, Germany, &

Leibniz Institute for Science and Mathematics Education, Kiel, Germany

Jens Möller

University of Kiel

Highlights

- Negative relations between math (verbal) achievement and math (verbal) test anxiety

- Positive relations between verbal achievement and math test anxiety

- Within- and cross-domain relations mediated through academic self-concept

- Higher direct and indirect achievement relations for worry than for emotionality

- Gender invariance regarding the relations between the constructs

Abstract

The present study builds on two strands of research: (1) the recently established generalized internal/external frame of reference (GI/E) model assuming social (comparing one's achievement in one domain with the achievement of one's peers in the same domain) and dimensional (comparing one's achievement in one domain with one's achievement in another domain) comparison processes in the formation of motivational constructs and self-perceptions, and (2) research on domain-specific facets of test anxiety. Using a sample of 5135 German seventh grade students, it is tested whether and how both comparison processes are involved in the formation of domain-specific facets of test anxiety when considering both the emotionality and worry components of test anxiety, and whether the relation between achievement and test anxiety is mediated through academic self-concept. When applying the GI/E model to test anxiety, the results showed negative relations between achievement and test anxiety within math and verbal (German) domains, but partially positive relations across domains. This pattern of relations emerged for both the worry and emotionality components while stronger achievement relations were found for worry. These findings indicate that dimensional achievement comparison processes operate in the formation of domain-specific test anxiety. Domain-specific academic self-concepts were found to mediate the relations between achievement and test anxiety within and across domains, the mediation being stronger for worry than for emotionality as an outcome. Boys and girls did not differ regarding direct and indirect relations among constructs. Implications for research on dimensional comparison processes and test anxiety are discussed.

Keywords: dimensional comparisons; gender; I/E model; test anxiety; self-concept

Within research on test anxiety, one approach conceptualizes test anxiety as a domain-unspecific construct (e.g., Bandalos, Yates, & Thorndike-Christ, 1995; Cassady & Johnson, 2002; Elliot & McGregor, 1999; Gierl & Todd, 1996) while another approach considers the domain specificity of test anxiety (e.g., Goetz, Frenzel, Pekrun, & Hall, 2006; Goetz, Frenzel, Pekrun, Hall, & Lüdtke, 2007; Gogol, Brunner, Martin, Preckel, & Goetz, 2017; Marsh & Yeung 1996; Schnabel, 1998; Sparfeldt, Schilling, Rost, Stelzl, & Peipert, 2005). Within the domain-specific approach to test anxiety, researchers seek to explain the formation of domain-specific facets of test anxiety. In this regard, it is possible to draw on the generalized internal/external frame of reference (GI/E) model (Möller, Müller-Kalthoff, Helm, Nagy, & Marsh, 2015) which itself builds upon the internal/external frame of reference (I/E) model (Marsh, 1986, 1990) and has been derived from dimensional comparison theory (Möller & Marsh, 2013). The GI/E model assumes that both social (comparing one's achievement in one domain with the achievement of one's peers in the same domain) and dimensional (comparing one's achievement in one domain with one's achievement in another domain) achievement comparisons are involved in the formation of domain-specific outcomes, which might also involve test anxiety. In this case, achievement and test anxiety might be negatively associated within matching domains (e.g., math achievement and math test anxiety), but positively related across non-matching domains (e.g., verbal achievement and math test anxiety). The present study tests this assumption and therefore examines whether social and dimensional achievement comparisons are involved in the formation of students' domain-specific test anxiety when considering both the worry and emotionality components of test anxiety (Liebert & Morris, 1967). Moreover, this study investigates whether the within-domain and cross-domain relations between achievement and test anxiety are mediated through domain-specific academic self-concept. Finally, we examine gender differences versus gender invariance in the direct and indirect (i.e., mediated through academic self-concepts) relations

between achievement and test anxiety (including worry and emotionality), in order to test the generalizability of the findings across student characteristics.

1. Models of Test Anxiety

Test anxiety has been shown to be negatively related to achievement (e.g.,Chapell et al., 2005; Gogol et al., 2017; Frenzel, Pekrun, & Goetz, 2007a; Hembree, 1988, 1990; Ho et al., 2000; Pajares & Graham, 1999), cognitive functioning or working memory (e.g., Ashcraft & Kirk, 2001; Eysenck, Derakshan, Santos, & Calvo, 2007), study skills including metacognition, motivation, self-regulation, and information organization (e.g., Ashcraft, 2002; Cassady, 2004; Ma, 1999; Schutz & Davis, 2000), and global well-being including self-esteem (e.g., Beidel & Turner, 1988; Pekrun et al., 2004). These detrimental effects of test anxiety have invoked high levels of research interest in and attention devoted to test anxiety, which has thus become the probably most central or at least the most profoundly studied academic emotion (Bodas & Ollendick, 2005; Lee, 2009; Stöber & Pekrun, 2004; Zeidner, 1998, 2007). Correspondingly, high levels of effort have been invested into research on test anxiety which addresses the state versus trait differentiation of test anxiety (Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013; Spielberger, Gorsuch, & Lushene, 1970; Zeidner, 1998), its internal structure (Gogol et al., 2017), its stability versus change across time (Gogol, Brunner, Preckel, Goetz, & Martin, 2016), and gender differences in its mean levels (Cassadey & Johnson, 2002; Chapell et al., 2005; Else-Quest, Hyde, & Linn, 2010; Hong & Karstensson, 2002; Martin, 2007; Miller & Bichsel, 2004). In the following, we elaborate on (1) the differentiation between worry and emotionality components of test anxiety, and (2) the domain-unspecific and domain-specific approaches to test anxiety as most relevant to the present study.

1.1 The Differentiation between Worry and Emotionality

Liebert and Morris (1967) proposed a structural model of test anxiety according to which test anxiety encompasses two facets – worry and emotionality (Gierl & Todd, 1996;

Stöber, 2004; Zeidner, 2007). The worry facet targets the cognitive component and includes rumination, negative expectations about potential consequences of failure, and detrimental self-evaluations. The emotionality facet addresses the physiological component and depicts arousal or nervousness, demonstrated by, for example, increased levels of heart rate, sweating, feelings of panic, or stomach troubles.

To date, this model of test anxiety has been addressed by many empirical studies (e.g., Cassady & Johnson, 2002; Elliot & McGregor 1999; Hong & Karstensson, 2001). In essence, worry and emotionality have been shown to constitute separate factors which bear differential relations to achievement (Everson, Millsap, & Rodriquez, 1991; Morris, Davis, & Hutchings, 1981). The worry component has been found to consistently show negative associations with achievement (Hembree, 1988). For emotionality, some studies did not suggest a relation to achievement at all (Elliot & McGregor 1999; Hong, 1999; Morris et al., 1981), while others revealed a negative achievement relation which is yet weaker than the relation between the worry component and achievement (Cassady & Johnson, 2002; Hembree, 1988; Hong & Karstensson, 2001; Seipp, 1991).

1.2 Domain-unspecific and Domain-specific Approaches

Test anxiety can be considered to be a domain-unspecific construct. In this case, test anxiety is measured without any relation to a specific content domain but students are asked regarding their test anxiety in general (Bandalos et al., 1995; Cassady & Johnson, 2002; Elliot & McGregor, 1999; Gierl & Todd, 1996). Another approach considers test anxiety to be a domain-specific construct. Here, students are assumed to display differential facets of test anxiety for different school subjects (Frenzel et al., 2007a; Goetz, Cronjaeger, Frenzel, Lüdtke, & Hall, 2010; Goetz, Frenzel et al., 2006, 2007; Goetz, Pekrun, Hall, & Haag, 2006; Marsh & Yeung 1996), within which students further distinguish between worry and emotionality components (Sparfeldt et al., 2005; Wigfield & Meece, 1988). The domain specificity of test anxiety becomes evident in confirmatory factor analyses (CFA; Brown,

2006) which establish separate factors for test anxiety in relation to different domains and show weak relations among these domain-specific factors (i.e., between test anxiety experienced in math and verbal domains).

When considering the relation between achievement and test anxiety, research adapting a domain-specific approach has often focused on one domain only. In this context, numerous studies have demonstrated a negative relation between test anxiety and achievement in math (Frenzel et al., 2007a; Hembree, 1990; Ho et al., 2000; Ma, 1999; Pajares & Graham, 1999). Goetz et al. (2010) included other domains than math and stated separate models for investigating the relation between achievement and test anxiety in various specific domains (i.e., separate models for the relation between achievement and test anxiety in math, physics, German, and English). However, within the domain-specific approach to test anxiety, little effort has so far been invested in examining the relations between achievement and test anxiety *across* different domains (e.g., the relation between math achievement and verbal test anxiety). Dimensional comparison theory (Möller & Marsh, 2013) and the GI/E model (Möller et al., 2015) may offer an adequate theoretical framework for studying the relations between achievement and test anxiety across different domains.

2. Dimensional Comparison Theory

Dimensional Comparison Theory (DCT; Möller & Marsh, 2013) bases on the I/E model (Marsh, 1986, 1990; Möller, Pohlmann, Köller, & Marsh, 2009) which illustrates the relations between math and verbal achievements and math and verbal self-concepts (i.e., students' self-perceptions of competence in math and verbal domains). Math and verbal self-concepts have consistently been found to be only weakly correlated although math and verbal achievements have been found to be substantially correlated (Marsh, 1986, 1990). Moreover, math (verbal) self-concepts and achievements have been found to be positively related, but math (verbal) achievement has been found to be negatively related to verbal (math) self-concept. These surprising findings have been explained by assuming an interplay of both

social and dimensional achievement comparison processes in the formation of domain-specific (i.e., math and verbal) academic self-concepts. In the social (external) comparison process, students compare their own achievement in one domain to their classmates' achievements in the same domain. This kind of comparison process is supposed to lead to a positive correlation between math and verbal self-concepts since math and verbal achievements are highly correlated (Möller et al., 2009). In addition, the social comparison process invokes a positive relation between achievements and self-concepts of matching domains (i.e., within-domain achievement–self-concept relations; e.g., math achievement and math self-concept). In a dimensional (internal) comparison process, students compare their individual achievement in one domain (e.g., math) to their individual achievement in another domain (e.g., verbal). This comparison process results in a negative correlation between math and verbal self-concepts and in negative relations between achievement and self-concept of non-matching domains (i.e., cross-domain achievement–self-concept relations). As such, higher math (verbal) achievement leads to lower verbal (math) self-concept. This phenomenon is known as a contrast effect. The negative correlation between math and verbal self-concepts resulting from the dimensional comparison process and the positive correlation resulting from the social comparison process outbalance one another, leading to a low or near-zero correlation between math and verbal self-concepts.

The I/E model has invoked increased attention and interest regarding the phenomenon of dimensional comparison processes. While the original I/E model (Marsh, 1986, 1990; Möller et al., 2009) focuses on the dimensional comparison processes at play in the formation of math and verbal self-concepts, recent research has become more interested in dimensional comparison processes in terms of a general psychological phenomenon. DCT has subsequently been established as the theory behind dimensional comparison processes, elaborating on the antecedents, psychological processes, and consequences associated with dimensional comparison processes. In this context, DCT widens the perspective on

dimensional comparison processes and provides a broader definition. Generally speaking, dimensional comparisons take place when individuals compare perceptions of aspects of a particular domain A with perceptions of aspects of a particular domain B, bearing consequences for any kind of outcomes related to these domains. DCT thus provides a theoretical foundation for empirical research on dimensional comparison processes. The corresponding work program is described by the GI/E model.

2.1 The GI/E Model

Building upon DCT which itself provides a broader perspective on dimensional comparison processes, the GI/E model extends the original I/E model. Accordingly, the GI/E model allows other predictor variables than math and verbal achievements which are subject to dimensional comparison processes. In addition, in the GI/E model, the rationale of the original I/E model was expanded to other outcome variables beyond math and verbal self-concepts so that a variety of outcome variables are considered which may be influenced by dimensional comparison processes. In this context, student ratings on self-regulated learning (Miller, 2000), intrinsic motivation (Marsh, Abduljabbar et al., 2015), perceptions of the learning environment (Arens & Möller, 2016), or interest (Schurtz, Pfost, Nagengast, & Artelt, 2014) have so far been found to be affected by dimensional achievement comparisons. When restricted to math and verbal domains, the GI/E model pattern becomes salient in (1) positive achievement–outcome relations within the math and verbal domains illustrating social comparison processes, and (2) simultaneous negative achievement–outcome relations across math and verbal domains indicating dimensional achievement comparison processes and contrast effects between math and verbal domains. Such a pattern of relations was also found when investigating the relations between math and verbal achievements and math and verbal enjoyment, the latter representing a positive academic emotion (Goetz, Frenzel, Hall, & Pekrun, 2008). This finding suggests that academic emotions might also be influenced by

dimensional achievement comparisons and may help explain the recently found domain specificity of test anxiety.

2.2 Application of the GI/E Model to Math and Verbal Test Anxiety

When applying the GI/E model to math and verbal test anxiety, the most basic precondition is the assessment of math and verbal achievements along with students' math and verbal test anxiety. Ideally, both the worry and emotionality components of test anxiety should be measured to examine whether the GI/E model assumptions hold for both components of test anxiety. Some studies have suggested higher negative relations between achievement and worry than between achievement and emotionality, yet respective studies primarily included only domain-unspecific measures of test anxiety (Cassady & Johnson, 2002; Hembree, 1988; Hong & Karstensson 2001; Seipp, 1991; Zeidner & Schleyer, 1999). Based on these findings, one could expect that the worry component shares higher relations with achievement than the emotionality component when also considering within-domain and cross-domain relations between achievement and domain-specific measures of test anxiety. However, when focusing on within-domain relations for math, some studies (Ho et al., 2000; Wigfield & Meece, 1988) demonstrated higher achievement relations for the emotionality component. So far, no study has examined differential achievement relations for the worry and emotionality component when considering cross-domain relations.

A (latent) regression model should be preferably chosen as the statistical approach used to test the GI/E model assumptions as it estimates the paths leading from domain-specific achievements to domain-specific facets of test anxiety while controlling for the other relations. Significant within-domain relations between achievement and test anxiety would indicate the operation of social achievement comparison processes, and significant cross-domain relations would indicate the operation of dimensional achievement comparison processes in the formation of domain-specific facets of test anxiety. Here, it should be mentioned that negative within-domain relations between achievement and test anxiety depict

the operation of social achievement comparison processes, and positive cross-domain relations depict the operation of contrast effects resulting from dimensional achievement comparison processes (e.g., higher achievement in math leads to lower test anxiety in math but to higher test anxiety in verbal domains). Readers familiar with the original I/E model might be used to positive within-domain relations indicating social comparison processes, while negative cross-domain paths between math and verbal achievement and self-concept measures represent contrast effects due to dimensional comparisons (Marsh, 1986, 1990; Möller et al., 2009).

These suggestions for applying the GI/E model to test anxiety have in part been addressed by a few studies, but so far no study has realized the full approach. Goetz, Frenzel et al. (2007) reported the correlations among test anxiety and achievement measures referring to German, English, math, and physics. Supporting the domain specificity of test anxiety, higher correlations were found between achievement and test anxiety measures addressing the same domain (e.g., math achievement and math test anxiety), while negligible correlations were demonstrated between achievement and test anxiety measures of non-matching domains (e.g., math achievement and test anxiety in German). However, Goetz, Frenzel et al. (2007) focused on correlational analyses rather than estimating a regression model. Thus, he could not provide insights into the paths linking achievement and test anxiety within and across domains, which would be the more appropriate approach to testing the GI/E model assumptions. Besides, the study did not discriminate between worry and emotionality components of test anxiety.

A differentiation between worry and emotionality components of test anxiety was yet realized by Sparfeldt et al. (2005), who examined the correlations among achievement, worry, and emotionality related to the four school subjects of math, physics, German, and English. Again supporting the domain specificity of test anxiety, achievement and worry respectively emotionality were substantially correlated within matching domains, but revealed negligible

correlations across non-matching domains. Given its restriction to correlational analyses, this study failed to fully test the GI/E model in terms of a regression model.

A latent regression model was yet realized by Marsh (1988) who offered first insights into the applicability of the GI/E model assumptions to test anxiety. Higher math achievement was found to be associated with lower test anxiety in math but with higher test anxiety in English. In parallel, English achievement was demonstrated to be negatively associated with test anxiety in English, but showed a positive relation to test anxiety in math. Hence, this study has provided evidence for the operation of social and dimensional achievement comparisons in the formation of domain-specific facets of test anxiety, and has shown a contrast effect between math and verbal domains. However, Marsh (1988) did not distinguish between the worry and emotionality components of test anxiety and could thus not probe for differential within-domain and cross-domain achievement relations for the worry relative to the emotionality component. The present study therefore aims to examine the validity of the GI/E model assumptions for the relations between math and verbal achievement and test anxiety when including both the worry and emotionality components of test anxiety.

3. Mediation through Academic Self-concept

The within-domain and cross-domain relations between math and verbal achievement and test anxiety as stated in the GI/E model might be mediated through academic self-concept [see Marsh (1988) who discussed this issue when supporting the GI/E model for test anxiety but did not explicitly test this possibility]. Achievement feedback such as school grades or test results might offer students an idea for self-evaluating their academic competence (i.e., academic self-concept) which then impacts on further motivational and emotional constructs. As such, the GI/E model-like relations between math and verbal achievement on the one hand and math and verbal enjoyment (Goetz et al., 2008) or math and verbal interest (Schurtz et al., 2014) on the other hand were found to be mediated through students' domain-specific academic self-concepts (see also Guo, Marsh, Parker, Morin, & Dicke, 2017).

Academic self-concept seems to be a reasonable mediator in the relation between achievement and test anxiety because academic self-concept shares relations to both achievement and test anxiety. First, academic self-concept might be substantially associated with achievement as academic self-concept reflects self-perceived achievement. Indeed, many empirical studies document substantial relations between achievement and self-concept, whereby the pattern of achievement–self-concept relations has been found to be domain-specific in nature. For example, math achievement is more highly related to math than to verbal self-concept (Huang, 2011; Marsh & Craven, 2006; Swann, Chang-Schneider, & Larsen McClarty, 2007; Valentine, DuBois, & Cooper, 2004).

Second, academic self-concept might be substantially associated with test anxiety. Theoretically, a relation between self-concept and test anxiety matches the control-value theory of achievement emotions (Pekrun, 2006; Pekrun & Stephens, 2010). This theory states two kinds of appraisals, i.e., control appraisals and value appraisals as antecedents of achievement emotions including test anxiety. The control appraisal encompasses competence beliefs such as academic self-concept in order that academic self-concept might influence test anxiety. Correspondingly, a number of studies documented substantial cross-sectional and longitudinal relations between students' academic self-concept and test anxiety within matching domains (e.g., Ahmed, Minnaert, Kuyper, & van der Werf, 2012; Frenzel et al., 2007a; Goetz et al., 2010; Gogol et al., 2017; Lee, 2009; Meece, Wigfield, & Eccles, 1990; Pekrun, 2006). Academic self-concept is conceptualized as students' self-perceptions of competence (Marsh & Craven, 2006) and has thus a cognitive focus covering positive and negative self-evaluations which are also included in the worry component of test anxiety (Liebert & Morris, 1967). Hence, relatively stronger associations between the worry component and academic self-concept than between the emotionality component and academic self-concept can be expected. Zeidner and Schleyer (1999) presented some first

corresponding evidence although their study was based on domain-unspecific measures of test anxiety rather than pursuing a domain-specific approach.

4. Generalizability across Gender

The original I/E model has been found to be generalizable across a variety of student characteristics such as gender (Möller et al., 2009; Skaalvik & Rankin, 1990), age (Marsh, Abduljabbar et al., 2015; Möller et al., 2009), or cultural background (Marsh, Abduljabbar et al., 2015; Marsh & Hau, 2004). However, one should not simply transfer evidence for the robustness of the original I/E model across student characteristics to the GI/E model tested here, i.e., when using test anxiety as an outcome variable. To this aim, in this study, we additionally investigate the invariance of the proposed GI/E model across boys and girls. In the context of the mediation model proposed above, it is also interesting to examine whether boys and girls differ in their indirect relations between math and verbal achievement and test anxiety, mediated through domain-specific academic self-concepts.

5. The Present Study

In the present study, math and verbal (i.e., German) achievement and test anxiety were measured with a sample of seventh grade German secondary school students. We investigated the relations between achievement and test anxiety within [i.e., between math (verbal) achievement and math (verbal) test anxiety] and across [i.e., between math (verbal) achievement and verbal (math) test anxiety] domains. In addition, we examined whether the within-domain and cross-domain relations between math and verbal achievement and math and verbal test anxiety are mediated by domain-specific self-concepts. Finally, we examined gender differences versus gender invariance in the direct and indirect within-domain and cross-domain relations between math and verbal achievement and math and verbal test anxiety. All these analytic steps are completed while differentiating between worry and emotionality components of math and verbal test anxiety.

6. Method

The data analyzed in this study were retrieved from the large-scale longitudinal project "Learning Processes, Educational Careers, and Psychosocial Development in Adolescence and Young Adulthood (BIJU)" conducted under the aegis of the Max Planck Institute for Human Development, Berlin, Germany (for more information on this data set, see for example Baumert et al., 1996; Schnabel, 1998). In this study, we focus on the first measurement wave of the BIJU study which took place at the beginning of the school year 1991/1992 when students attended grade level 7. Participating students came from three German federal states (North Rhine-Westphalia, Mecklenburg-Western Pomerania, and Saxony-Anhalt). The sample analyzed here consists of 5135 students with 2459 (47.9%) boys and 2676 (52.1%) girls attending 275 different classes of 145 different schools. As expected for grade 7 German students, the students had a mean age of 12.88 years ($SD = 0.645$). The majority of participating students [$N = 4531$ (88.2%)] had German citizenship. To maintain a representative sample, it was stratified by region and school type, randomly sampling schools by school type and two seventh grade classes per school. Data were assessed in entire classes by trained research assistants.

In the German educational system, students are allocated to different achievement tracks for secondary schooling. Hence, primarily contingent upon students' accomplishments in elementary school, they attend the academic track (Gymnasium), the intermediate track (Realschule), the low achievement track (Hauptschule), or the comprehensive track (Gesamtschule). The sample of the present study encompasses students from all achievement tracks of secondary schools in Germany: academic track: $N = 2126$; intermediate track: $N = 675$; low achievement track: $N = 1513$; comprehensive track: $N = 435$; no information available: $N = 386$.

6.1 Measures

6.1.1 Test anxiety. A set of seven items retrieved from Helmke (1992, see also Schnabel, 1998) was used to measure the worry component of test anxiety in math ($\alpha = .812$)

and German (α = .755), respectively. These items address students' negative self-beliefs, expectations and ruminations about the consequences of failure, and undesirable social reactions in the case of failure (e.g.,"I thought of all the things I can't do."; "I thought about who of the other students would probably do best."; "I thought: What will the teacher think about my performance?"; "My mind was distracted from focusing on the tasks."). Another set of 10 items retrieved from Hodapp, Laux, and Spielberger (1982, see also Schnabel, 1998) was used to assess the domain-specific emotionality components (math: α = .929; German: α = .894). These items ask about students' physiological states (e.g., "My hand was trembling while I was writing."; "My heart was beating very fast."; "I had a strange feeling in my stomach."; "I started to sweat."). The same items were used for measuring worry and emotionality in both math and German. However, the students were instructed to respond to the items once when remembering a classroom test situation in math, and once when remembering a classroom test situation in German as a school subject. All items were rated on a 5-point Likert-type response scale ranging from "not at all" to "very often". Hence, higher values reflect higher levels of worry or emotionality.

6.1.2 Academic self-concept. In order to measure students' academic self-concept related to math and German, we used five items from Jopt (1978) and Jerusalem (1984) which are still used in contemporary self-concept research (see for example Möller, Zimmermann, & Köller, 2014; Stäbler, Dumont, Becker, & Baumert, 2017; Zimmermann, Möller, & Köller, 2017). These items ask for students' self-perceptions of competence. The items were formulated in parallel across the two domains as they had the same wordings and only differed in their targeted domain (e.g., "Nobody is perfect but I am just not good at math/German."; "I am not particularly good at math/German."; "Math/German just isn't my thing."). The reliability of the items was satisfactory for both math (α = .855) and German (α = .771). Students had to respond to the items on a four point Likert scale ranging from totally

true to not true. Before the analyses, the items were recoded in the way that high values consistently indicated high levels of self-concept.

6.1.3 Achievement. The school grades in math and German as school subjects the students had obtained in their last school report served as achievement indicators. In Germany, school grades range from 1 to 6, with 1 representing the best and 6 the poorest grade. To facilitate interpretation of the results, the grades were reversely coded before all analyses, thus higher values indicate higher levels of achievement.

6.2 Statistical Analyses

All analyses were conducted within the framework of structural equation modeling (SEM; e.g., Kline, 2005) using the statistical package of Mplus 7.4 (Muthén & Muthén, 1998-2015). All models were estimated by applying the robust maximum likelihood estimator (MLR) which has been shown to be robust against violations of normality assumptions and accounts for the treatment of item responses on a Likert-type scale as continuous variables (Beauducel & Herzberg, 2006). Missing values on all variables were estimated by the Full Information Maximum Likelihood (FIML) implemented in Mplus. The FIML approach is known to be reliable in handling missing data, making less restrictive assumptions than, for example, listwise deletion (Enders, 2010; Graham, 2009). The amount of missing values ranged between 10.5% and 11.7% for math worry, between 13.0% and 15.6% for German worry, between 11.7% and 12.3% for math emotionality, between 15.6% and 17.3% for German emotionality, between 5.7% and 6.1% for math self-concept, and between 8.1% and 8.6% for German self-concept. With regard to achievement measures, there were 8.4% missing values on math grades, and 9.2% missing values on German grades.

The data set has a multilevel structure since the participating students were nested into classes (Raudenbush & Bryk, 2002). Students attending the same class can be assumed to be more similar to each other than students attending different classes. Hence, the student ratings cannot be considered as independent observations. Therefore, all analyses were conducted

using the Mplus option "type = complex" with students' classes treated as clustering variables. This option corrects for possible biased standard errors resulting from the hierarchical nature of the data. Furthermore, all models considered correlated uniquenesses to account for potential shared method variance attributed to the wordings of the items (Marsh et al., 2013). Correlated uniquenesses were allowed between the parallel-worded items for measuring math and verbal self-concepts and between the same items assessing worry and emotionality related to math respectively German.

The analyses started with a CFA model to examine the integrity of the test anxiety and achievement measures (Brown, 2006). Hence, we estimated a measurement model stating separate factors for worry, emotionality, and achievement in math and German (3 x 2 = 6 factors; Model 1 in Table 1). The achievement factors were single-item factors defined by students' school grades, while the worry and emotionality factors were defined by the respective domain-specific set of items. We then tested the GI/E model assumptions by estimating a latent regression model in which worry and emotionality in both math and German served as outcome variables and were regressed on the math and German achievement factors.

We then turned to models also including academic self-concept. To warrant the integrity of the self-concept measures, we estimated a CFA model assuming eight factors (i.e., one factor for worry, emotionality, self-concept, and achievement in math and German, respectively; Model 2 in Table 1). We finally included math and verbal self-concepts as mediator variables in the GI/E model in order to test the assumption that the relations between achievement and worry (Model 3), between achievement and emotionality (Model 4), and between achievement and both worry and emotionality (Model 5) are mediated through academic self-concept. The path coefficients for the indirect relations were estimated in Mplus by specifying them through the "model indirect" option.

For the purpose of model fit evaluation, we follow the advice to consider a wide range of descriptive goodness-of-fit indices (e.g., Marsh, Hau, & Wen, 2004). Accordingly, we report the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR). For the CFI and TLI, values above .90 and .95 represent an adequate respectively good model fit (Hu & Bentler, 1999). For the RMSEA, values should be below .05 for a close fit, or between .05 and .08 for a reasonable fit (Browne & Cudeck, 1993). Regarding the SRMR, Hu and Bentler (1999) propose values below .08 as indicative of a good model fit.

A series of invariance models were used to test gender differences in the direct and indirect (i.e., mediated through academic self-concept) relations between math and verbal achievement and math and verbal test anxiety (including worry and emotionality) (Millsap, 2011). To this aim, gender was used as a grouping factor in Model 2 (i.e., the CFA model assuming separate factors for worry, emotionality, self-concept, and achievement in math and verbal domains). The taxonomy of invariance models (Models I1 to I12 in Table 1) started with a model of configural invariance assuming the same factor structure (i.e., the same number of factors defined by the same items; Model I1) across gender (Meredith, 1993). This model was increasingly expanded by additionally integrating invariant factor loadings (weak measurement invariance; Meredith, 1993; Model I2), invariant factor loadings and item intercepts (strong measurement invariance; Meredith, 1993; Model I3), and invariant factor loadings, item intercepts, and item uniquenesses (strict measurement invariance; Meredith, 1993; Model I4). In order to examine gender differences in the direct and indirect relations between math and verbal achievement and test anxiety, we subsequently tested the invariance of factor variances (Model I5). In the case of invariant factor variances, tests of invariant factor correlations can be realized by testing for invariant factor covariances (Marsh, 1994; Model I6).

The final aim was to gain insight into the generalizability of the mediation models across gender. First using worry as an outcome variable, we stated a regression model including gender-invariant factor loadings, item intercepts, item uniquenesses, and factor variances along with freely estimated paths for the relations among math and verbal achievement, self-concept, and worry (Model I7). This model was compared to a model in which the relations (i.e., covariances and path coefficients) among math and verbal achievement, self-concept, and worry were restricted to be of the same size for boys and girls (Model I8). These two modeling steps were repeated when using emotionality as an outcome variable (Models I9 and I10) and when using both worry and emotionality as outcome variables (Models I11 and I12).

In order to evaluate the invariance models, we follow the recommendation to examine changes in the descriptive goodness-of-fit indices between more and less restrictive models. According to the guidelines proposed by Cheung and Rensvold (2002) and Chen (2007), invariance can be assumed as long as the CFI and TLI do not drop more than .01, and the RMSEA does not increase more than .015. The proposed cut-off values for the various goodness-of-fit indices to evaluate the fit of latent models including nested models should rather be treated as guidelines instead of "golden rules". As well as considering a range of resulting fit indices, researchers are advised to base their final model evaluation on different types of information including the resulting parameter estimates, statistical conformity, and theoretical adequacy of the models (Marsh et al., 2004).

7. Results

The CFA model assuming separate factors for worry, emotionality, and achievement in the domains of math and German resulted in an adequate model fit (Model 1 in Table 1; CFI = .944; TLI = .938; RMSEA = .031; SRMR = .035), suggesting the integrity of the model and measured constructs. The factor correlations resulting from this model (Table S1 of the Online Supplements) indicated the domain specificity of test anxiety as test anxiety in math

and test anxiety in German were substantially, yet not perfectly, correlated to each other irrespective of whether the worry ($r = .633$, $p < .05$) or emotionality ($r = .642$, $p < .05$) components were considered. Moreover, the results supported the distinction between a worry and an emotionality component as the correlations were far from perfect within the math ($r = .575$, $p < .05$) and German domains ($r = .545$, $p < .05$). Both components of test anxiety demonstrated negative relations to achievement, but the worry component was found to be more highly correlated with achievement than the emotionality component. In addition, the pattern of relations between the test anxiety components and achievement was found to be domain-specific as worry (emotionality) in math was more highly correlated with math achievement ($r = -.183$, resp. $r = -.058$; both $p < .05$) than with German achievement ($r = -.044$, $p < .05$, resp. $r = .017$, ns). In parallel, worry (emotionality) in German displayed a higher negative association with German achievement ($r = -.178$, $p < .05$, resp. $r = -.046$, $p < .05$) compared to math achievement ($r = -.129$, $p < .05$, resp. $r = -.010$, ns).

7.1 Testing the GI/E Model

Subsequently, we tested the GI/E model including math and verbal achievements as predictor variables and both the worry and emotionality components of math and verbal test anxiety as outcome variables. This model (Table 2)[1] showed significant negative relations between achievement and test anxiety within the domains of math and German with higher relations for the worry (math: $\beta = -.258$, German: $\beta = -.161$; both $p < .05$) than for the emotionality component (math: $\beta = -.115$, German: $\beta = -.065$; both $p < .05$). When looking at the cross-domain relations, German achievement was found to be positively related to the worry ($\beta = .119$, $p < .05$) as well as to the emotionality component ($\beta = .089$, $p < .05$) of math test anxiety, with a higher relation between German achievement and math worry than between German achievement and math emotionality. Math achievement was not significantly related to the worry and emotionality components of German test anxiety. Math and German achievements were highly correlated ($r = .632$, $p < .05$). The worry ($r = .645$, $p <$

.05) and emotionality ($r = .649$, $p < .05$) components of test anxiety were also substantially

correlated across the math and verbal domains.

7.2 Mediation through Academic Self-concept

As a prerequisite to testing the mediated GI/E model, we evaluated a measurement

model (Model 2 in Table 1) including separate factors for worry, emotionality, self-concept,

and achievement in math and German. The goodness-of-fit indices indicated the adequacy of

this model (CFI = .935; TLI = .929; RMSEA = .031; SRMR = .031). The factor correlations

resulting from this model (Table 3) supported the domain specificity of academic self-concept

and were in line with previous research as math self-concept was more highly related to math

achievement ($r = .360$, $p < .05$) than to German achievement ($r = .056$, $p < .05$), whereas

German self-concept was more highly related to German achievement ($r = .333$, $p < .05$) than

to math achievement *($r = .120$, $p < .05$)*. Furthermore, math self-concept demonstrated a

higher negative relation to test anxiety in math (worry: $r = -.534$, emotionality: $r = -.350$; both

$p < .05$) than to test anxiety in German (worry: $r = -.293$, emotionality: $r = -.205$; both $p <$

.05). In parallel, German self-concept showed a higher relation to test anxiety in German

(worry: $r = -.480$, emotionality: $r = -.278$; both $p < .05$) compared to test anxiety in math

(worry: $r = -.162$, emotionality: $r = -.105$; both $p < .05$). Academic self-concept was more

highly correlated with the worry component of test anxiety than with the emotionality

component. This pattern of findings could be observed within matching domains [i.e., the

relation between math (German) self-concept and math (German) worry was higher (math: $r =$

-.534, German: $r = -.480$; both $p < .05$) than the relation between math (German) self-concept

and math (German) emotionality (math: $r = -.350$, German: $r = -.278$; both $p < .05$)] and

across non-matching domains [i.e., the relation between math self-concept and German worry

($r = -.293$, $p < .05$) was higher than the relation between math self-concept and German

emotionality ($r = -.205$, $p < .05$), and the relation between German self-concept and math

worry ($r = -.162$, $p < .05$) was higher than the relation between German self-concept and math emotionality ($r = -.105$, $p < .05$)].

Regarding the achievement–self-concept relations, the original I/E model pattern was replicated in an additional latent regression model[2]. Based on these findings supporting the integrity of the used self-concept measures, we turned to the mediation models. In a first model (Model 3 in Table 1), the worry component was considered as an outcome variable. Hence, the relation between math achievement on the one hand and math and German worry on the other hand is assumed to be mediated through math self-concept, and the relation between German achievement on the one hand and math and German worry on the other hand is assumed to be mediated through German self-concept (Figure 1a). In this model, all indirect effects were significant and only one direct effect (i.e., the effect from German achievement to German worry; $\beta = -.060$, $p < .05$) remained significant (Table S2 of the Online Supplements). Both indirect within-domain relations were negative (the relation between math achievement and math worry mediated through math self-concept: $\beta = -.284$, $p < .05$; the relation between German achievement and German worry mediated through German self-concept: $\beta = -.184$, $p < .05$). This was expected since these relations resulted from positive within-domain relations between achievement and self-concept and negative within-domain relations between self-concept and worry. Both indirect cross-domain relations were positive (the relation from math achievement to German worry mediated through German self-concept: $\beta = .062$, $p < .05$; the relation from German achievement to math worry mediated through math self-concept: $\beta = .184$, $p < .05$) due to the negative relations between achievement and self-concept of non-matching domains (see the original I/E model) and between self-concept and worry of matching domains. Moreover, the original I/E model was retained as obvious in the resulting achievement–self-concept relations which were positive within domains (math: $\beta = .541$, German: $\beta = .425$; both $p < .05$) and negative across domains (math achievement and German self-concept: $\beta = -.144$; German achievement and math self-

concept: $\beta = -.281$; both $p < .05$). Finally, self-concept and worry were found to be negatively related with stronger relations within matching domains (math: $\beta = -.524$, German: $\beta = -.433$; both $p < .05$) than across non-matching domains (math self-concept and German worry: $\beta = .257$; German self-concept and math worry: $\beta = -.095$; both $p < .05$).

When using emotionality as an outcome variable (Figure 1b; Model 4 in Table 1), all indirect effects were significant indicating that the within-domain and cross-domain relations between achievement and emotionality were mediated through academic self-concept (Table S2 of the Online Supplements). The direct cross-domain relation between math achievement and German emotionality remained significant ($\beta = .112$, $p < .05$) indicating that this relation was partially mediated through German self-concept. The direct within-domain relation between math achievement and math emotionality was significantly positive ($\beta = .081$, $p < .05$). This result contradicts the theoretically assumed and consistently demonstrated negative associations between achievement and test anxiety within domains, but might originate from a suppression effect in the mediation model (Shieh, 2006). Corresponding to the original I/E model, the within-domain relations between achievement and self-concept were positive, while the cross-domain relations were negative. Self-concept and emotionality were negatively related to each other and these relations were higher within matching than across non-matching domains.

Finally, we stated a GI/E mediation model (Model 5 in Table 1) in which worry and emotionality in math and German both served as outcome variables. Hence, this complex model included math and verbal achievements as predictor variables, math and verbal self-concepts as mediator variables, and worry and emotionality in math and German as outcome variables. The results (Table 4) show that among the eight possible direct relations, only three relations reached statistical significance. Two of them correspond to the findings from the non-mediated models and theoretical assumptions indicating negative within-domain and positive between-domain relations (German achievement and German worry: $\beta = -.061$, $p <$

.05; math achievement and German emotionality: β = .111, $p < .05$). However, the significantly positive relation between math achievement and emotionality in math (β = .081, $p < .05$) again did not fit to the previous findings from the non-mediated models and to theoretical assumptions, but might rather originate from a suppression effect (Shieh, 2006). All indirect effects were statistically significant and in the expected direction. Hence, the relations between achievement on the one hand and worry and emotionality on the other hand mediated through domain-specific self-concepts were negative within matching domains (i.e., within the math and verbal domains), but positive across non-matching domains.

7.3 Invariance across Gender

The CFA model assuming separate factors for worry, emotionality, self-concept, and achievement in math and German (cf. Model 2 in Table 1) was found to display measurement invariance across gender as the changes in the descriptive goodness-of-fit indices remained below ΔCFI/TLI $\leq$ -.01 and ΔRMSEA $\leq$ +.015 across the models of configural invariance (Model I1), weak measurement invariance (i.e., invariant factor loadings, Model I2), strong measurement invariance (i.e., invariant factor loadings and item intercepts, Model I3) and strict measurement invariance (i.e., invariant factor loadings, item intercepts, and item uniquenesses, Model I4).

When further adding invariance constraints on the factor variances (Model I5 in Table 1), the TLI and RMSEA did not change relative to the preceding model, and the CFI only dropped by CFIΔ = -.001. This result supports equal-sized factor variances for boys and girls, allowing for testing the invariance of factor correlations by restricting the factor covariances to invariance (Marsh, 1994). As such, each possible pair of factor covariances was constrained to be invariant across gender in the next step (Model I6). In this case, the values for the CFI and RMSEA remained stable and the TLI even increased (TLIΔ = +.001) due to greater model parsimony. Thus, the findings supported the invariance of factor covariances

indicating that the relations between worry, emotionality, self-concept, and achievement factors regarding math and German did not vary between boys and girls.

The final set of models (Models I7 to I12 in Table 1) examined the invariance of the mediation models. The fit did not decline between models where the direct and indirect relations were freely estimated across gender (Model I7 for worry, Model I9 for emotionality, Model I11 for both worry and emotionality as outcomes) and models with invariance constraints on these relations (Model I8 for worry, Model I10 for emotionality; Model 12 for both worry and emotionality as outcomes).[3] Hence, the mediation models seemed to be invariant across gender when considering worry and emotionality separately as well as jointly as outcome variables.

8. Discussion

In research on test anxiety, one approach addresses domain-specific facets of test anxiety (Goetz et al., 2010; Goetz, Frenzel et al., 2006, 2007; Goetz, Pekrun et al., 2006; Marsh & Yeung 1996; Sparfeldt et al., 2005). In research on dimensional comparison processes (Möller & Marsh, 2013), the GI/E model (Möller et al., 2015) has been established inspiring researchers to investigate outcome variables beyond academic self-concepts that might be influenced by dimensional achievement comparisons. Based on and combining these two theoretical strands, the present study aimed to examine whether the GI/E model applies to domain-specific facets of test anxiety and thus whether dimensional achievement comparisons are at play in the formation of domain-specific facets of test anxiety.

8.1 The GI/E Model for Test Anxiety

In general, the results indicated that the GI/E model is applicable to test anxiety. The GI/E model pattern was first evident in negative within-domain relations between achievement and test anxiety, replicating previous findings on negative relations between achievement and test anxiety in specific domains such as math (Frenzel et al., 2007a; Hembree, 1990; Ho et al., 2000; Ma, 1999; Pajares & Graham, 1999) and indicating that

social achievement comparisons are at play in the formation of domain-specific test anxiety. Second, the GI/E model pattern became evident in a positive cross-domain relation between verbal achievement and test anxiety in math. This finding suggests that dimensional achievement comparison processes might operate in the formation of test anxiety in math.

Math achievement, however, was not significantly related to test anxiety in German, indicating that dimensional comparisons between math and verbal achievements might be less important for the formation of test anxiety in verbal domains. This finding corresponds to findings from other studies showing that self-perceptions (including self-concept, motivation, and emotion constructs) in verbal domains are less subject to dimensional achievement comparison processes than self-perceptions addressing math. Correspondingly, the meta-analysis on the original I/E model published by Möller et al. (2009) demonstrated higher negative cross-paths leading from verbal achievement to math self-concept than from math achievement to verbal self-concept (see also Schurtz et al., 2014 showing more negative relations between English performance and math interest than between math performance and English interest). This might be explained by students' different experiences regarding math and verbal domains. Students' experiences with math are mainly restricted to the school context, while students' experiences with the verbal domain go beyond the school context, spanning a range of academic and non-academic life domains. As such, students might use information only obtained in the school context such as their math and verbal accomplishments to establish their self-perceptions, motivation, and emotions in math. On the other hand, when forming verbal self-perceptions, motivation, and emotions, students might rely on a wide range of information and feedback which may not only refer to their school accomplishments but might also encompass verbal skills in non-academic settings. Hence, the formation of verbal motivation, self-perceptions, and emotions including test anxiety might be more multi-faceted. This assumption is further corroborated by the observation made here and in previous studies (Möller et al., 2009; Schurtz et al., 2014) that students seem also to rely

less on social achievement comparison processes in the formation of verbal outcomes (self-concept, interest, and test anxiety). This becomes evident from lower within-domain relations between achievement and outcomes (e.g., self-concept, motivation, and emotion) within the verbal than within the math domain. Hence, in sum, students seem to base their self-perceptions in the verbal domain to a lesser extent on achievement comparison processes in general.

8.2 The Differentiation between Worry and Emotionality

Applying the GI/E model to test anxiety, a similar pattern of results was found for the worry and emotionality components. However, relatively stronger achievement relations were found for worry than for emotionality and this finding pertained to within-domain as well as to cross-domain relations. This observation corresponds to previous findings demonstrating stronger achievement relations for worry than for emotionality (Gierl & Todd, 1996; Hembree, 1988; Hong & Karstensson 2001; Seipp, 1991; Stöber, 2004; Zeidner, 2007; Zeidner & Schleyer, 1999) and extends it to domain-specific facets of test anxiety and to within-domain and cross-domain achievement relations. Hence, worry and emotionality can be conceptualized as separate components of domain-unspecific test anxiety (Everson et al., 1991; Stöber, 2004; Zeidner, 2007) and domain-specific test anxiety (Sparfeld et al., 2005; this study) showing differential relations with achievement.

The GI/E model for test anxiety revealed a high correlation between math and verbal test anxiety. More concretely, the correlation between math and verbal worry and the correlation between math and verbal emotionality was found to be of similar size as the correlation between math and verbal achievement. This finding is interesting with regard to both theoretical strands of research combined in this study, i.e., DCT including the GI/E model, and research on test anxiety. A core assumption of the original I/E model refers to the negligible correlation between math and verbal self-concepts which should be at least lower than the correlation between math and verbal achievements (Marsh, 1986, 1990). The broader

GI/E model (Möller et al., 2015) established in DCT (Möller & Marsh, 2013) is more liberal in this regard since it admits the possibility of dimensional comparison processes even in the case of a substantial correlation between outcome variables (i.e., test anxiety in this study). Hence, the finding of a substantial correlation between domain-specific test anxiety factors supports the GI/E model as a theoretical framework for investigating the consequences of dimensional achievement comparison processes on a wide range of outcome variables. Furthermore, the found substantial correlation between math and verbal test anxiety matches previous findings indicating that the nature of test anxiety is less domain-specific than academic self-concept or even other academic emotions such as enjoyment (Goetz, Frenzel et al., 2006; Zeidner, 1998).

8.3 Mediation through Academic Self-concept

The findings resulting from this study further argue for the existence of indirect effects since both the within-domain relations as well as the cross-domain relations between achievement and test anxiety were found to be mediated through academic self-concept irrespective of whether the worry, the emotionality, or both the worry and emotionality components were considered. The mediation yet seemed to be stronger when using worry rather than emotionality as an outcome variable, as only one direct effect was maintained in the model for worry while two direct effects were retained in the model for emotionality. This finding substantiates the assumption of a stronger conceptual overlap between worry and academic self-concept, which is plausible given the cognitive and self-evaluative nature of both constructs (Zeidner & Schleyer, 1999). The latter conjecture is further substantiated by the demonstrated higher within-domain and cross-domain relations between worry and self-concept than between emotionality and self-concept.

8.4 Generalizability across Gender

Boys and girls were not found to differ in direct and indirect relations between achievement and test anxiety including worry and emotionality. This finding indicates that

boys and girls use the same mechanisms to establish domain-specific facets of test anxiety, including social and dimensional achievement comparison processes. This conclusion corresponds to previous findings demonstrating gender invariance of the original I/E model pattern (Möller et al., 2009; Skaalvik & Rankin, 1990).

8.5 Practical and Theoretical Implications

In sum, the present study offers interesting and innovative findings which are associated with various practical and theoretical implications. On a practical level, efforts have been invested in discovering effective interventions to remedy students' test anxiety (Ergene, 2003; Hembree, 1988). Respective approaches might benefit from the knowledge that dimensional achievement comparison processes are involved in the formation of domain-specific facets of test anxiety. Hence, applied researchers and practitioners are advised not to only focus on students' domain-unspecific, global level test anxiety or on only one content domain such as test anxiety in math. They should rather take into account that students' accomplishments and experiences in other domains also matter in the establishment of domain-specific test anxiety. Interventions might further benefit from the finding of the separation between a worry component and an emotionality component within domain-specific facets of test anxiety. Hence, both components should be considered in intervention approaches which might require differential enhancement strategies. Against the background of the mediation models supported in this study, it is also recommended to combine interventions targeting students' test anxiety and academic self-concept interventions (O'Mara, Marsh, Craven, & Debus, 2006). In view of the found gender-invariant relations among achievement, self-concept and test anxiety (including worry and emotionality), the same intervention approaches can be applied to boys and girls, although girls seem to have a higher need for test anxiety interventions given their higher mean levels as documented in other previous studies (Cassadey & Johnson, 2002; Chapell et al., 2005; Else-Quest et al.,

2010; Frenzel et al., 2007a; Hembree, 1990; Hong & Karstensson, 2002; Martin, 2007; Miller & Bichsel, 2004).

On a theoretical level, this study contributes to research on dimensional comparison processes and research on test anxiety. In essence, it provides further empirical support for the GI/E model and illustrates that dimensional achievement comparison processes are at play in the formation of a variety of outcome variables including test anxiety. Hence, the domain-specific approach to test anxiety can benefit from DCT and the related GI/E model in order to explain the formation of domain-specific facets of test anxiety. Moreover, the found mediation through domain-specific academic self-concepts gives insight into the mechanism underlying the consistently observed relation between test anxiety and achievement within and across domains.

8.6 Limitations and Directions for Future Research

A general limitation of this study refers to its cross-sectional and correlational design. Hence, the findings do not allow any temporal or even casual interpretations. The examination of the relations between domain-specific achievement, self-concept, and test anxiety would thus benefit from longitudinal studies. In research on the I/E model, the reciprocal internal/external frame of reference (RI/E) model (Möller, Retelsdorf, Köller, & Marsh, 2011; Niepel, Brunner, & Preckel, 2011) has been established to describe within-domain and cross-domain relations between achievement and self-concept measures across time. Following the logic of the GI/E model (Möller et al., 2015), the RI/E model should be extended to other motivational variables than academic self-concept. A longitudinal approach including multiple measurement waves which preferably span various grade levels would also allow insights into age-dependent variations in the patterns of findings. The sample of this study only consisted of seventh grade students so it was not possible to take age effects into account. Besides pursuing the matter of generalizability across students' age as an avenue for future research, it is necessary to test the generalizability of the present findings across student

samples including students from different immigrant background, socioeconomic status (SES), achievement track, or even cultures or educational systems. Replication of these findings with more recent data sets seems also needed since the data are from 1991/1992 and thus rather dated.

From the perspective of research on dimensional comparison processes, a limitation of this study is that it only took math and verbal domains into account. This approach fits the traditional I/E model (Marsh, 1986) but a larger variety of school subjects was considered in recent publications allowing the examination and juxtaposition of contrast and assimilation effects originating from dimensional comparison processes (Marsh et al., 2014; Marsh, Lüdtke et al., 2015; Möller, Streblow, Pohlmann, & Köller, 2006). Moreover, only school grades were used as achievement indicators. School grades and standardized achievement test scores present the two most commonly used achievement indicators, but they are associated with differential characteristics and bear differentially sized relations to motivational constructs including self-concept (Marsh et al., 2014; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005). Future studies should combine and compare both types of achievement indicators to examine whether they are similarly subject to dimensional achievement comparison processes in the formation of test anxiety. Finally, from the perspective of the GI/E model (Möller et al., 2015), this study examines test anxiety as another outcome variable (i.e., as a variable which is influenced by dimensional comparison processes) beyond academic self-concept as stated in the original I/E model (Marsh, 1986, 1990; Möller et al., 2009). More future research, however, is also needed with respect to further possible predictor variables since in this study, similar to the original I/E model, domain-specific achievements were still assumed as the predictor variables, i.e., as the variables that are compared across domains.

From the perspective of research on test anxiety, a limitation of the present study is that it did not take the differentiation between state and trait anxiety into account (Spielberger

et al., 1970; Zeidner, 1998) although state and trait test anxiety might reveal differential

achievement relations and gender effects (Goetz et al., 2013). The items applied in the present

study for measuring test anxiety are hard to unambiguously categorize as measures for state

versus measures for trait anxiety but they seem to encompass both trait and state components.

In fact, students are asked to respond to the items when remembering a test situation in math

and verbal classes. The focus on a specific test situation might be evocative of a state

measure. On the other hand, students had to respond to this item retrospectively and might

think of different test situations bringing this measure closer to a trait measure. Hence, an

avenue for future research might be to replicate the analyses when using test anxiety measures

which can clearly be classified as trait or state measures. For measuring students' state test

anxiety related to math respectively German, students should be directly asked in their math

or German classes, preferably immediately before and during a test situation (Goetz et al.,

2013; Goetz, Preckel, Pekrun, & Hall, 2007). For measuring students' trait test anxiety related

to math respectively German, students should rate their habitual experience of test anxiety

associated with math and German as school subjects. Moreover, in this study, the students

were asked to remember a test/exam situation for the measurement of test anxiety. Thus, test

anxiety is assessed with a special focus on test anxiety related to test/exam situations.

However, other situations or academic settings in which academic emotions including test

anxiety can occur should also be taken into account. For instance, the Achievement Emotions

Questionnaire (AEQ; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011) measures test anxiety

during class, while studying, and when taking tests or exams. Similarly, Goetz et al. (2012)

distinguished between anxiety while doing homework and anxiety experience during class.

Regarding the internal structure of test anxiety, domain-unspecific and domain-specific

approaches might be integrated so that they coexist rather than being mutually exclusive.

Recent studies (Gogol et al., 2016, 2017) provided evidence of a bifactor model approach to

test anxiety with a general factor at the apex, and domain-specific factors related to math,

French, and German as nested subject-specific factors. The general factor has been found to be strong in terms of its explained variance. This result matches findings from this and other studies (Goetz, Frenzel et al., 2006; Zeidner, 1998; see above) according to which test anxiety shows a weaker pattern of domain specificity compared to other academic emotions. These novel advances in modeling the internal structure of test anxiety should be considered when testing within-domain and cross-domain achievement relations.

To sum up, the present study offers strong support for the role of social and dimensional achievement comparison processes in the formation of students' domain-specific test anxiety. Still, students' learning environments such as teachers' instruction practices or classroom management (Frenzel, Pekrun, & Goetz, 2007b; Hattie 2009), parents' reactions, and peer feedback or peer norms (Gunderson, Ramirez, Levine, & Beilock, 2012; Zeidner & Schleyer, 1999) might also impact on individual students' levels of test anxiety and should thus be addressed by future studies. As such, we linked the theoretical approach of DCT (Möller & Marsh, 2013) with research on test anxiety, thereby proving new mechanisms of the formation of domain-specific worry and emotionality components of test anxiety. This study thus offers interesting insights into research on dimensional comparison processes and research on test anxiety, but also indicates directions for future research.

9. Endnote

[1] This model is statistically equivalent to the confirmatory factor analysis model (Model 1 in Table 1) assuming separate factors for worry, emotionality, and achievement in math and German since the factor correlations were only replaced by path coefficients. Hence, Models 1 and 2 result in the same fit: χ^2 (564) = 3393.758, CFI = .944, TLI = .938, RMSEA = .031, SRMR = .035.

[2] The fit of the original I/E model using math and verbal achievements as predictor variables and math and verbal self-concepts as outcome variables was good: χ^2 (45) = 587.039, CFI = .970, TLI = .955, RMSEA = .048, SRMR = .043. Corresponding to the original I/E model

(Marsh, 1990; Möller et al., 2009), math achievement demonstrated a positive relation to math

self-concept (β = .541, p < .05), but a negative relation to German self-concept (β = -.150, p <

.05), and German achievement showed a positive relation to German self-concept (β = .427, p

< .05), but a negative relation to math self-concept (β = -.286, p < .05). Math and verbal

achievements were found to be substantially correlated (r = .631, p < .05), whereas math and

verbal self-concepts shared only a small relation (r = .225, p < .05).

[3] The indirect relations between achievement and test anxiety mediated through academic self-

concept within and across the math and verbal domains were specified by using the "model

constraint" option in Mplus.

10. References

Ahmed, W., Minnaert, A., Kuyper, H., & van der Werf, G. (2012). Reciprocal relationships

between math self-concept and math anxiety. *Learning and Individual Differences, 22*,

385–389.

Arens, A.K., & Möller, J. (2016). Dimensional comparisons in students' perceptions of the

learning environment. *Learning and Instruction, 42,* 22–30.

Ashcraft, M.H. (2002). Math anxiety: Personal, educational, and cognitive consequences.

Current Directions in Psychological Science, 11, 181–185.

Ashcraft, M.H., & Kirk, E.P. (2001). The relationships among working memory, math

anxiety, and performance. *Journal of Experimental Psychology: General, 130*, 224–

237.

Bandalos, D.L., Yates, K., & Thorndike-Christ, T. (1995). Effects of math self-concept,

perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal

of Educational Psychology, 87*, 611–623.

Baumert, J., Roeder, P.M., Gruehn, S., Heyn, S., Köller, O., Rimmele, R., . . . Seipp, B.

(1996). Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). In

K.-P. Treumann, G. Neubauer, R. Möller, & J. Abel (Eds.), *Methoden und*

Anwendungen empirischer pädagogischer Forschung (pp. 170–180). Münster: Waxmann.

Beauducel, A., & Herzberg P.Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling, 1,* 186–203.

Beidel, D.C., & Turner, S.M. (1988). Comorbidity of test anxiety and other anxiety disorders in children. *Journal of Abnormal Child Psychology, 16*, 275–287.

Bodas, J., & Ollendick, T.H. (2005). Test anxiety: A cross-cultural perspective. *Clinical Child and Family Psychology Review, 8*, 65–88.

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

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K.A. Bollen, & J.S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Cassady, J.C. (2004). The influence of cognitive test anxiety across the learning–testing cycle. *Learning and Instruction, 14*, 569–592.

Cassady, J.C., & Johnson, R.E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology, 27*, 270–295.

Chapell, M.S., Blanding, Z.B., Silverstein, M.E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*, 268–274.

Chen, F.F. (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504.

Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Elliot, A.J., & McGregor, H.A. (1999). Test anxiety and the hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 76*, 628–644.

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

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

Ergene, T. (2003). Effective interventions on test anxiety reduction: A meta-analysis. *School Psychology International, 24*, 313–328.

Everson, H.T., Millsap, R.E., & Rodriguez, C.M. (1991). Isolating gender differences in test anxiety: A confirmatory factor analysis of the test anxiety Inventory. *Educational and Psychological Measurement, 51*, 243–251.

Eysenck, M.W., Derakshan, N., Santos, R., & Calvo, M.G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*, 336–353.

Frenzel, A.C., Pekrun, R., & Goetz, T. (2007a). Girls and mathematics - A "hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education, 22*, 497–514.

Frenzel, A.C., Pekrun, R., & Goetz, T. (2007b). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*, 478–493.

Gierl, M.J., & Todd, R.W. (1996). A confirmatory factor analysis of the Test Anxiety Inventory using Canadian high school students. *Educational and Psychological Measurement, 56*, 315–324.

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N.C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science, 24*, 2079–2087.

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*, 44–58.

Goetz, T., Frenzel, A.C., Hall, N.C., & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology, 33*, 9–33.

Goetz, T., Frenzel, A.C., Pekrun, R., & Hall, N.C. (2006). The domain specificity of academic emotional experiences. *The Journal of Experimental Education, 75*, 5–29.

Goetz, T., Frenzel, A.C., Pekrun, R., Hall, N.C., & Lüdtke, O. (2007). Between- and within-domain relations of students' academic emotions. *Journal of Educational Psychology, 99*, 715–733.

Goetz, T., Nett, U. E., Martiny, S.E., Hall, N.C., Pekrun, R., Dettmers, S., & Trautwein, U. (2012). Students' emotions during homework: Structures, self-concept antecedents, and achievement outcomes. Learning and Individual Differences, 22, 225–234.

Goetz, T., Pekrun, R., Hall, N.C., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*, 289–308.

Goetz, T., Preckel, F., Pekrun, R., & Hall, N.C. (2007). Emotional experiences during test taking: Does cognitive ability make a difference? Learning and Individual Differences, 17, 3–16.

Gogol, K., Brunner, M., Martin, R., Preckel, F., & Goetz, T. (2017). Affect and motivation within and between school subjects: Development and validation of an integrative structural model of academic self-concept, interest, and anxiety. Contemporary Educational Psychology, 49, 46–65.

Gogol, K., Brunner, M., Preckel, F., Goetz, T., & Martin, R. (2016). Developmental dynamics of general and school-subject-specific components of academic self-concept, academic interest, and academic anxiety. Frontiers in Psychology, 7, 356.

Graham, J.W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549–576.

Gunderson, E.A., Ramirez, G., Levine, S. C., & Beilock, S.L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles, 66*, 153–166.

Guo, J., Marsh, H.W., Parker, P.D., Morin, A.J.S., & Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in science: Dimensional comparison processes and expectancy-by-value interactions. Learning and Instruction, 49, 81–91.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* London: Routledge.

Helmke, A. (1992). *Selbstvertrauen und schulische Leistungen*. Göttingen: Hogrefe.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58*, 47–77.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Reserach in Mathematics Education, 21*, 33–46.

Ho, H.-Z., Senturk, D., Lam, A.G., Zimmer, J.M., Hong, S., Okamoto, Y., & Chiu, S.-Y. (2000). The affective and cognitive dimensions of math anxiety: A cross-national study. *Journal for Research in Mathematics Education, 31*, 362–379.

Hodapp, V., Laux, L., & Spielberger, C. (1982). Theorie und Messung der emotionalen und kognitiven Komponente der Prüfungsangst. *Zeitschrift für Differentielle und Diagnostische Psychologie, 3*, 169–184.

Hong, E. (1999). Test anxiety, perceived test difficulty, and test performance: Temporal patterns of their effects. *Learning and Individual Differences, 11*, 431–447.

Hong, E., & Karstensson, L. (2002). Antecedents of state test anxiety. *Contemporary Educational Psychology, 27*, 348–367.

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,* 1–55.

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology, 49,* 505–528.

Jerusalem, M. (1984). *Selbstbezogene Kognitionen in schulischen Bezugsgruppen. Eine Längsschnittstudie [Self-related cognitions in school reference groups: A longitudinal study]*. Berlin, Germany: Free University Press.

Jopt, U.-J. (1978). *Selbstkonzept und Ursachenerklärung in der Schule. Zur Attribuierung von Schulleistungen [Self-concept and causal attributions in schools. On attributions of school achievement]*. Bochum, Germany:

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

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences, 19,* 355–365.

Liebert, R.M., & Morris, L.W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports, 20*, 975–978.

Ma, X. (1999). A meta-analysis of the relationship between anxiety towards mathematics and achievement in mathematics. *Journal for Research in Mathematics Education, 30,* 520–540.

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

Marsh, H.W. (1988). The context specificity of math and English anxieties: The High School and Beyond Study. *Anxiety Research, 1*, 137–149.

Marsh, H.W. (1990). Influences of internal and external frames of reference on the formation of math and English self-concepts. *Journal of Educational Psychology, 82*, 107–116.

Marsh, H.W. (1994). Confirmatory factor analysis models of factorial invariance: A multifaceted approach. *Structural Equation Modeling, 1*, 5–34.

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

Marsh, H.W., Abduljabbar, A.S., Parker, P., Morin, A.J.S., Abdelfattah, F., Nagengast, B., . . . Abu-Hillal, M.M. (2015). The internal/external frame of reference model: Age-cohort and cross-national differences in paradoxical relations between TIMSS mathematics and science achievement, self-concept and intrinsic motivation. *American Educational Research Journal, 52,* 168–202.

Marsh, H.W., & Craven, R.G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective. Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1,* 133–163.

Marsh, H.W., & Hau, K.-T. (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*, 56–67.

Marsh, H.W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320–341.

Marsh, H.W., Kuyper, H., Seaton, M., Parker, P.D., Morin, A.J.S., Möller, J., & Abduljabbar, A.S. (2014). Dimensional comparison theory: An extension of the internal/external frame of reference effect on academic self-concept formation. *Contemporary Educational Psychology, 39,* 326–341

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

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

Marsh, H.W., & Yeung, A.S. (1996). The distinctiveness of affects in specific school subjects: An application of confirmatory factor analysis with the National Educational Longitudinal Study of 1988. *American Educational Research Journal, 33*, 665–689.

Martin, A.J. (2007). Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology, 77,* 413–440.

Meece, J.L., Wigfield, A., & Eccles, J.S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82*, 60–70.

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

Miller, J.W. (2000). Exploring the source of self-regulated learning: The influence of internal and external comparisons. *Journal of Instructional Psychology, 26,* 47–52.

Miller, H., & Bichsel, J. (2004). Anxiety, working memory, gender, and math performance. *Personality and Individual Differences, 37*, 591–606.

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

    Routledge.

Möller, J., Müller-Kalthoff, H., Helm, F., Nagy, N., & Marsh, H.W. (2015). The generalized

    internal/external frame of reference model: An extension to dimensional comparison

    theory. *Frontline Learning Research, 3*, 37–47.

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

    120,* 544–560.

Möller, J., Pohlmann, B., Köller, O., & Marsh, H.W. (2009). A meta-analytic path analysis of

    the internal/external frame of reference model of academic achievement and academic

    self-concept. *Review of Educational Research, 79*, 1129–1167.

Möller, J., Retelsdorf, J., Köller, O. & Marsh, H.W. (2011). The reciprocal internal/external

    frame of reference model: An integration of models of relations between academic

    achievement and self-concept. *American Educational Research Journal, 48*, 1315–

    1346.

Möller, J., Streblow, L., Pohlmann, B., & Köller, O. (2006). An extension of the

    internal/external frame of reference model to two verbal and numerical domains.

    European Journal of Psychology of Education, 21, 467–487.

Möller, J. Zimmermann, F., & Köller, O. (2014). The reciprocal internal/external frame of

    reference model using grades and test scores. *British Journal of Educational

    Psychology, 84*, 591–611.

Morris, L.W., Davis, M.A., & Hutchings, C.H. (1981). Cognitive and emotional components

    of anxiety: Literature review and a revised worry-emotionality scale. *Journal of

    Educational Psychology, 73*, 541–555.

Muthén, L.K., & Muthén, B.O. (1998-2015). *Mplus User's Guide. Seventh Edition.* Los

    Angeles, CA: Muthén & Muthén.

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

O'Mara, A.J., Marsh, H.W., Craven, R.G., & Debus, R.L. (2006). Do self-concept
interventions make a difference? A synergistic blend of construct validation and meta-
analysis. *Educational Psychologist, 41,* 181-206.

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics
performance of entering middle school students. *Contemporary Educational
Psychology, 24,* 124–139.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions,
corollaries, and implications for educational research and practice. *Educational
Psychology Review, 18*, 315–341.

Pekrun, R., Goetz, T., Frenzel, A.C., Barchfeld, P., & Perry, R. P. (2011). Measuring
emotions in students' learning performance: The Achievement Emotions
Questionnaire (AEQ). *Contemporary Educational Psychology, 36,* 36–48.

Pekrun, R., Goetz, T., Perry, R.P., Kramer, K., Hochstadt, M., & Molfenter, S. (2004).
Beyond test anxiety: Development and validation of the Test Emotions Questionnaire
(TEQ). *Anxiety, Stress & Coping, 17*, 287–316.

Pekrun, R., & Stephens, E.J. (2010). Achievement emotions: A control-value approach. *Social
and Personality Psychology Compass, 4*, 238–255.

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

Schnabel, K. (1998). *Prüfungsangst und Lernen: Empirische Analysen zum Einfluss
fachspezifischer Leistungsängstlichkeit auf schulischen Lernfortschritt*. Münster [u.a.]:
Waxmann Verlag.

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

Schutz, P.A., & Davis, H.A. (2000). Emotions and self-regulation during test taking. *Educational Psychologist, 35*, 243–256.

Seipp, B. (1991) Anxiety and academic performance: A meta-analysis of findings, *Anxiety Research, 4*, 27–41.

Shieh, G. (2006). Suppression situations in multiple linear regression. *Educational and Psychological Measurement, 66*, 435–447.

Skaalvik, E.M., & Rankin, R.J. (1990). Math, verbal, and general academic self-concept: The internal/external frame of reference model and gender differences in self-concept structure. *Journal of Educational Psychology, 82*, 546–554.

Sparfeldt, J.R., Schilling, S.R., Rost, D.H., Stelzl, I., & Peipert, D. (2005). Leistungsängstlichkeit: Facetten, Fächer, Fachfacetten? *Zeitschrift für Pädagogische Psychologie, 19*, 225–236.

Spielberger, C.D., Gorsuch, R.L., & Lushene, R.E. (1970). *Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Stäbler, F., Dumont, H., Becker, M., & Baumert, J. (2017). What happens to the fish's achievement in a little pond? A simultaneous analysis of class-average achievement effects on achievement and academic self-concept. *Journal of Educational Psychology, 109,* 191–207.

Stöber, J. (2004). Dimensions of test anxiety: Relations to ways of coping with pre-exam anxiety and uncertainty. *Anxiety, Stress, & Coping, 17,* 213–226.

Stöber, J., & Pekrun, R. (2004). Advances in test anxiety research. *Anxiety, Stress & Coping, 17*, 205–211.

Swann, W.B., Chang-Schneider, C., & Larsen McClarty, K. (2007). Do people's self-views

matter? Self-concept and self-esteem in everyday life. *American Psychologist, 62*, 84–

94.

Valentine, J.C., DuBois, D.L., & Cooper, H. (2004). The relation between self-beliefs and

academic achievement: A meta-analytic review. *Educational Psychologist, 39*, 111–

133.

Wigfield, A., & Meece, J.L. (1988). Math anxiety in elementary and secondary school

students. *Journal of Educational Psychology, 80*, 210–216.

Zeidner, M. (1998). *Test anxiety: The state of the art.* New York: Plenum Press.

Zeidner, M. (2007). Test anxiety in educational contexts: Concepts, findings, and future

directions. In P. Schutz, & R. Pekrun (Eds.). *Emotion in education* (pp. 165–184). San

Diego, CA: Academic Press.

Zeidner, M., & Schleyer, E.J. (1999). Test anxiety in intellectually gifted school students.

Anxiety, Stress & Coping, 12, 163–189.

Zimmermann, F., Möller, J., & Köller, O. (2017). When students doubt their teachers'

diagnostic competence: Moderation in the internal/external frame of reference model.

Journal of Educational Psychology. Online first: doi: 10.1037/edu0000196

Standardized results of the mediation model (Model 3) for worry
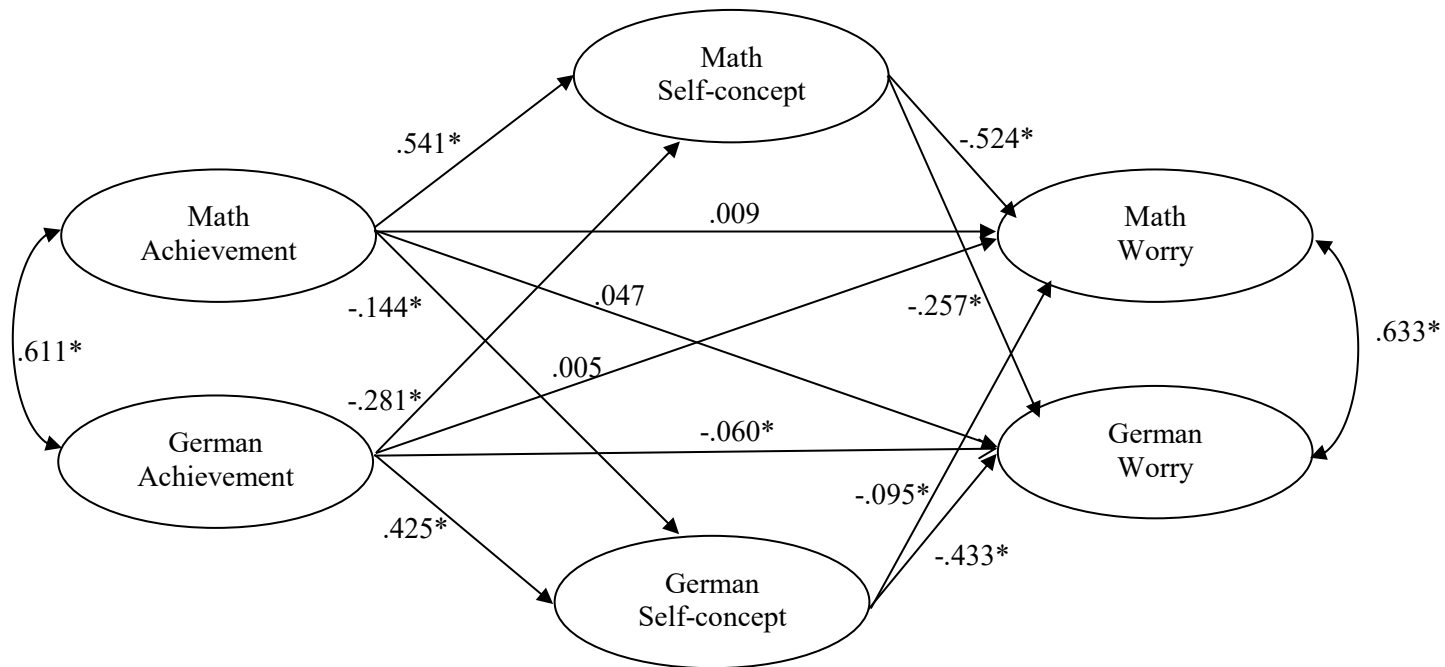
Note. The (standardized) indirect effects are all statistically significant: Math achievement → math self-concept → math worry: β = -.284*; Math achievement → German self-concept → German worry: β = .062*; German achievement → German self-concept → German worry: β = -.184*; German achievement → math self-concept → math worry: β = .148*.
* $p < .05$.

Figure 1b

Standardized results of the mediation model (Model 4) for emotionality

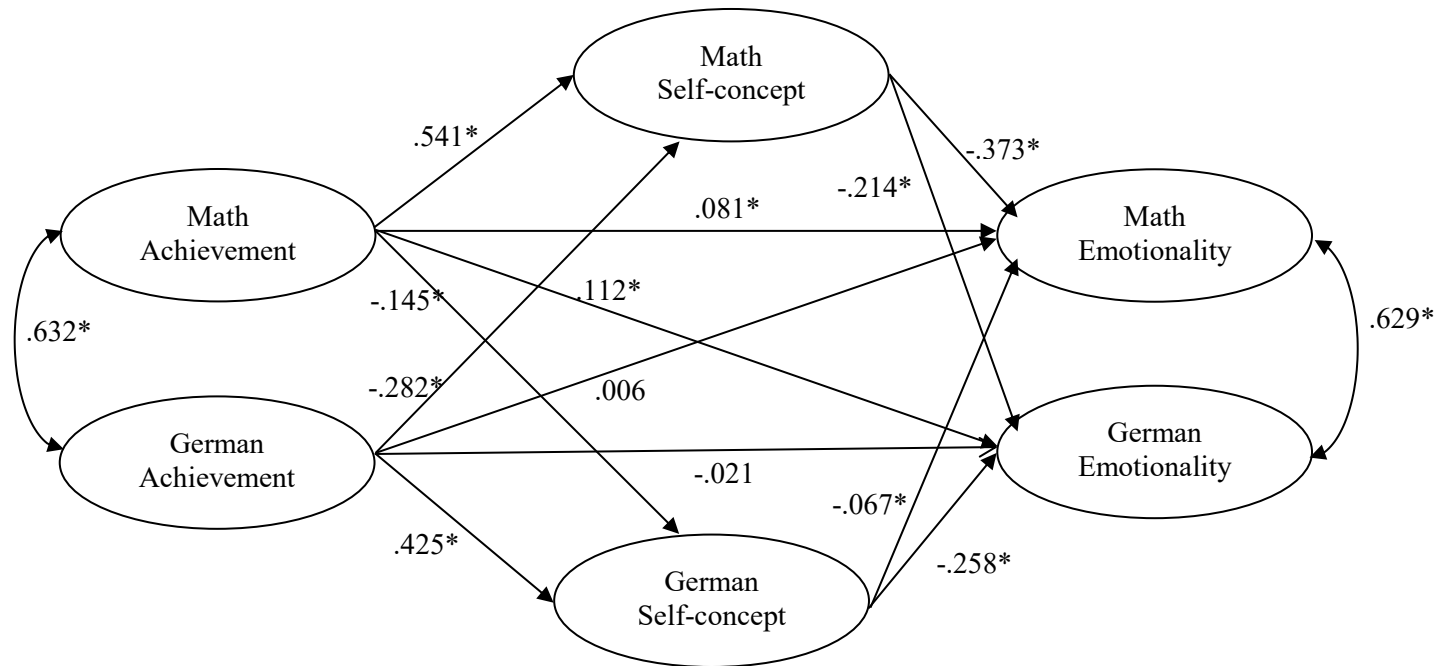*Note.* The (standardized) indirect effects are all statistically significant: Math achievement → math self-concept → math emotionality: β = -.202*;
Math achievement → German self-concept → German emotionality: β = .037*; German achievement → German self-concept → German emotionality: β = -.110*; German achievement → math self-concept → math emotionality: β = .105*.
* *p* < .05.

Table 1

Goodness-of-fit Indices

Model		χ^2	df	CFI	TLI	RMSEA	SRMR
1	Measurement model with six factors (i.e., worry, emotionality, achievement in math and German)	3393.759	564	.944	.938	.031	.035
2	Measurement model with eight factors: (i.e., worry, emotionality, self-concept, achievement in math and German)	5509.319	941	.935	.929	.031	.041
3	GI/E mediation model for worry as the outcome variable	3020.300	275	.922	.907	.044	.061
4	GI/E mediation model for emotionality as the outcome variable	2651.811	437	.955	.949	.031	.043
5	GI/E mediation model for worry and emotionality as outcome variables	5625.932	942	.933	.927	.031	.045
Invariance across Gender							
I1	Configural invariance	6558.767	1882	.933	.927	.031	.042
I2	Invariance of factor loadings	6728.871	1920	.932	.926	.031	.045
I3	Invariance of factor loadings and item intercepts	7098.480	1958	.927	.923	.032	.045
I4	Invariance of factor loadings, item intercepts, and item uniquenesses	7504.778	2002	.922	.919	.033	.047
I5	Invariance of factor loadings, item intercepts, item uniquenesses, and factor variances	7550.576	2010	.921	.919	.033	.051
I6	Invariance of factor loadings, item intercepts, item uniquenesses, factor variances, and factor covariances	7607.032	2038	.921	.920	.033	.051
I7	GI/E mediation model for worry; free across gender	3677.997	620	.912	.907	.044	.066
I8	GI/E mediation model for worry; invariance across gender	3711.308	634	.911	.909	.044	.067
I9	GI/E mediation model for emotionality; free across gender	3839.327	962	.941	.939	.034	.054
I10	GI/E mediation model for emotionality; invariance across gender	3891.796	976	.940	.939	.034	.057
I11	GI/E mediation model for worry and emotionality; free across gender	7667.112	2012	.920	.917	.033	.053
I12	GI/E mediation model for worry and emotionality; invariance across gender	7739.594	2039	.919	.918	.033	.055

Note. df = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean squared residual; GI/E model = generalized internal/external frame of reference model. All models integrated correlated uniquenesses between parallel-worded items and were conducted with the MLR estimator. All χ^2 values are significant ($p < .001$).

Table 2

Standardized Path Coefficients and Factor Correlations of the GI/E Model for Test Anxiety

Path Coefficients

	β	SE
Achievement German → Worry German	-.161*	0.028
Achievement German → Emotionality German	-.065*	0.025
Achievement German → Worry Math	.119*	0.025
Achievement German → Emotionality Math	.089*	0.025
Achievement Math → Worry Math	-.258*	0.022
Achievement Math → Emotionality Math	-.115*	0.022
Achievement Math → Worry German	-.028	0.025
Achievement Math → Emotionality German	.031	0.021

Correlations

	Worry Math	Emotionality Math	Worry German	Achievement German
Emotionality Math	.572*			
Worry German	.645*	.355*		
Emotionality German	.411*	.649*	.547*	
Achievement Math				.632*

Note. The fit of this model is equivalent to the confirmatory factor analysis model (Model 1 in Table 1) assuming separate factors for worry, emotionality, and achievement in math and German: χ^2 (564) = 3393.758 ($p < .001$); CFI = .944; TLI = .938; RMSEA = .031; SRMR = .035.
* $p < .05$.

Table 3

Standardized Factor Correlations of Model 2

	Worry Math	Worry German	Emotionality Math	Emotionality German	Self-concept Math	Self-concept German	Achievement Math
Worry German	.626*						
Emotionality Math	.580*	.350*					
Emotionality German	.400*	.550*	.643*				
Self-concept Math	-.534*	-.293*	-.350*	-.205*			
Self-concept German	-.162*	-.480*	-.105*	-.278*	.163*		
Achievement Math	-.190*	-.135*	-.058*	-.010	.360*	.120*	
Achievement German	-.051*	-.189*	.013	-.048*	.056*	.333*	.632*

Note. * $p < .05$.

Table 4

Standardized Path Coefficients and Factor Correlations of the Mediated GI/E Model for Test Anxiety (Model 5 in Table 1)

Standardized direct effects from achievement to anxiety	
Achievement German → Worry German	-.061 (0.028)*
Achievement German → Worry Math	.004 (0.024)
Achievement Math → Worry Math	.011 (0.021)
Achievement Math → Worry German	.048 (0.025)
Achievement German → Emotionality German	-.020 (0.029)
Achievement German → Emotionality Math	.005 (0.026)
Achievement Math → Emotionality Math	.081 (0.023)*
Achievement Math → Emotionality German	.111 (0.024)*
Standardized direct effects from achievement to academic self-concept	
Achievement German → Self-concept German	.426 (0.025)*
Achievement German → Self-concept Math	-.281 (0.023)*
Achievement Math → Self-concept Math	.541 (0.021)*
Achievement Math → Self-concept German	-.144 (0.024)*
Standardized direct effects from academic self-concept to anxiety	
Self-concept German → Worry German	-.439 (0.024)*
Self-concept German → Worry Math	-.095 (0.022)*
Self-concept Math → Worry Math	-.531 (0.019)*
Self-concept Math → Worry German	-.260 (0.021)*
Self-concept German → Emotionality German	-.258 (0.022)*
Self-concept German → Emotionality Math	-.066 (0.020)*
Self-concept Math → Emotionality Math	-.374 (0.019)*
Self-concept Math → Emotionality German	-.215 (0.021)*
Standardized indirect effects	
Achievement Math → Self-concept Math → Worry Math	-.287 (0.016)*
Achievement German → Self-concept German → Worry German	-.187 (0.016)*
Achievement Math → Self-concept German → Worry German	.063 (0.011)*
Achievement German → Self-concept Math → Worry Math	.149 (0.014)*
Achievement Math → Self-concept Math → Emotionality Math	-.202 (0.013)*
Achievement German → Self-concept German → Emotionality German	-.110 (0.012)*
Achievement Math → Self-concept German → Emotionality German	.037 (0.007)*
Achievement German → Self-concept Math → Emotionality Math	.105 (0.011)*

Note. * $p < .05$.

Online Supplements for

"Social and Dimensional Comparisons in Math and Verbal Test Anxiety: Within- and Cross-domain Relations with Achievement and the Mediating Role of Academic Self-concept"

Table S1

Standardized Factor Correlations of Model 1

	Worry Math	Worry German	Emotionality Math	Emotionality German	Achievement Math
Worry German	.633*				
Emotionality Math	.575*	.346*			
Emotionality German	.399*	.545*	.642*		
Achievement Math	-.183*	-.129*	-.058*	-.010	
Achievement German	-.044*	-.178*	.017	-.046*	.632*

Note. * $p < .05$.

Table S2

Results (Regression or Correlation Coefficients and Standard Errors in Parentheses) from the Mediation Models (Models 3 and 4 in Table 1)

	Worry (Model 3)	Emotionality (Model 4)
Standardized direct effects from achievement to anxiety		
Achievement German → Anxiety German	-.060 (0.028)*	-.021 (0.028)
Achievement German → Anxiety Math	.005 (0.021)	.006 (0.026)
Achievement Math → Anxiety Math	.009 (0.021)	.081 (0.023)*
Achievement Math → Anxiety German	.047 (0.025)	.112 (0.024)*
Standardized direct effects from achievement to academic self-concept		
Achievement German → Self-concept German	.425 (0.025)*	.425 (0.025)*
Achievement German → Self-concept Math	-.281 (0.023)*	-.282 (0.023)*
Achievement Math → Self-concept Math	.541 (0.021)*	.541 (0.021)*
Achievement Math → Self-concept German	-.144 (0.024)*	-.145 (0.024)*
Standardized direct effects from academic self-concept to anxiety		
Self-concept German → Anxiety German	-.433 (0.024)*	-.258 (0.022)*
Self-concept German → Anxiety Math	-.095 (0.022)*	-.067 (0.020)*
Self-concept Math → Anxiety Math	-.524 (0.019)*	-.373 (0.019)*
Self-concept Math → Anxiety German	-.257 (0.021)*	-.214 (0.021)*
Standardized indirect effects		
Achievement Math → Self-concept Math → Anxiety Math	-.284 (0.016)*	-.202 (0.013)*
Achievement German → Self-concept German → Anxiety German	-.184 (0.016)*	-.110 (0.012)*
Achievement Math → Self-concept German → Anxiety German	.062 (0.011)*	.037 (0.007)*
Achievement German → Self-concept Math → Anxiety Math	.148 (0.014)*	.105 (0.011)*
Correlations		
Anxiety Math ↔ Anxiety German	.611 (0.017)*	.629 (0.020)*
Achievement Math ↔ Achievement German	.633 (0.012)*	.632 (0.012)*

Note. * $p < .05$.