



# How Classmates' Gender Stereotypes Affect Students' Math Self-Concepts: A Multilevel Analysis

Fabian Wolff\*

*Institute of Psychology, University of Koblenz-Landau, Koblenz, Germany*

The present research is the first to examine how students' individual and their classmates' math-related gender stereotypes, endorsing that math would be a typically male domain, relate to students' math self-concepts. To this end, data of  $N = 1,424$  secondary school students from Germany were analyzed using multilevel analyses. As expected, strong individual beliefs in the math-related gender stereotype were related to lower math self-concepts for girls, but to higher math self-concepts for boys. Moreover, classmates' shared beliefs in this stereotype showed a negative relation to girls' self-concepts, whereas no significant relation between classmates' shared beliefs and boys' self-concepts was found. These relations also persisted after controlling for students' math grades and age. In sum, the results demonstrated that gender stereotypes shared by students' classmates can have a substantial impact on students' math self-concepts, beyond their individual gender stereotypes. This finding emphasizes the significance of classmates as important socializing peers in the process of students' self-concept formation.

**Keywords:** academic self-concept, beliefs, classmates, contextual effects, gender stereotypes, math self-concept, peers

## OPEN ACCESS

### Edited by:

Yves Karlen,  
University of Applied Sciences  
and Arts Northwestern Switzerland,  
Switzerland

### Reviewed by:

Michael D. Toland,  
University of Toledo, United States  
Malte Jansen,  
Institute for Educational Quality  
Improvement (IQB), Germany

### \*Correspondence:

Fabian Wolff  
wolff@uni-koblenz.de

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 26 August 2020

**Accepted:** 08 April 2021

**Published:** 12 May 2021

### Citation:

Wolff F (2021) How Classmates' Gender Stereotypes Affect Students' Math Self-Concepts: A Multilevel Analysis. *Front. Psychol.* 12:599199. doi: 10.3389/fpsyg.2021.599199

## INTRODUCTION

Although gender differences in math performance have largely vanished today, female students still tend to consider themselves as being less competent in math than their male counterparts (e.g., Hyde et al., 2008; Niepel et al., 2020). These gender differences in students' math self-concepts remain a cause for concern—not only because subject-specific self-concepts strongly relate to students' attitudes and affects toward school subjects (e.g., Goetz et al., 2008; Schurtz et al., 2014), but also because students include their self-concepts in their career choices (e.g., Parker et al., 2012, 2014). For that reason, the lower math self-concepts of females might hamper current political efforts to reduce the underrepresentation of women in fields related to math, science, technology, and engineering [Niepel et al., 2020; Organisation for Economic Co-operation and Development (OECD), 2020].

In the last decades, gender differences in students' math self-concepts have been examined in plenty of studies (e.g., Eccles-Parsons et al., 1983; Skaalvik and Skaalvik, 2004; Preckel et al., 2008). Among these, some studies have addressed the question of how gender differences in students' math self-concepts depend on gender stereotypes. Interestingly enough, these studies found that the gender differences in students' math self-concepts are larger, the more students believe that

math would be a subject for boys rather than for girls (e.g., Nosek et al., 2002; Steffens and Jelenec, 2011; Passolunghi et al., 2014). However, all of these studies only considered the relations between students' math self-concepts and their math-related gender stereotypes at the individual level, while neglecting the potential effects of math-related gender stereotypes shared by students' classmates. Given prior findings on the impact of significant adults' (teachers' and parents') math-related gender stereotypes on students' attitudes toward math (e.g., Gunderson et al., 2012), as well as the role of classmates as important socializing peers (e.g., Berndt and Murphy, 2002), it would seem reasonable to assume that classmates' math-related gender stereotypes affect students' math self-concepts beyond their individual stereotypes. Moreover, Muntoni et al. (2021) provided empirical evidence that students' self-concepts are affected by gender stereotypes shared in the classroom. In this study, the authors found that boys' reading self-concepts were lower, the more their classmates believed in the stereotype that reading would be a typically female domain. However, although most research on the relations between subject-specific gender stereotypes and students' self-concepts has focused on the math domain, no empirical study has so far examined whether Muntoni et al.'s (2021) findings would also apply to the math domain (i.e., whether girls show lower math self-concepts, the more their classmates believe in the stereotype favoring boys in math).

The present research addresses this research gap. To this end, I will analyze data collected in a large sample of secondary school students from Germany who reported both their math self-concepts and their math-related gender stereotypes. For the first time, I will examine whether and how girls' and boys' math self-concepts are affected by math-related gender stereotypes held by themselves and shared by their classmates. Thus, this study will make an important contribution to the literature by significantly enhancing our knowledge of the emergence of gender differences in students' math self-concepts. Beyond that, it is the first study to test whether Muntoni et al.'s (2021) findings—which have received considerable attention not only in the scientific community [e.g., Gesellschaft für Empirische Bildungsforschung (GEBF), 2019] but also in the media (e.g., Deutsches Schulportal, 2020; Society for Research in Child Development, 2020)—will generalize to another domain. This research will therefore provide us with valuable information to assess whether Muntoni et al. (2021) have discovered a psychological mechanism that operates across domains, or whether the relations shown in their study may be more specific to the domain of reading.

## Impact of Students' Own Gender Stereotypes on Their Math Self-Concepts

Stereotypes can be defined as shared beliefs about traits that are characteristic of members of a certain social category (e.g., Greenwald and Banaji, 1995). Although they may facilitate human behavior and decision-making in complex environments, they often come along with negative effects for the stereotyped group (e.g., Steele and Aronson, 1995; Schmader et al., 2004; Ihme

and Möller, 2015). Concerning math, prior studies on gender stereotypes have repeatedly found support for the prevalence of the stereotype endorsing that math would be a typically male domain (e.g., Nosek et al., 2002; Steffens et al., 2010; Cvencek et al., 2011; Steffens and Jelenec, 2011; but see Plante et al., 2009). This finding is somewhat paradoxical in that girls do not (any longer) perform significantly worse than boys in math [e.g., Hyde et al., 2008; Else-Quest et al., 2010; Lindberg et al., 2010; Niepel et al., 2020; Organisation for Economic Co-operation and Development (OECD), 2020] and even receive slightly better math grades (Voyer and Voyer, 2014). Yet, the prevailing view that males would be more competent in math may be a key reason as to why girls on average report lower math self-concepts than boys (e.g., Skaalvik and Skaalvik, 2004; Niepel et al., 2020) and thus develop less positive attitudes and affects toward math (e.g., Else-Quest et al., 2010; Goetz et al., 2013; Schurtz et al., 2014; Arens et al., 2017).

A prominent theory predicting these relations is Eccles' expectancy-value theory—a comprehensive model for the explanation of achievement-related choices and behaviors (EVT; Eccles-Parsons et al., 1983; Wigfield and Eccles, 2000; Eccles, 2009). EVT was originally developed to help explain gender differences in the likelihood of studying math and science. Among others, it assumes that stereotypes shared within one culture, such as beliefs about gender roles, affect students' self-concepts, and consequently their subject-specific task values and other achievement-related outcomes. However, EVT also assumes that students' perceptions of these stereotypes significantly account for the stereotypes' impact on students' self-concepts. In line with this assumption, several studies have shown that the influence of the math-related gender stereotype favoring boys in math on students' math self-concepts depends on students' beliefs in this stereotype (e.g., Nosek et al., 2002; Kurtz-Costes et al., 2008; Evans et al., 2011; Steffens and Jelenec, 2011; Passolunghi et al., 2014). More specifically, these studies revealed that girls tend to show lower math self-concepts, whereas boys typically have higher math self-concepts, the more they believe in this stereotype. Furthermore, some studies have shown that the effect of math-related gender stereotypes can even be reversed in favor of girls if students—against cultural conventions—are convinced that girls would be more competent in math than boys (e.g., Passolunghi et al., 2014).

To sum up, the stereotype endorsing that math would be a typically male domain does not affect each student's math self-concept in the same way. Rather, girls' and boys' beliefs regarding the validity of this stereotype are central determinants for its influence on their math self-concepts. Based on these findings, the present study further examines how students' beliefs in the gender stereotype favoring boys in math affect their math self-concepts by considering, for the first time, the potential influence of shared beliefs about this stereotype in students' classrooms.

## Impact of Classmates' Gender Stereotypes on Students' Self-Concepts

Besides the impact of students' individual gender stereotypes, previous research has found that students' self-concepts are also

affected by gender stereotypes held by significant others. These studies have dealt with the influence of parents' and teachers' beliefs (e.g., Jacobs and Eccles, 1992; Bouchey, 2004; Gunderson et al., 2012; Rowley et al., 2013; Muntoni and Retelsdorf, 2019). For example, Jacobs and Eccles (1992), analyzing longitudinal data, found that the math-related gender stereotypes of mothers influenced their perceptions of their children's math abilities in such a way that mothers with a strong belief in the stereotype favoring boys in math enhanced their perceived abilities of their sons but devalued their perceived abilities of their daughters. In turn, these ability perceptions affected the children's own math self-concepts, beyond students' actual gender and their teachers' beliefs about their abilities. These findings are also in accord with EVT, which emphasizes the role of socializers' beliefs and predicts that these beliefs, such as those about gender stereotypes, affect students' self-concepts even beyond students' own beliefs about gender stereotypes.

Unlike the relatively high number of studies examining the impact of parents' and teachers' gender stereotypes on students' math self-concepts, the impact of classmates' gender stereotypes on students' math self-concepts has not been examined so far. Nevertheless, such an influence seems plausible, given that peers also represent important socializers for students, especially during adolescence (e.g., Berndt, 1979; Hartup, 1996; Espelage et al., 2003; Brown, 2011). In particular, previous research has shown that peers can be of high importance for imparting behavior connoting appropriateness based on gender (e.g., Lamb et al., 1980; Zucker and Bradley, 1995; Bussey and Bandura, 1999; Ruble et al., 2007). Moreover, several studies have found that students' individual beliefs and values can be significantly affected by those of their classmates (e.g., Ryan, 2001; Bouchey, 2004; Frenzel et al., 2010; Studsrød and Bru, 2011). For example, with regard to math, Bouchey (2004) showed in a cross-sectional study that classmates' ratings of students' math ability were positively related to students' math self-concepts, after controlling for students' math achievement. Besides, Frenzel et al. (2010) found in a longitudinal study that students' math interest was higher, the more value their classmates expressed toward math. These authors also noted that classmates can be assumed to play an important role in the formation of students' achievement-related beliefs and values, especially in Germany, as students in the German school system stay within the same group of students across most subjects within a school year.

More specific evidence for the assumption that students' math self-concepts might be influenced by their classmates' math-related gender stereotypes stems from two longitudinal studies, in which data collected in Germany were re-analyzed. In the one study, Salikutluk and Heyne (2017) investigated how classmates' gender norms impact on students' math achievement. They found that girls performed worse in classes where traditional gender norms were strongly present, whereas girls' and boys' math achievement did not differ in classes where traditional gender norms were low or absent. In the other study, Muntoni et al. (2021) examined the relations between reading-related gender stereotypes and reading self-concepts. As already stated in the Introduction, they found

that boys developed lower reading self-concepts, the more their classmates believed in the stereotype favoring girls in reading, whereas classmates' beliefs in this stereotype did not affect girls' reading self-concepts. Similar findings also emerged for students' reading self-efficacy, reading motivation, and reading achievement.

Taken together, the results of both studies suggest that high subject-specific gender stereotypes shared within the classroom might negatively impact on the stereotyped group. However, it should be noted that the gender norms examined by Salikutluk and Heyne (2017) did not refer to school subjects, but particularly addressed the question of whether men should use violence to defend their families. Thus, it is unclear whether Salikutluk and Heyne's findings would be replicated if gender norms (or gender stereotypes) referred to math. In particular, it is worth noting that students' individual gender norms in Salikutluk and Heyne's study were negatively related to math achievement not only for girls but also for boys—presumably because a strong belief that men should use violence might go along with more antisocial behavior at school and thus with lower math achievement for boys. Moreover, it is questionable to what extent the effects of gender norms on students' math achievement correspond to those on their math self-concept, given that the gender gap in math self-concept is significantly stronger than in math achievement (e.g., Niepel et al., 2020).

In contrast to Salikutluk and Heyne (2017), Muntoni et al. (2021) investigated the impact of classmates' subject-specific gender stereotypes on students' self-concepts more specifically. Notwithstanding this, it cannot be taken for granted that their results would generalize to the math domain. This is especially true because the stereotyped gender groups in the domains of reading and math differ from each other. For example, an important difference between stereotypes relating to boys' reading competence versus girls' math competence is that boys, in fact, perform worse in reading, whereas the gender difference in students' math performance is at least very small [e.g., Organisation for Economic Co-operation and Development (OECD), 2020]. Thus, one could speculate that classmates' subject-specific gender stereotypes affect boys' reading self-concepts more strongly than girls' math self-concepts because the stereotype favoring girls in reading is objectively more tenable. Aside from that, it is worth mentioning that the contextual effect on boys' reading self-concepts found by Muntoni et al. (2021) strongly resulted from an absence of the effect of boys' individual reading-related gender stereotypes on their reading self-concepts. However, concerning the math domain, such a finding would contradict the various findings on the impact of students' individual beliefs in the stereotype favoring boys in math on their math self-concepts (e.g., Passolunghi et al., 2014). This also makes it necessary to investigate whether Muntoni et al.'s (2021) findings can be generalized to the domain of math.

To summarize, it seems likely that classmates' math-related gender stereotypes affect students' math self-concepts beyond their individual gender stereotypes. Nevertheless, this is an empirical question, which has not yet been addressed in any study. The present study aims to close this research gap.

## THE PRESENT RESEARCH

This research is the first to investigate the relations between students' and their classmates' math-related gender stereotypes and students' math self-concepts. Its main purpose is to examine the influence of classmates' gender stereotypes on students' self-concepts. Thus, for the first time, it aims to test whether Muntoni et al.'s (2021) much-noticed findings concerning the relations between classmates' reading-related gender stereotypes and students' reading self-concepts would generalize to the domain of math. Moreover, this study seeks to dive further into the question as to what extent students' individual beliefs concerning math-related gender stereotypes affect their math self-concepts by examining these relations, separately for girls and boys, after controlling for the relations between gender stereotypes and self-concepts at the class level.

Based on the theoretical reflections and empirical findings presented in the previous sections, I assumed to find a negative relation between girls' individual beliefs in the gender stereotype favoring boys in math and their math self-concepts (*Hypothesis 1a*). On the contrary, I assumed to find a positive relation between boys' individual beliefs in this stereotype and their math self-concepts (*Hypothesis 1b*). Furthermore, I expected that girls' math self-concepts showed a negative relation to their classmates' beliefs in the gender stereotype favoring boys in math (*Hypothesis 2*), given that previous research suggests that high subject-specific gender stereotypes shared within the classroom might negatively impact on the stereotyped group (Salikutluk and Heyne, 2017; Muntoni et al., 2021). That is, I assumed that classmates' math-related gender stereotypes would show a negative relation to girls' math self-concepts, after controlling for students' individual differences in their math-related gender stereotypes. I left it as an open research question whether classmates' beliefs in the gender stereotype favoring boys in math would also relate to boys' math self-concepts. For example, whereas some studies have found that stereotypes questioning the ability of an outgroup can lead to performance boosts in the ingroup ("stereotype lift"; e.g., Walton and Cohen, 2003), other studies have found that they can decrease the performance in the ingroup ("choking under pressure"; e.g., Cheryan and Bodenhausen, 2000). Nevertheless, given that Muntoni et al. (2021) found no significant effect of classmates' beliefs in the stereotype favoring girls in reading on girls' reading self-concepts, I considered it as likely that boys' math self-concepts would not significantly relate to their classmates' beliefs in the stereotype favoring boys in math.

I tested my hypotheses in a large sample of secondary school students from Germany. To account for potential alternative explanations, I also tested the relations between students' math-related gender stereotypes and math self-concepts after controlling for students' math grades and age. I controlled for students' math grades, as students form their subject-specific self-concepts especially based on their grades. In particular, previous research has shown that students' math self-concepts show significantly stronger relations to their math grades than to their results from standardized math tests (e.g., Wolff et al., 2019b; Möller et al., 2020). I took into account students' age since

students' math self-concepts usually decrease with increasing age (e.g., Wolff et al., 2020b; Orth et al., 2021). I expected to find support for all stated hypotheses even after controlling for the covariates.

## MATERIALS AND METHODS

### Sample

The sample consisted of  $N = 1,424$  students (age:  $M = 15.1$ ,  $SD = 2.01$ ; 56.2% female, 42.6% male, 1.2% not specified) between Grade 7 and Grade 13. These students attended 90 classes of 11 secondary schools in the German federal state of Schleswig-Holstein. The number of classes per school ranged from 2 to 14 with an average of about 8 classes per school. The sample stemmed from a larger project with the major aim to examine comparison effects in the process of students' self-concept formation (Wolff et al., 2021). The schools were recruited by direct requests to individual school offices. Participating classes were selected by school administrators. In many cases, there was direct contact with specific teachers at the schools who offered to participate in the study, particularly with classes they taught themselves. The participation was voluntary and the informed consent of the parents was required if students were underage. A lack of informed parental consent was the main reason why students did not participate in the study. It mostly resulted from the fact that the students had not handed over the parent letter in advance of the study.

### Procedure

The data collection took place in the spring of 2018. The students answered paper-and-pencil questionnaires during regular school lessons. The constructs were measured in the following order: math self-concept, math-related gender stereotype, math grade, and demographics (age and gender). Approval for the whole procedure was obtained by the local ministry of education.

## Measures

### Math Self-Concepts

Students' math self-concepts were measured using six items: (1) "With a number of things in math, I immediately know: I will never understand this," (2) "Although I try my best, math is hard for me," (3) "I simply have no natural aptitude in math," (4) "I would much rather do math if the subject were not so difficult," (5) "Math does not come naturally to me," (6) "It comes easily to me to understand tasks and solve problems in math." These items had already been used successfully in contemporary self-concept research (e.g., Wolff et al., 2019a). The students responded to all items on a 6-point Likert scale ranging from 1 = *strongly disagree* to 6 = *strongly agree*. The negatively phrased items were reverse coded in the way that higher scores indicated higher self-concepts. The reliability of the scale according to Geldhof et al. (2014) was high at the within-level ( $\alpha = 0.92$ ) and the between-level ( $\alpha = 0.97$ ). Moreover, invariance tests provided evidence for the invariance of the scale across gender groups (see **Table 1**).



**TABLE 1** | Measurement invariance of math self-concept between gender groups.

Model	Equality constraints	$\chi^2$	df	p	CFI	TLI	RMSEA
1	Baseline	179.54	18	<0.001	0.996	0.993	0.113
2	Thresholds	183.46	36	<0.001	0.996	0.997	0.076
3	Thresholds and loadings	161.19	41	<0.001	0.997	0.998	0.065

The table shows the pooled results over  $m = 50$  imputed data sets. They stem from single-level analyses in MPlus in which the complex modeling procedure (type = complex) was used to correct the estimated standard errors for the nested data structure of students in classes. The self-concept items were treated as categorical indicators and measurement invariance was tested according to Svetina et al. (2020). In Model 1, the indicators were allocated to one factor, but the thresholds and loadings were estimated freely across the groups. In Model 2, the thresholds were constrained to be equal across the groups. In Model 3, the thresholds and loadings were constrained to be equal across the groups. Measurement invariance is supported since the model fit improves with increasing equality constraints. CFI, comparative fit index; TLI, Tucker-Lewis index; RMSEA, root mean square error of approximation.  $N = 1,424$  ( $n = 800$  girls,  $n = 607$  boys,  $n = 17$  not specified).

## Gender Stereotypes

Students' beliefs in the stereotype endorsing that math would be a typically male subject were measured using three items: (1) "Boys are simply more gifted at math," (2) "If there was a typical boy subject, math would be one," (3) "Math is a subject that is usually more fun for boys." These items had been developed by Pohlmann (2005). The students responded to all items on a 6-point Likert scale ranging from 1 = *strongly disagree* to 6 = *strongly agree*. The reliability of the scale (Geldhof et al., 2014) was high at the within-level ( $\alpha = 0.90$ ) and the between-level ( $\alpha = 0.99$ ).

## Grades

The students reported their math grades from their latest report cards. Previous research has shown that self-reports of grades in the German school system are sufficiently reliable (e.g., Dickhäuser and Plenter, 2005). All grades were coded according to the German 15-point grading system, which ranges from 0 = *insufficient* to 15 = *perfect*. This grading system is mainly used at the upper secondary level in Germany, which was also the case in the present study. When teachers assigned grades according to the 6-point grading system, ranging from 1 = *excellent* to 6 = *insufficient* (as the case in most of the lower classes in the sample of the present study), I converted these grades into the 15-point system, using the official transformation key, according to which grade 1 equals 14 points, grade 2 equals 11 points, grade 3 equals 8 points, grade 4 equals 5 points, grade 5 equals 2 points, and grade 6 equals 0 points. I preferred this transformation as it did not lead to any loss of information. Nevertheless, I note that the findings of this study fully replicated if transforming the grades from the 15-point system to the 6-point system instead (i.e., 13–15 points to grade 1, 10–12 points to grade 2, etc.).

## Statistical Analyses

The statistical analyses were similar to Muntoni et al. (2021). I conducted multilevel analyses in MPlus 8.4 (Muthén and Muthén, 2017), with students (within-level) being clustered in classes (between-level). More specifically, I specified doubly latent multiple group two-level structural equation models to examine whether students' math self-concepts were affected by their individual math-related gender stereotypes and the average math-related gender stereotypes within their classes. In these models, students' gender stereotypes were estimated

as latent variables at the within- and between-level, which allowed controlling for measurement and sampling errors (Marsh et al., 2009). Since group membership (i.e., gender) was included as a within-level grouping variable (because each class consisted of girls and boys), the between-level data were not independent between the within-level groups. To account for this dependency, I applied multilevel mixture models, in which a latent class variable indicating the group membership was specified (Asparouhov and Muthén, 2012). Thus, a latent class could be set up at the within-level and the group membership variable was specified as a perfect indicator of the latent class variable (Muntoni et al., 2021). To handle missing values (on average, 2.2% of the data were missing per variable), I applied multiple imputation, including  $m = 50$  imputed data sets (Graham et al., 2007).

Overall, I calculated two models. In Model 1, I only took students' gender stereotypes (independent variable) and self-concepts (dependent variable) into consideration. In Model 2, I additionally controlled for students' math grades and age (control variables). In both models, I also included 10 dummy variables, indicating students' affiliation to 1 of 11 schools, which allowed me to take account of the three-level data structure. For testing Hypotheses 1a and 1b, I examined whether the effects of students' gender stereotypes on their math self-concepts at the within-level were significantly negative for girls and significantly positive for boys. For testing Hypothesis 2, I examined whether contextual effects of class-average gender stereotypes on students' math self-concepts occurred. To this end, I calculated additional parameters indicating whether the effects of gender stereotypes on self-concepts at the between- and within-level were significantly different from each other. This was necessary because the variables which appeared at both levels were implicitly group-mean centered. Accordingly, the effects of class-average gender stereotypes on math self-concepts at the between-level were no direct estimates of the contextual effects (Marsh et al., 2009). As suggested in Hypothesis 2, the contextual effects should be significantly negative for girls.

I calculated Tymms' (2004)  $\Delta$  to facilitate the interpretation of the effect sizes. For this purpose, I used the formula  $\Delta = (2 \times B \times SD_{predictor})/\sigma$ , where  $B$  is the unstandardized regression coefficient,  $SD_{predictor}$  is the standard deviation of the predictor variable (for the contextual effect: at the between-level), and  $\sigma$  is the total standard deviation of the outcome variable (Marsh et al., 2009). Tymms'  $\Delta$  can be interpreted similar to

Cohen's (1988) *d*. Thus,  $|\Delta| = 0.2$  represents a small effect,  $|\Delta| = 0.5$  a moderate effect, and  $|\Delta| = 0.8$  a large effect.

## RESULTS

### Preliminary Analyses

Table 2 presents the descriptive statistics. In line with previous research, boys showed higher math self-concepts compared to girls ( $\Delta M = 0.43$ ,  $\beta = 0.16$ ,  $p < 0.001$ ), whereas boys' and girls' math grades did not significantly differ from each other ( $\Delta M = -0.07$ ,  $\beta = -0.01$ ,  $p = 0.71$ ). Boys also held a slightly stronger belief in the stereotype favoring boys in math ( $\Delta M = 0.19$ ,  $\beta = 0.08$ ,  $p = 0.02$ ). Moreover, boys' math-related gender stereotypes were positively related to their math self-concepts ( $r = 0.18$ ,  $p < 0.001$ ) and math grades ( $r = 0.12$ ,  $p < 0.01$ ), whereas girls' math-related gender stereotypes were negatively related to their math self-concepts ( $r = -0.26$ ,  $p < 0.001$ ) and math grades ( $r = -0.21$ ,  $p < 0.001$ ). For both genders, older students showed lower math grades (both  $r = -0.14$ ,  $p < 0.01$ ). In addition, older girls showed slightly stronger beliefs in the stereotype favoring boys in math compared to younger girls ( $r = 0.13$ ,  $p < 0.01$ ) and older boys showed slightly lower math self-concepts compared to younger boys ( $r = -0.15$ ,  $p < 0.01$ ). The correlations between students' math self-concepts and math grades were strongly positive for girls ( $r = 0.61$ ,  $p < 0.001$ ) and boys ( $r = 0.58$ ,  $p < 0.001$ ). The intraclass correlation coefficients (ICCs), indicating the proportion of total variance that was accounted for by between-class variance, were 0.05 or higher for all four variables. In particular, the ICC of 0.11 found for students' beliefs in the stereotype favoring boys in math is noteworthy, as it indicated that students of the same class showed a substantial agreement in their belief in this stereotype. Furthermore, this value suggested that it was appropriate to examine the relations between students' math-related gender stereotypes and their math self-concepts using multilevel analyses.

### Hypotheses Testing

Table 3 presents the results of the multilevel analyses. Model 1 tested the effects of students' math-related gender stereotypes on their math self-concepts without considering control variables. As predicted in Hypothesis 1, girls' individual math-related gender stereotypes showed a negative effect on their math self-concepts ( $B = -0.26$ ,  $p < 0.001$ ,  $\Delta = -0.43$ ), whereas this effect was positive for boys ( $B = 0.21$ ,  $p < 0.001$ ,  $\Delta = 0.35$ ). Moreover, as predicted in Hypothesis 2, there was a negative contextual effect on girls' math self-concepts ( $B = -0.38$ ,  $p \leq 0.05$ ,  $\Delta = -0.23$ ), implying that classmates' math-related gender stereotypes were negatively related to girls' math self-concepts, after controlling for individual differences in gender stereotypes. The contextual effect on boys' math self-concepts was non-significant ( $B = 0.07$ ,  $p = 0.78$ ). Model 2 demonstrated that these relations also persisted after controlling for students' math grades and age. Whereas the negative effect of girls' individual math-related gender stereotypes on their math self-concepts ( $B = -0.15$ ,  $p < 0.001$ ,  $\Delta = -0.24$ )

TABLE 2 | Means, standard deviations, bivariate correlations, and intraclass correlations [with 95% confidence intervals] of the manifest variables.

Variables	<i>M</i> <sub>girls</sub>	<i>M</i> <sub>boys</sub>	<i>SD</i> <sub>girls</sub>	<i>SD</i> <sub>boys</sub>	<i>r</i> <sub>math self-concept</sub>	<i>r</i> <sub>gender stereotype</sub>	<i>r</i> <sub>math grade</sub>	<i>r</i> <sub>age</sub>	ICC
Math self-concept	3.85 [3.72, 3.97]	4.27 [4.14, 4.40]	1.38 [1.33, 1.43]	1.28 [1.21, 1.34]	–	0.18 [0.10, 0.27]	0.58 [0.52, 0.64]	–0.15 [–0.24, –0.06]	0.05 [0.02, 0.08]
Gender stereotype	2.73 [2.58, 2.87]	2.92 [2.78, 3.06]	1.25 [1.20, 1.31]	1.27 [1.21, 1.33]	–0.26 [–0.34, –0.19]	–	0.12 [0.04, 0.21]	0.08 [–0.02, 0.18]	0.11 [0.07, 0.14]
Math grade	8.43 [8.12, 8.74]	8.36 [8.05, 8.67]	3.03 [2.87, 3.18]	3.10 [2.87, 3.30]	0.61 [0.56, 0.66]	–0.21 [–0.29, –0.13]	–	–0.14 [–0.23, –0.05]	0.09 [0.05, 0.12]
Age	15.21 [14.76, 15.66]	14.95 [14.52, 15.37]	2.05 [1.80, 2.26]	1.94 [1.70, 2.15]	–0.09 [–0.18, 0.01]	0.13 [0.04, 0.23]	–0.14 [–0.23, –0.05]	–	0.90 [0.87, 0.92]

The table shows the pooled results over  $m = 50$  imputed data sets. The means, standard deviations, and bivariate correlations stem from a single-level analysis in MPlus in which the complex modeling procedure (type = complex) was used to correct the estimated standard errors for the nested data structure of students in classes. The correlations below the diagonal refer to the subsample of girls. The correlations above the diagonal refer to the subsample of boys. The intraclass correlation coefficients (ICCs) indicate the proportion of total variance that is accounted for by between-class variance. They stem from additional multilevel analyses. Math self-concept and gender stereotype had a possible range of 1 = strongly disagree to 6 = strongly agree. Math grade had a possible range of 0 = insufficient to 15 = perfect. Bold values are significantly different from zero ( $p < 0.05$ ).  $N = 1,424$  ( $n = 800$  girls,  $n = 607$  boys,  $n = 17$  not specified).

**TABLE 3** | Unstandardized regression coefficients, standard deviations, and Tymms'  $\Delta$  [with 95% confidence intervals] of the multilevel analyses predicting students' math self-concept.

Variables	Model 1						Model 2					
	Girls			Boys			Girls			Boys		
	B	SD	$\Delta$	B	SD	$\Delta$	B	SD	$\Delta$	B	SD	$\Delta$
<b>Within-level</b>												
Gender stereotype	<b>-0.26</b> [-0.35, -0.16]	<b>1.13</b> [1.06, 1.19]	<b>-0.43</b> [-0.59, -0.27]	<b>0.21</b> [0.11, 0.30]	<b>1.13</b> [1.06, 1.19]	<b>0.35</b> [0.19, 0.51]	<b>-0.15</b> [-0.22, -0.07]	<b>1.13</b> [1.06, 1.19]	<b>-0.24</b> [-0.37, -0.11]	<b>0.13</b> [0.06, 0.21]	<b>1.13</b> [1.06, 1.19]	<b>0.24</b> [0.10, 0.37]
Math grade	-	-	-	-	-	-	<b>0.28</b> [0.25, 0.31]	<b>2.93</b> [2.79, 3.06]	<b>1.19</b> [1.07, 1.31]	<b>0.24</b> [0.21, 0.27]	<b>2.93</b> [2.79, 3.06]	<b>1.12</b> [0.98, 1.26]
Age	-	-	-	-	-	-	0.09 [-0.01, 0.19]	<b>0.65</b> [0.56, 0.73]	0.09 [-0.01, 0.18]	-0.06 [-0.19, 0.08]	<b>0.65</b> [0.56, 0.73]	-0.06 [-0.19, 0.08]
<b>Between-level</b>												
Gender stereotype	<b>-0.63</b> [-0.99, -0.28]	<b>0.40</b> [0.33, 0.47]	<b>-0.38</b> [-0.59, -0.17]	0.28 [-0.22, 0.79]	<b>0.40</b> [0.33, 0.47]	0.17 [-0.14, 0.48]	<b>-0.58</b> [-0.95, -0.21]	<b>0.40</b> [0.33, 0.46]	<b>-0.34</b> [-0.56, -0.13]	-0.03 [-0.52, 0.46]	<b>0.40</b> [0.33, 0.46]	-0.02 [-0.33, 0.29]
Math grade	-	-	-	-	-	-	0.08 [-0.06, 0.22]	<b>0.90</b> [0.67, 1.08]	0.11 [-0.08, 0.29]	-0.04 [-0.27, 0.19]	<b>0.90</b> [0.67, 1.08]	-0.06 [-0.39, 0.27]
Age	-	-	-	-	-	-	-0.05 [-0.10, 0.01]	<b>1.91</b> [1.70, 2.10]	-0.13 [-0.29, 0.03]	-0.07 [-0.16, 0.03]	<b>1.91</b> [1.70, 2.10]	-0.20 [-0.48, 0.08]
<b>Contextual effect</b>												
Gender stereotype	<b>-0.38</b> [-0.76, -0.00]	<b>0.40</b> [0.33, 0.47]	<b>-0.23</b> [-0.46, -0.00]	0.07 [-0.44, 0.59]	<b>0.40</b> [0.33, 0.47]	0.05 [-0.27, 0.36]	<b>-0.44</b> [-0.82, -0.05]	<b>0.40</b> [0.33, 0.46]	<b>-0.26</b> [-0.48, -0.03]	-0.16 [-0.67, 0.34]	<b>0.40</b> [0.33, 0.46]	-0.10 [-0.42, 0.22]

The table shows the pooled results over  $m = 50$  imputed data sets. Tymms' (2004)  $\Delta$  was calculated using the formula  $\Delta = (2 \times B \times SD_{predictor})/\sigma$ , where  $B$  is the unstandardized regression coefficient,  $SD_{predictor}$  is the standard deviation of the predictor variable (for the contextual effect: at the between-level), and  $\sigma$  is the total standard deviation of math self-concept (Model 1:  $\sigma = 1.35$  for girls,  $\sigma = 1.32$  for boys; Model 2:  $\sigma = 1.37$  for girls,  $\sigma = 1.27$  for boys). Both models also included 10 dummy variables at the between-level, indicating students' affiliation to 1 of 11 schools (not depicted). Bold values are significantly different from zero ( $p < 0.05$ ).  $N = 1,424$  ( $n = 800$  girls,  $n = 607$  boys,  $n = 17$  not specified).

and the positive effect of boys' individual math-related gender stereotypes on their math self-concepts ( $B = 0.13$ ,  $p < 0.001$ ,  $\Delta = 0.24$ ) were slightly reduced, the negative contextual effect on girls' math self-concepts was even slightly stronger ( $B = -0.44$ ,  $p = 0.03$ ,  $\Delta = -0.26$ ). The contextual effect on boys' math self-concepts was still non-significant ( $B = -0.16$ ,  $p = 0.53$ ). Students' individual math grades showed strong positive effects on their math self-concepts (girls:  $B = 0.28$ ,  $p < 0.001$ ,  $\Delta = 1.19$ ; boys:  $B = 0.24$ ,  $p < 0.001$ ,  $\Delta = 1.12$ ). The effects of math grades at the between-level and of age at both levels were non-significant (all  $|B| \leq 0.09$ ,  $p \geq 0.07$ ).

## DISCUSSION

The present research substantially extends our knowledge of the relations between subject-specific gender stereotypes and self-concepts. For the first time, I investigated how students' individual and their classmates' math-related gender stereotypes affect girls' and boys' math self-concepts. In line with my hypotheses, and with findings from previous research (e.g., Nosek et al., 2002; Steffens and Jelenec, 2011; Passolunghi et al., 2014), I found that girls showed lower math self-concepts, the more they believed in the stereotype favoring boys in math, whereas boys showed higher math self-concepts, the more they believed in this stereotype. Furthermore, and most central for this research, I found a negative contextual effect of classmates' beliefs in the stereotype favoring boys in math on girls' math self-concepts. Thus, girls attending classes in which students strongly believed in the stereotype that math would be a typically male domain showed lower math self-concepts than girls attending classes in which the students did not believe in this stereotype, after controlling for individual gender stereotypes. The contextual effect of classmates' beliefs in the stereotype favoring boys in math on boys' math self-concepts was close to zero and indicated that classmates' shared beliefs in this stereotype do not significantly affect boys' math self-concepts.

### Theoretical Implications

The findings of this research have important theoretical implications as they further support the role of classmates as significant socializing peers. Specifically, it was shown that girls' math self-concepts were lower the more their classmates believed in the stereotype favoring boys in math. In contrast, boys' math self-concepts were not related to the gender stereotypes shared in students' classrooms. Remarkably, both the significant contextual effect on girls' math self-concepts and the non-significant contextual effect on boys' math self-concepts found in the present research were in accord with the results of Muntoni et al. (2021), who found a negative contextual effect of shared reading-related gender stereotypes, endorsing that reading would be a typically female domain, on boys' reading self-concepts, but not on girls' reading self-concepts. Taken together the results of both studies, it seems that gender stereotypes shared in the classroom negatively affect the self-concepts of students of the stereotyped gender group, whereas they do not affect the self-concepts of students of the non-stereotyped gender group.

Muntoni et al. (2021) had already speculated, when they compared the effects of gender stereotypes between the domains of math and reading, "that the underlying processes are rather similar" (p. 190). Nevertheless, an empirical test of this assumption seemed indicated, especially because math-related gender stereotypes are even less tenable than reading-related gender stereotypes, according to differences in girls' and boys' performance in standardized math and reading tests: Whereas girls, in fact, do not perform significantly worse in math than boys (which also corresponds to the similar math grades found for girls and boys in the present study), they still outperform boys in reading [e.g., Organisation for Economic Co-operation and Development (OECD), 2020]. It would, therefore, also have been plausible if girls' math self-concepts had not been related to shared math-related gender stereotypes in the present study, although Muntoni et al. (2021) found a negative effect of shared reading-related gender stereotypes on boys' reading self-concepts.

However, the findings of the present study suggest that the impact of math-related gender stereotypes on girls' math self-concepts may be even stronger than the impact of reading-related gender stereotypes on boys' reading self-concepts. This is due to the fact that the contextual effect on girls' math self-concepts occurred beyond the negative effect of girls' individual math-related gender stereotypes on their math self-concepts, whereas the contextual effect on boys' reading self-concepts in Muntoni et al.'s (2021) study, which was similarly strong as the contextual effect on girls' math self-concepts in the present study, occurred along with a non-significant effect of boys' individual reading-related gender stereotypes on their reading self-concepts. More precisely, Muntoni et al. (2021) examined two models. In the first model, they only controlled for demographics. In this model, they found positive effects of girls' individual reading-related gender stereotypes on their reading self-concepts and negative effects of boys' individual reading-related gender stereotypes on their reading self-concepts, but no contextual effects. The contextual effect on boys' reading self-concepts was only shown in the second model, in which the authors additionally controlled for prior reading self-concepts. Yet, in this model, the effects of gender stereotypes at the within-level were non-significant. The fact that the present research revealed simultaneous effects of students' individual and shared gender stereotypes on their self-concepts constitutes an important difference to Muntoni et al.'s (2021) results. Although the results of both studies suggest that shared gender stereotypes have specific effects on the self-concepts of the stereotyped group, it seems that there are also some differences in the mechanisms of how gender stereotypes affect students' self-concepts in the domains of math and reading.

### Practical Implications

The findings of this research are of high practical importance as they illustrate, more differentiated than in previous research, how girls' math self-concepts are impaired by math-related gender stereotypes. On the one hand, it was demonstrated that girls show lower math self-concepts, the more they believe in the stereotype favoring boys in math. This finding was



already known from prior studies (e.g., Passolunghi et al., 2014). However, in the present study, it was shown for the first time that these relations also hold in multilevel analyses. On the other hand, the present research revealed that math-related gender stereotypes shared in the classroom negatively relate to girls' math self-concepts, after controlling for girls' individual math-related gender stereotypes. To the best of my knowledge, this finding has not been shown in any prior empirical study yet. To conclude, the math-related gender stereotype endorsing that math would be a typically male domain is not only incompatible with empirical findings that showed no substantial differences in girls' and boys' math performance [e.g., Organisation for Economic Co-operation and Development (OECD), 2020]. It also seems to be disadvantageous for girls in two ways. Teachers are, therefore, advised to avoid the emergence of gender stereotypes and to remove existing stereotypes whenever possible. In particular, this applies if gender stereotypes manifest themselves in the classroom and thus affect students' self-concepts (see, e.g., Frawley, 2005, for interventional strategies).

It is worth noting that the effects of students' individual and shared math-related gender stereotypes found in the present study were only small—similarly to the effects of students' individual and shared reading-related gender stereotypes found by Muntoni et al. (2021). Nevertheless, it seems likely that these effects will have a substantial impact over time and could even become a self-fulfilling prophecy. Several studies have demonstrated that students' academic self-concepts are not only formed based on their prior achievements but also affect their subsequent achievements to a substantial degree (e.g., Marsh and Craven, 2006; Wolff et al., 2020b). If girls' math self-concepts suffer from math-related stereotypes, this can consequently impair their future math performance. In particular, girls might then be prone to believe that they indeed perform worse in math than boys, with the result that their math self-concepts and achievements could further worsen—and that they might decide against a career in math, science, technology, or engineering, even though they have the potential to succeed in these disciplines.

## Strengths, Limitations, and Directions for Future Research

The present research has some limitations that should be discussed in greater depth. A first limitation is its cross-sectional design. Accordingly, the findings of this study allow us to make conclusions about how students' math-related gender stereotypes are related to their math self-concepts measured at a certain point in time, but not to changes in these self-concepts across time. Future research should, therefore, aim to examine how individual and shared math-related gender stereotypes relate to changes in students' self-concepts. In particular, it would seem worthwhile to conduct such investigations during phases of school transitions, which have shown to come along with significant changes in students' self-concepts (e.g., Wigfield et al., 1991; Wolff et al., 2020c), given that these changes might also result from changes in gender stereotypes shared within the classroom. Moreover, it would be advisable for future research to measure students'

gender stereotypes (and other covariates) some time before the self-concepts as such a design would be more suitable to allow an approximation to causal conclusions than the simultaneous assessment of these constructs.

A second limitation of the present study can be seen in the fact that the classes were selected by the school administrators, rather than chosen randomly. Accordingly, the examined sample was not representative. However, this limitation seemed acceptable since the aim of this research was not to examine gender stereotypes and self-concepts in a representative student sample, but the relations between these constructs at the individual and class level. For this purpose, the analyzed data seemed quite suitable as they showed several favorable characteristics. For example, the internal consistencies of the self-concept scale and the scale assessing students' gender stereotypes were very high, the ICC of students' gender stereotypes was also high, and the number of missing values was very low. Beyond that, the data were collected in the year 2018, specifically for the present study. Given that gender stereotypes have changed across the decades (e.g., Eagly et al., 2020), this up-to-datedness of the analyzed data can be seen as a strength of this study.

A third limitation of the present research is its restriction to the investigation of the relations between students' subject-specific gender stereotypes and self-concepts in the domain of math. For this reason, it was not possible to test whether Muntoni et al.'s (2021) findings, which referred to these relations in the domain of reading, would have replicated within the present sample. Moreover, it was not possible to examine the interplay between gender stereotypes and self-concepts in different subjects. For example, numerous studies examining the relations between students' math and verbal achievements and self-concepts have demonstrated that students' achievement in one domain negatively affects their self-concept in the other domain, after controlling for achievement in the other domain (Möller et al., 2020). It would be conceivable that similar results also emerged for the relations between math and verbal gender stereotypes and self-concepts. Hence, future research should investigate the relations between math- and reading-related gender stereotypes and self-concepts within the same sample.

Finally, a limitation of the present research involves the fact that students' stereotypes were only measured through self-reports. It is conceivable that students' implicit gender stereotypes (shared in the classroom) affect their self-concepts beyond what is explicitly expressed. This seems particularly plausible considering that previous studies have found some gender-specific differences in the existence of explicit and implicit gender stereotypes, although the results were somewhat inconsistent (e.g., Steffens and Jelenec, 2011; vs. Passolunghi et al., 2014; see also Nosek et al., 2002). Researchers should thus feel encouraged to measure students' subject-specific gender stereotypes using both explicit and implicit measures in future research. Furthermore, it would seem worthwhile to supplement explicit self-concept measures with implicit measures of subject-specific self-concepts (e.g., Wolff et al., 2020a), given that gender stereotypes, especially if measured implicitly, might affect students' self-concepts especially at an unconscious level.

## CONCLUSION

The present research provided new significant insights into the interaction of gender and gender stereotypes in the formation of students' academic self-concepts. For the first time, it was shown that not only students' individual math-related gender stereotypes but also the math-related gender stereotypes shared in their classrooms affect students' math self-concepts. By analogy with Muntoni et al. (2021), who found a negative contextual effect of classmates' shared reading-related gender stereotypes on boys' reading self-concepts, this research revealed a negative contextual effect of classmates' shared math-related gender stereotypes on girls' math self-concepts. Taken together, it seems that stereotyped gender groups suffer from both individual and shared gender stereotypes. Given this, it is to be hoped that gender stereotypes favoring girls or boys in a certain domain will continue to decrease in the future.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

## REFERENCES

- Arens, A. K., Becker, M., and Möller, J. (2017). Social and dimensional comparisons in math and verbal test anxiety: within- and cross-domain relations with achievement and the mediating role of academic self-concept. *Contemp. Educ. Psychol.* 51, 240–252. doi: 10.1016/j.cedpsych.2017.08.005
- Asparouhov, T., and Muthén, B. O. (2012). *Multiple Group Multilevel Analysis. Mplus Web Notes, 16*. Available online at: <http://www.statmodel.com/examples/webnotes/webnote16.pdf> (accessed April 7, 2021).
- Berndt, T. J. (1979). Developmental changes in conformity to peers and parents. *Dev. Psychol.* 6, 608–616. doi: 10.1037/0012-1649.15.6.608
- Berndt, T. J., and Murphy, L. M. (2002). Influences of friends and friendships: myths, truths, and research recommendations. *Adv. Child Dev. Behav.* 30, 275–310. doi: 10.1016/S0065-2407(02)80044-1
- Bouchev, H. A. (2004). Parents, teachers, and peers: discrepant or complementary achievement socializers? *N. Dir. Child Adolesc. Dev.* 106, 35–53. doi: 10.1002/cd.115
- Brown, B. B. (2011). "Popularity in peer group perspective: the role of status in adolescent peer systems," in *Popularity in the Peer System*, eds A. H. N. Cillessen, D. Schwartz, and L. Mayeux (New York, NY: Guilford), 165–192.
- Bussey, K., and Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychol. Rev.* 106, 676–713. doi: 10.1037/0033-295X.106.4.676
- Cheryan, S., and Bodenhausen, G. V. (2000). When positive stereotypes threaten intellectual performance: the psychological hazards of "model minority" status. *Psychol. Sci.* 11, 399–402. doi: 10.1111/1467-9280.00277
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.
- Cvencek, D., Meltzoff, A. N., and Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Dev.* 82, 766–779. doi: 10.1111/j.1467-8624.2010.01529.x
- Deutsches Schulportal (2020). *Vorurteile beeinträchtigen Jungen beim Lesen [Prejudices Affect Boys Reading]*. Available online at: <http://deutsches-schulportal.de/expertenstimmen/vorurteile-beeintraechtigen-jungen-beim-lesen/> (accessed April 7, 2021).

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the Local Legislation and Institutional Requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

FW designed the study, prepared the research materials, organized the data collection, conducted the analyses, and wrote the manuscript.

## FUNDING

The publication was funded by the Open Access Fund of the University of Koblenz-Landau.

## ACKNOWLEDGMENTS

The author thanks Vivien Bogdanski for copyediting the manuscript.

- Dickhäuser, O., and Plenter, I. (2005). Letztes Halbjahr stand ich zwei": Zur Akkuratheit selbst berichteter Noten [On the accuracy of self-reported school marks]. *Z. Pädagog. Psychol. [German J. Educ. Psychol.]* 19, 219–224. doi: 10.1024/1010-0652.19.4.219
- Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., and Sczesny, S. (2020). Gender stereotypes have changed: a cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *Am. Psychol.* 75, 301–315. doi: 10.1037/amp0000494
- Eccles, J. S. (2009). Who am I and what am I going to do with my life? personal and collective identities as motivators of action. *Educ. Psychol.* 44, 78–89. doi: 10.1080/00461520902832368
- Eccles-Parsons, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). "Expectancies, values, and academic behaviors," in *Achievement and Achievement Motivation*, ed. J. T. Spence (San Francisco, CA: Freeman), 75–146.
- Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053
- Espelage, D. L., Holt, M. K., and Henkel, R. R. (2003). Examination of peer-group contextual effects on aggression during early adolescence. *Child Dev.* 74, 205–220. doi: 10.1111/1467-8624.00531
- Evans, A. B., Copping, K. E., Rowley, S. J., and Kurtz-Costes, B. (2011). Academic self-concept in black adolescents: do race and gender stereotypes matter? *Self Identity* 10, 263–277. doi: 10.1080/15298868.2010.485358
- Frawley, T. (2005). Gender bias in the classroom: current controversies and implications for teachers. *Child. Educ.* 81, 221–227. doi: 10.1080/00094056.2005.10522277
- Frenzel, A. C., Goetz, T., Pekrun, R., and Wattl, H. M. G. (2010). Development of mathematics interest in adolescence: influences of gender, family, and school context. *J. Res. Adolesc.* 20, 507–537. doi: 10.1111/j.1532-7795.2010.00645.x
- Geldhof, G. J., Preacher, K. J., and Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychol. Methods* 19, 72–91. doi: 10.1037/a0032138
- Gesellschaft für Empirische Bildungsforschung (GEBF) (2019). *[Society for Empirical Educational Research]. GEBF-Preise 2019 [GEBF Awards 2019]*.

- Available online at: <https://www.gefb-ev.de/deutsch/gefb-preise-2019/> (accessed April 7, 2021).
- Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., and Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychol. Sci.* 24, 2079–2087. doi: 10.1177/0956797613486989
- Goetz, T., Frenzel, A. C., Hall, N. C., and Pekrun, R. (2008). Antecedents of academic emotions: testing the internal/external frame of reference model for academic enjoyment. *Contemp. Educ. Psychol.* 33, 9–33. doi: 10.1016/j.cedpsych.2006.12.002
- Graham, J. W., Olchowski, A. E., and Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev. Sci.* 8, 206–213. doi: 10.1007/s11121-007-0070-9
- Greenwald, A. G., and Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychol. Rev.* 102, 4–27. doi: 10.1037/0033-295x.102.1.4
- Gunderson, E. A., Ramirez, G., Levine, S. C., and Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles* 66, 153–166. doi: 10.1007/s11199-011-9996-2
- Hartup, W. W. (1996). The company they keep: friendships and their developmental significance. *Child Dev.* 67, 1–13. doi: 10.2307/1131681
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., and Williams, C. C. (2008). Gender similarities characterize math performance. *Science* 321, 494–495. doi: 10.1126/science.1160364
- Ihme, T. A., and Möller, J. (2015). “He who can, does; he who cannot, teaches?”: stereotype threat and preservice teachers. *J. Educ. Psychol.* 107, 300–308. doi: 10.1037/a0037373
- Jacobs, J. E., and Eccles, J. S. (1992). The impact of mothers’ gender-role stereotypic beliefs on mothers’ and children’s ability perceptions. *J. Pers. Soc. Psychol.* 63, 932–944. doi: 10.1037//0022-3514.63.6.932
- Kurtz-Costes, B., Rowley, S. J., Harris-Britt, A., and Woods, T. A. (2008). Gender stereotypes about mathematics and science and self-perceptions of ability in late childhood and early adolescence. *Merrill Palmer Q.* 54, 386–409. doi: 10.1353/mpq.0.0001
- Lamb, M. E., Easterbrooks, M. A., and Holden, G. W. (1980). Reinforcement and punishment among preschoolers: characteristics, effects, and correlates. *Child Dev.* 51, 1230–1236. doi: 10.2307/1129565
- Lindberg, S. M., Hyde, J. S., Petersen, J. L., and Linn, M. C. (2010). New trends in gender and mathematics performance: a meta-analysis. *Psychol. Bull.* 136, 1123–1135. doi: 10.1037/a0021276
- Marsh, H. W., and Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspectives. *Perspect. Psychol. Sci.* 1, 133–163. doi: 10.1111/j.1745-6916.2006.00010.x
- Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., et al. (2009). Doubly-latent models of school contextual effects: integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behav. Res.* 44, 764–802. doi: 10.1080/00273170903333665
- Möller, J., Zitzmann, S., Helm, F., Machts, N., and Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Rev. Educ. Res.* 90, 376–419. doi: 10.3102/0034654320919354
- Muntoni, F., and Retelsdorf, J. (2019). At their children’s expense: how parents’ gender stereotypes affect their children’s reading outcomes. *Learn. Instr.* 60, 95–103. doi: 10.1016/j.learninstruc.2018.12.002
- Muntoni, F., Wagner, J., and Retelsdorf, J. (2021). Beware of stereotypes: are classmates’ stereotypes associated with students’ reading outcomes? *Child Dev.* 92, 189–204. doi: 10.1111/cdev.13359
- Muthén, L. K., and Muthén, B. O. (2017). *Mplus User’s Guide*, 8th Edn. Los Angeles, CA: Muthén & Muthén.
- Niepel, C., Stadler, M., and Greiff, S. (2020). Seeing is believing: gender diversity in STEM is related to mathematics self-concept. *J. Educ. Psychol.* 111, 1119–1130. doi: 10.1037/edu0000340
- Nosek, B. A., Banaji, M. R., and Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *J. Pers. Soc. Psychol.* 83, 44–59. doi: 10.1037//0022-3514.83.1.44
- Organisation for Economic Co-operation and Development (OECD) (2020). *PISA 2018 Results (Vol. 2): Where All Students Can Succeed*. Paris: OECD Publishing.
- Orth, U., Dapp, L. C., Erol, R. Y., Krauss, S., and Luciano, E. C. (2021). Development of domain-specific self-evaluations: a meta-analysis of longitudinal studies. *J. Pers. Soc. Psychol.* 120, 145–172. doi: 10.1037/pspp0000378
- Parker, P. D., Nagy, G., Trautwein, U., and Lüdtke, O. (2014). “Predicting career aspirations and university majors from academic ability and self-concept: a longitudinal application of the internal/external frame of reference model,” in *Gender Differences in Aspirations and Attainment: A Life Course Perspective*, eds I. Schoon and J. S. Eccles (Cambridge: Cambridge University Press), 224–246. doi: 10.1017/cbo9781139128933.015
- Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., and Eccles, J. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: a multi-context study. *Dev. Psychol.* 48, 1629–1642. doi: 10.1037/a0029167
- Passolunghi, M. C., Ferreira, T. I. R., and Tomasetto, C. (2014). Math-gender stereotypes and math-related beliefs in childhood and early adolescence. *Learn. Individ. Differ.* 34, 70–76. doi: 10.1016/j.lindif.2014.05.005
- Plante, I., Théorêt, M., and Favreau, O. E. (2009). Student gender stereotypes: contrasting the perceived maleness and femaleness of mathematics and language. *Educ. Psychol.* 29, 385–405. doi: 10.1080/01443410902971500
- Pohlmann, B. (2005). *Konsequenzen dimensionaler Vergleiche [Consequences of Dimensional Comparisons]*. Münster: Waxmann.
- Preckel, F., Goetz, T., Pekrun, R., and Kleine, M. (2008). Gender differences in gifted and average-ability students: comparing girls’ and boys’ achievement, self-concept, interest, and motivation in mathematics. *Gift. Child Q.* 52, 146–159. doi: 10.1177/0016986208315834
- Rowley, S., Kurtz-Costes, B., and Rouland, K. (2013). Self views of African American youth are related to the gender stereotypes and academic attributions of their mothers. *Self Identity* 12, 382–399. doi: 10.1080/15298868.2012.682360
- Ruble, D. N., Martin, C. L., and Berenbaum, S. A. (2007). “Gender development,” in *Handbook of Child Psychology: Social, Emotional, and Personality Development*, eds N. Eisenberg, W. Damon, and R. M. Lerner (Hoboken, NJ: Wiley), 858–932.
- Ryan, A. M. (2001). The peer group as a context for the development of young adolescent motivation and achievement. *Child Dev.* 72, 1135–1150. doi: 10.1111/1467-8624.00338
- Salikutluk, Z., and Heyne, S. (2017). Do gender roles and norms affect performance in maths? The impact of adolescents’ and their peers’ gender conceptions on maths grades. *Eur. Soc. Rev.* 33, 368–381. doi: 10.1093/esr/jcx049
- Schmader, T., Johns, M., and Barquissau, M. (2004). The costs of accepting gender differences: the role of stereotype endorsement in women’s experience in the math domain. *Sex Roles* 50, 835–850. doi: 10.1023/B:SERS.0000029101.74557.a0
- Schurtz, I. M., Pfost, M., Nagengast, B., and Artelt, C. (2014). Impact of social and dimensional comparisons on student’s mathematical and english subject-interest at the beginning of secondary school. *Learn. Instr.* 34, 32–41. doi: 10.1016/j.learninstruc.2014.08.001
- Skaalvik, S., and Skaalvik, E. M. (2004). Gender differences in math and verbal self-concept, performance expectations, and motivation. *Sex Roles* 50, 241–252. doi: 10.1023/B:SERS.0000015555.40976.e6
- Society for Research in Child Development (2020). *Do Girls Read Better Than Boys? If So, Gender Stereotypes May be to Blame*. Available online at: [https://www.eurekalert.org/pub\\_releases/2020-02/sfri-dgr021920.php](https://www.eurekalert.org/pub_releases/2020-02/sfri-dgr021920.php) (accessed April 7, 2021).
- Steele, C. M., and Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *J. Pers. Soc. Psychol.* 69, 797–811. doi: 10.1037/0022-3514.69.5.797
- Steffens, M. C., and Jelenec, P. (2011). Separating implicit gender stereotypes regarding math and language: implicit ability stereotypes are self-serving for boys and men, but not for girls and women. *Sex Roles* 64, 324–335. doi: 10.1007/s11199-010-9924-x
- Steffens, M. C., Jelenec, P., and Noack, P. (2010). On the leaky math pipeline: comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *J. Educ. Psychol.* 102, 947–963. doi: 10.1037/a0019920

- Studsrod, I., and Bru, E. (2011). Perceptions of peers as socialization agents and adjustment in upper secondary school. *Emot. Behav. Diffic.* 16, 159–172. doi: 10.1080/13632752.2011.569401
- Svetina, D., Rutkowski, L., and Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines: an illustration using Mplus and the lavaan/semTools packages. *Struct. Equ. Model. Multidiscip. J.* 27, 111–130. doi: 10.1080/10705511.2019.1602776
- Tymms, P. (2004). “Effect sizes in multilevel models,” in *But What Does It Mean? The Use of Effect Sizes in Educational Research*, eds I. Schagen and K. Elliot (London: National Foundation for Educational Research), 55–66.
- Voyer, D., and Voyer, S. D. (2014). Gender differences in scholastic achievement: a meta-analysis. *Psychol. Bull.* 140, 1174–1204. doi: 10.1037/a0036620
- Walton, G. M., and Cohen, G. L. (2003). Stereotype lift. *J. Exp. Soc. Psychol.* 39, 456–467. doi: 10.1016/S0022-1031(03)00019-2
- Wigfield, A., and Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps.1999.1015
- Wigfield, A., Eccles, J. S., Mac Iver, D., Reuman, D. A., and Midgley, C. (1991). Transitions during early adolescence: changes in children’s domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Dev. Psychol.* 27, 552–565. doi: 10.1037/0012-1649.27.4.552
- Wolff, F., Helm, F., Junge, F., and Möller, J. (2020a). Are dimensional comparisons performed unconsciously? An investigation of the internal/external frame of reference model using implicit self-concepts. *J. Educ. Psychol.* 112, 397–415. doi: 10.1037/edu0000375
- Wolff, F., Helm, F., and Möller, J. (2019a). Integrating the 2I/E model into dimensional comparison theory: towards a comprehensive comparison theory of academic self-concept formation. *Learn. Instr.* 62, 64–75. doi: 10.1016/j.learninstruc.2019.05.007
- Wolff, F., Nagy, G., Retelsdorf, J., Helm, F., Köller, O., and Möller, J. (2019b). The 2I/E model: integrating temporal comparisons into the internal/external frame of reference model. *J. Educ. Psychol.* 111, 1131–1161. doi: 10.1037/edu0000319
- Wolff, F., Sticca, F., Niepel, C., Goetz, T., Van Damme, J., and Möller, J. (2020b). The reciprocal 2I/E model: an investigation of mutual relations between achievement and self-concept levels and changes in the math and verbal domain across three countries. *J. Educ. Psychol. Adv. Online Publ.* 1–21. doi: 10.1037/edu0000632
- Wolff, F., Wigfield, A., Möller, J., Dicke, A.-L., and Eccles, J. S. (2020c). Social, dimensional, and temporal comparisons by students and parents: an investigation of the 2I/E model at the transition from elementary to junior high school. *J. Educ. Psychol.* 112, 1644–1660. doi: 10.1037/edu0000440
- Wolff, F., Zitzmann, S., and Möller, J. (2021). Moderators of dimensional comparison effects: a comprehensive replication study putting prior findings on five moderators to the test and going beyond. *J. Educ. Psychol.* 113, 621–640. doi: 10.1037/edu0000505
- Zucker, K. J., and Bradley, S. J. (1995). *Gender Identity Disorder and Psychosexual Problems in Children and Adolescents*. New York, NY: Guilford.

**Conflict of Interest:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Wolff. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.