Sex Roles (2011) 64:324–335
DOI 10.1007/s11199-010-9924-x

ORIGINAL ARTICLE

Separating Implicit Gender Stereotypes regarding Math and Language: Implicit Ability Stereotypes are Self-serving for Boys and Men, but not for Girls and Women

Melanie C. Steffens · Petra Jelenec

Abstract We investigated implicit gender stereotypes related to math and language separately, using Go/No-go Association Tasks. Samples were grade 9 adolescents ($N=187$) and university students ($N=189$) in Germany. Research questions concerned the existence of and gender differences in implicit stereotypes. While typical explicit-stereotyping findings were replicated, implicit math-male stereotypes were found in male, but not in female participants. Females revealed strong language-female stereotypes, whereas males showed language-male counterstereotypes. Thus, females' implicit math-gender stereotypes were the only ones that did not link own gender to the respective academic domain in a self-serving way. Further, females' stronger stereotypes were related to lower and males' to higher scores on constructs related to math ability, corroborating implicit stereotypes' importance.

Keywords Math-gender stereotypes · Implicit stereotypes · Gender differences

Introduction

In many countries, women are still underrepresented in math-intensive careers and earn only a small percentage of university diplomas in these fields (e.g., 15.4% of computer science diplomas in Germany in 2004) (Ramm and Bargel 2005). Cross-cultural variations in the gender gap in math achievement are huge, ranging from a large girls' advantage in many countries—Bahrain, Cyprus, Jordan, Singapore, Russia, and Kazakhstan—to the large boys' advantage common in Western countries (Bos et al. 2008; Mullis et al. 2003). Moreover, there are cultural variations in the strength of the relationship between gender, interest, and achievement (Evans et al. 2002). This speaks against a biological calling for a cultural explanation of such gender gaps.

Gender stereotypes regarding academic domains contribute to gender differences in performance, as U.S.-based studies have shown. Negative stereotypes concerning women's math ability can undermine women's performance goals (Smith 2006); they can further undermine math performance and interest in math (e.g., Davies et al. 2002; Gresky et al. 2005; Lesko and Henderlong Corpus 2006; Spencer et al. 1999); finally, gender stereotypes regarding academic domains influence ability self-concepts that, in turn, are crucial for career choices (e.g., Eccles 1994). Women who endorse math-gender stereotypes have more negative math ability self-concepts and are less interested in math careers (Schmader et al. 2004). Whereas several recent studies have demonstrated the power of *implicit gender stereotypes*, none of them has assessed implicit math-gender stereotypes separately from other stereotypic domains. The aim of the present study was doing so in female and male German adolescents and university students.

When asked directly, male and female students in the U.S.A. sometimes disavow math-gender stereotypes, claiming that men and women have similar math abilities (e.g.,

M. C. Steffens (✉) · P. Jelenec
Institut für Psychologie, Friedrich-Schiller-Universität Jena,
Am Steiger 3, Haus 1,
07743 Jena, Germany
e-mail: melanie.steffens@uni-jena.de

Ambady et al. 2001; Hyde et al. 1990; Steele 2003). However, it is still possible that students possess negative stereotypes regarding women's math abilities. For instance, in an Italian sample, the math performance of the very girls who disavowed holding math-gender stereotypes was undermined by highlighting these stereotypes (Muzzatti and Agnoli 2007). Such stereotypes that are not openly admitted can be conceptualized as associations between gender and stereotypic attributes, for example *math-male* and *language-female*, and these associations can differ in strength. These so-called implicit stereotypes can be activated quickly and automatically without intention or control, and they may influence behavior without the person's awareness of that impact (see Greenwald and Banaji 1995). In other words, whether questionnaire responses indicate gender stereotypes or not cannot be equated with the presence or absence of implicit stereotypes.

Much evidence for implicit stereotypes has been accumulated using reaction-time based implicit measures. Specifically, Implicit Association Tests (IATs) (Greenwald et al. 1998) have shown that women and men hold implicit math-gender stereotypes. In detail, American college students revealed stereotypes, associating math with male and arts with female: They were able to react faster to a task where math and male-related words required one reaction and arts and female-related words required the other reaction, than to a task where math and female-related words required one reaction and arts and male-related words required the other reaction (Nosek et al. 2002a, b). These implicit math-gender stereotypes also demonstrated unique predictive power regarding math-related outcome variables. Stronger implicit stereotypes were related to a stronger math preference, higher math identification, and better math performance for men. In contrast, for women, stronger implicit stereotypes were related to a lower math preference, lower math identification, and lower math performance (Nosek et al. 2002b). In a prospective study with female college students in the U.S.A., stronger implicit math-gender stereotypes predicted lower math performance and lower interests in math-related careers (Kiefer and Sekaquaptewa 2007b). Further, implicit math-gender stereotypes appeared to play a main role in undermining women's math performance (Kiefer and Sekaquaptewa 2007a). In a cross-national comparison, average implicit science-male stereotypes predicted by-country gender gaps in science and math performance, with stronger stereotypes going along with a larger gap between males' and females' performance (Nosek et al. 2009).

As crucial career decisions are made during school years, it is important to consider children's and adolescents' implicit math-gender stereotypes. Evidence for implicit math-gender stereotypes has been found already in elementary school children in the U.S.A. (Cvencek et al. in press). Steffens et al. (2010) assessed implicit math-language gender stereotypes with IATs in children and adolescents in Germany aged, on average, 9, 13, and 15 years. Girls aged 9 years already revealed implicit math-language gender stereotypes, and adolescent girls aged 13 and 15 years showed stronger implicit math-gender stereotypes than boys. These implicit stereotypes, along with their explicit counterparts, were early established and then apparently remained stable across age. In contrast to girls, significant implicit gender stereotypes could not be detected in German boys of any age group.

Math-gender stereotypes have been assumed to affect gender differences in math-related ability self-concepts (Eccles 2005). Domain-specific ability self-concepts, in turn, have been shown to influence math achievement in Australia (Marsh and Yeung 1997), and they can have a greater impact on subsequent course selections than math grades, as a German study showed (Köller et al. 2000). As meta-analyses comprising samples from many countries have shown, boys' higher math self-concepts relative to girls' are particularly pronounced in adolescence, and they exceed by far actual performance differences (Hyde et al. 1990a, b, 2008). Muzzatti and Agnoli (2007) speculated that by eighth grade, Italian adolescents' math-gender stereotype is internalized, and "participants are not aware of (or deny) the stereotype, but it is present implicitly" (p. 757–758). In line with these assumptions, Steffens et al. (2010) showed that for adolescent girls in Germany, implicit math-gender stereotypes predicted academic self-concepts, achievement, and enrolment preferences above and beyond explicit math-gender stereotypes. Explicit math-gender stereotypes were reported by girls and boys of all investigated age groups (grades 4–9). In line with this, teachers in Germany often hold gender-stereotyped views of children's abilities (Rustemeyer 1999). In a ranking of 61 countries whose inhabitants scored highest on implicit gender-science stereotyping, Germany obtained rank 23 (and the U.S.A., rank 36) (Nosek et al. 2009, online Appendix, http://www.pnas.org/content/suppl/2009/06/30/0809921106.DCSupplemental/Appendix_PDF.pdf).

More general work on gender stereotypes in Germany has confirmed many similarities with those in the U.S.A (Eckes 1997). Specifically, in both countries, perceptions of change were greater for women's than men's gender roles; women were projected to increase in masculine personality, cognitive, and physical attributes, to increase in feminine cognitive characteristics, but to remain stable in feminine positive personality attributes; in contrast, the stereotypes of men were far less dynamic (Wilde and Diekman 2005).

The stronger implicit math-language-gender stereotypes in adolescent girls than boys found by Steffens et al. (2010) in Germany call for an explanation, particularly as gender differences were not found in math-arts-gender stereotypes in adult or child samples in the U.S.A. (Cvencek et al. in press; Nosek et al. 2002a, b). Moreover, implicit associations of the concepts *math* and *language* (or *math* and *arts*) are intertwined in an IAT effect and cannot be separated (Nosek et al. 2005). In other words, if adolescent girls show stronger implicit stereotypes as compared to boys, these girls may hold either stronger math-boys stereotypes, or they may hold stronger language-girls stereotypes, or both. On the one hand, girls may have acquired stronger math-gender stereotypes than boys because these stereotypes are activated in girls more often than in boys, for example during math or science tests. On the other hand, as studies in the U.S.A. and Australia have shown, girls outperform boys on various verbal tasks and also have shown higher verbal self-concepts than boys (Hyde and Kling 2001; Hyde and Linn 1988; Marsh 1989). Being aware of this could have strengthened girls' as compared to boys' language stereotypes.

Summary of Hypotheses

Taken together, stronger implicit math-gender stereotypes in girls than boys were found in a previous study with German adolescents as participants (Steffens et al. 2010), but not with American college students (Nosek et al. 2002b). In order to shed light on these contrasting findings, participants in the present study were ninth graders and university students, comparable to the age groups in those studies. In order to investigate whether the stronger implicit math-gender stereotypes in girls than boys found by Steffens et al. (2010) are due to stronger math-boys or language-girls stereotypes, the present study used Go/No-Go Association Tasks (GNATs) (Nosek and Banaji 2001). GNATs have been developed as a measurement tool for implicit associations of a single concept with an attribute pair (e.g., *math* with *male* vs. *female*). Two separate GNATs were applied to measure implicit math-boys and language-girls stereotypes. We investigated the existence of implicit stereotypes as well as gender differences in them. Concurrently, explicit stereotypes, the respective ability self-concepts, and achievement were assessed both in order to describe our sample and for testing relations between implicit stereotypes and ability-related outcomes. Based on the IAT findings by Steffens et al. (2010) according to which females in Germany hold stronger joint implicit gender stereotypes concerning math and language than males do, we tested the following hypotheses:

Hypothesis 1: Females hold stronger *math* gender stereotypes than males.

Hypothesis 2: Females hold stronger *language* gender stereotypes than males.

Further, on the basis of studies with explicit measures described in the introduction, we tested whether these findings extend to implicit measures:

Hypothesis 3: Both females and males hold significant implicit math gender stereotypes.

Hypothesis 4: Both females and males hold significant implicit language gender stereotypes.

As implicit math-gender stereotypes have been found to be early established and then rather stable (Steffens et al. 2010), we did not expect differences between age groups. In order to test whether our sample generally shows typical gender differences and similarities, several explicit measures were used, among them explicit math and language gender stereotypes. Further, we investigated explicit ability self-concepts and school grades. Typically, gender differences in the math self-concept favor boys and gender differences in the language self-concept favor girls (or the self-concept related to the respective school subject, in our case, German); this was found both in Germany and Australia (see Hannover 1991; Marsh 1989). Further, whereas girls tend to outperform boys on language-related grades, often no gender differences are observed in math grades; for instance, this was shown both in Germany and in the U.S.A. (cf. Hannover 1991; Kimball 1989). An additional hypothesis was derived to test the power of implicit stereotypes:

Hypothesis 5: Implicit math and language gender stereotypes are related to explicit ability self-concepts and grades in the respective domain.

Method

Participants

The first part of our sample initially comprised data of $N=195$ participants attending ninth grade in various grammar schools (highest school track, Gymnasium) and secondary schools (intermediate school track, Realschule) in the East and West of Germany. Permissions to conduct the study were granted by school principles and parents. The adolescents participated in the study voluntarily during regular school hours. Eight participants who had higher error rates than 30% in at least one combined GNAT task were removed from all analyses, with $N=187$ participants remaining (mean age=14 years 10 months; 91 boys and 96 girls); 90 of them had always lived in Eastern Germany, 75, in Western Germany; only 14 indicated that German was not their first language. Neither school track nor East/West origin nor first language had any effects on results of implicit and explicit measures, so data were collapsed across these factors.

Additionally, 192 students at a large university in Eastern Germany participated in the study. Three participants with higher error rates than 30% in at least one combined GNAT task were excluded from all analyses. Data of 189 participants (mean age=22.1 years, range=18–35 years; 71 men, 118 women) were included. Among women, 85 (72%) studied a non-math major (i.e., liberal arts, social sciences, law), whereas 33 (28%) studied a math-intensive major (i.e., economics, science, math, medicine, engineering, computer science). Among men, 41 (58%) studied a non-math major and 30 (42%), a math-intensive major. These proportions were significantly different, $\chi^2_{(1)} = 4.07$, and are in line with gender-specific choices of majors in Germany. University students received either course credit or a chocolate bar for participation; ninth graders were rewarded with small gifts.

Materials

Implicit Measures

The math-gender GNAT consisted of the concept *math* (stimuli: computation, equation) and the concept pair *boys/men* (stimuli: boys, son) vs. *girls/women* (stimuli: girls, daughter) for ninth graders vs. adults, respectively. The language-gender GNAT used the concepts *language* (stimuli: poem, composition) and, again, *boys/men* vs. *girls/women*. Stimuli were selected to bear as few additional connotations as possible (Steffens et al. 2008). Further, two distractor stimuli related to the broader concept *school* were used in the GNATs (*school break, school bus* for ninth graders; *break, dorm* for adults): Adding stimuli of a super-ordinate category to the no-go trials contributed to a somewhat larger GNAT effect (Nosek and Banaji 2001).

For computing the GNAT effect in each GNAT, the difference between each participant's average reaction times in the two critical tasks was divided by the participant's overall standard deviation of the response latencies in these tasks (see Greenwald et al. 2003). Error reaction times were included in analyses. The GNAT effects for go trials with odd versus even position numbers in the critical tasks of the math-gender stereotype GNAT correlated with $r=.42$ and .39 (ninth graders and university students, respectively), those for the language-gender stereotype GNAT with $r=.34$ and .43, which is in the expected order of magnitude (e.g., Ebert et al. 2009).

Explicit Ability Self-concepts in Math and German and School Grades

Math self-concept was assessed with four items, for example, "I learn things quickly in math" (see Appendix). All explicit ratings were made on 1–5 scales, with smaller numbers indicating stronger agreement. Four parallel items were used for the German ability self-concept. All self-concept items were recoded; higher values indicate higher ability estimations. Both the math and the German ability self-concept scale revealed high internal consistencies, with Cronbach's $\alpha=.92$ (.91 and .93 for ninth graders and university students, respectively) and .88 (.86 and .90), respectively.

All participants were asked to indicate their latest report grades in math and German, with school grades in Germany ranging from "1" indicating "very good" to "5", "failed". School grades were then recoded so that larger values indicate better performance.

Explicit Gender Stereotypes

In order to assess stereotype endorsement, participants were asked about their agreement to four statements referring to the giftedness of boys and girls in math or German, "Boys (girls) are often talented for doing German (math)". Two further items captured comparative gender stereotypes, "Math (German) is rather a typical subject for ...", using *boys* and *girls* as anchor points of the scale (cf. Nosek et al. 2002b). In addition, stereotype awareness was assessed by asking participants to estimate to what extent they perceive most other people, in general, to hold gender stereotypes.

To obtain the score for explicit *math-boys* stereotyping, boys' giftedness rating in math was subtracted from girls' giftedness rating, with a higher value indicating stronger stereotype endorsement (possible range of the difference score: −4 to 4). The explicit *language-girls* stereotype score was computed likewise. In a second step, these indices were transformed to a value range between 1 and 5 and averaged with the respective comparative stereotype item. Correlations between computed scores and comparative scores were $r=.54$ (.52 and .48 for ninth graders and university students, respectively) and r=.46 (.52 and .36) for math and language, respectively.

Demographic Questions

Demographic questions were presented at the end of the study in order to avoid gender priming effects.

Procedure

After giving their informed consent, participants were tested in groups of up to seven by a female experimenter. GNATs and explicit measures were administered on iBooks. GNATs were always administered first to avoid priming stereotypes through explicit questions. Then, participants filled out the self-report measures in the order described above by clicking with the computer mouse on the response option they chose. Each question was presented separately and disappeared after a response was made. Finally, all participants were debriefed and rewarded. The study lasted 20–25 min.

Each GNAT consisted of two crucial reaction-time tasks. For example, in the stereotype-congruent task of a math-gender GNAT, participants are asked to respond to stimuli that are either math-related (e.g., calculus) or male-related (e.g., boy) by pressing the spacebar, whereas female-related and other distractor stimuli are to be ignored. In the stereotype-incongruent task, responses are required to math-related or female-related stimuli, whereas male stimuli and other distractor stimuli are ignored. Participants holding strong math-gender stereotypes should react faster in the *math-male* than in the *math-female* task.

Concepts were visible throughout a GNAT task, and stimuli were shown on the computer screen for 1,000 ms. Participants were asked to press the space bar as fast as possible if a stimulus belonged to one of the concepts (go trial). If a stimulus did not belong to either concept, participants were instructed to do nothing, and the stimulus disappeared after 1,000 ms (no-go trial). False responses were indicated by a flashing "F!". Half of the stimuli in a critical GNAT task required a go-response. Each critical task consisted of 60 trials (plus two practice trials at the beginning). Additional practice tasks requiring responses to only one concept (six trials) were inserted whenever new concepts were introduced.

The order of the GNATs (math vs. language) was counterbalanced. Further, half of the participants started with the stereotype-congruent task in both GNATs (*language-girls*, *math-boys*), the other half with the stereotype-incongruent task (*language-boys*, *math-girls*).

Design

Dependent variables were GNAT effects in the math-gender and language-gender GNAT. Gender and age were treated as independent variables. Two additional control factors (GNAT order and task order within the GNAT) were included, resulting in a 2×2×2×2 between-subjects design. Small gender differences in GNAT effects with an effect size of d=.30 could be detected with α=.05 and our sample size with a statistical power of $1-\beta$=.82 (Faul et al. 2007).

Results

All measures were tested for outliers, and violations of assumptions of statistical tests (e.g., homoscedasticity) were tested throughout the present studies; none were found. Unless indicated differently, all statistical tests in this article were conducted with α=.05, individual p-values are not reported for statistically significant effects, and η_p^2 is reported as an indicator of the effect size.

Description of Samples on Explicit Measures

Before turning to our substantive hypotheses, we look at average agreement and gender differences regarding explicit stereotyping and ability-related constructs in order to test whether typical findings are replicated in our sample. Table 1 shows male and female participants' average endorsement of explicit stereotypes along with ability self-concepts and grades with regard to math and German. As can be seen, all means for stereotypes exceed 3, showing on average stereotype endorsement and awareness both by boys and girls (see Table 2 for separate items). Females' ability self-concepts in German and math appear generally higher than males', mirroring grade differences, except for university students' math grades and ability self-concepts that are rather similar for women and men. In order to test for overall gender differences in explicit measures, a MANOVA with gender and age as independent variables was conducted on these eight dependent variables. It showed multivariate main effects of gender, F (8, 354)= 5.44, $\eta_p^2 = .11$, and of age, F (8, 354)=16.51, $\eta_p^2 = .27$, but no interaction, F (8, 354)<1.62. Respective univariate tests are reported below.

Explicit Gender Stereotypes

Whereas there was no gender difference in explicit math stereotypes, F (1, 361)<1.01, females more strongly endorsed stereotypes that females are talented for

Table 1 Mean explicit endorsement and awareness of gender stereotypes, mean ability self-concepts, and school grades (with standard deviations) reported by male and female ninth graders and university students

Participants	Stereotype endorsement		Stereotype awareness		Ability self-concepts		Grades	
	Math-male	Language-female	Math-male	Language-female	Math	German	Math	German
Ninth graders								
Male students	3.29[a] (.69)	3.54[a] (.63)	3.52[a] (1.15)	3.90[a] (.82)	3.19 (1.02)	3.26 (.77)	2.95 (1.00)	3.26 (.76)
Female students	3.26[a] (.66)	3.67[a] (.59)	3.97[a] (1.04)	4.06[a] (.88)	3.39 (1.03)	3.60 (.86)	3.46 (.97)	3.71 (.74)
University students								
Male students	3.62[a] (.45)	3.68[a] (.46)	4.45[a] (.60)	4.27[a] (.77)	3.00 (1.12)	3.34 (.95)	3.39 (1.11)	3.67 (.83)
Female students	3.73[a] (.55)	3.83[a] (.51)	4.56[a] (.66)	4.47[a] (.61)	3.08 (1.22)	3.67 (.92)	3.60 (1.10)	4.03 (.84)
Means across age groups								
Male students	3.43 (.62)	3.60[b] (.57)	3.93[b] (1.06)	4.06[b] (.82)	3.11 (1.07)	3.29[b] (.85)	3.15[b] (1.08)	3.44[b] (.82)
Female students	3.52 (.64)	3.76[c] (.55)	4.29[c] (.77)	4.29[c] (.77)	3.22 (1.15)	3.64[c] (.89)	3.54[c] (1.04)	3.89[c] (.81)

Higher values indicate stronger endorsement/higher ability, with possible values between 1 and 5. Stereotype awareness refer to questions starting with "How would most people judge..."

[a] Indicates that the group mean in stereotyping is significantly different from the scale mean

[b,c] Indicates that males' and females' scores averaged across age groups are significantly different

language than males did, F (1, 361)=6.19, $\eta_p^2 = .02$, and were more aware of stereotypes in their environment, F (1, 361)=11.88, $\eta_p^2 = .03$ and F (1, 361)=4.17, $\eta_p^2 = .01$, for math and language, respectively. All effect sizes associated with these gender differences are small, though. Age effects on all stereotyping indices showed that university students indicated to endorse stereotypes more strongly and were more aware of stereotypes regarding both math and language than ninth graders, Fs (1, 361)=38.85, 7.40, 67.29, and 22.12, $\eta_p^2 s = .10, .02, .16, \text{and} .06$, respectively. Effect sizes indicated that age differences in math stereotypes are larger than those in language stereotypes. There were no statistically significant age × gender interactions on endorsed and perceived math and language stereotypes (all Fs<3.36, all ps>.06, all $\eta_p^2 s < .01$).

One-sample *t*-tests against the neutral value of the scale (i.e., 3) were carried out separately for the four participant groups (adjusted α=.0125) (Bortz 1999). As expected, boys and girls, men and women endorsed traditional gender stereotypes regarding math and language (all ts>3.79 with dfs>70 and $\eta_p^2 s \geq .13$). Similarly, each group was aware of other people bearing gender stereotypes regarding math and language (all ts>4.28 with dfs>70 and $\eta_p^2 s \geq .17$). In other words, our sample is rather typical with regard to explicit stereotyping.

Explicit Ability Self-concepts and School Grades

Males and females did not differ in their math self-concepts, F (1, 361)<1.60. However, females showed the typical better language self-concept compared to males, F (1, 361)=13.84, $\eta_p^2 = .04$. In our sample, females earned both better math and German grades than males, F (1, 361)=10.58, $\eta_p^2 = .03$, and F (1, 361)=23.09, $\eta_p^2 = .06$, respectively. Thus, explicit ability self-concepts and school grades in German showed the typical advantage of females over males. Females even

Table 2 Mean stereotype endorsement (with mode; and standard deviation) on each explicit gender stereotype item, reported by male and female ninth graders and university students

Participants	Boys' math talent	Girls' math talent	Math is a boys'–girls' subject	Boys' German talent	Girls' German talent	German is a boys'-girls' subject
Ninth graders						
Male students	3.32 (4; .89)	3.07 (3; 1)	2.55 (3; .81)	2.71 (3; .73)	3.73 (4; .88)	3.57 (3; .79)
Female students	3.23 (3; .98)	3 (3; .87)	2.60 (3; .90)	2.45 (2; .77)	3.74 (4; .84)	3.70 (4; .81)
University students						
Male students	3.58 (4; .71)	2.77 (3; .77)	2.17 (2; .59)	2.59 (3; .71)	3.94 (4; .83)	3.69 (4; .67)
Female students	3.65 (4; .85)	2.72 (3; .65)	2.01 (2; .69)	2.45 (2; .65)	3.87 (4; .72)	3.96 (4; .70)

Higher values on talent questions indicate stronger endorsement ("1 = little talent, 5 = much talent"). School subject questions were anchored with "1 = boys' subject" and "5 = girls' subject" had better math grades than males. However, females seemed to underestimate their math ability because they did not rate their math ability higher than males did despite their better achievements. This pattern is therefore consistent with the more common finding of females receiving comparable math grades as males, but showing a lower math ability self-concept. The only age differences in explicit ability self-concepts and grades were better grades in our university student sample than in the ninth graders, F (1, 361)=7.50, $\eta_p^2 = .02$ and F (1, 361)=18.62, $\eta_p^2 = .05$ in math and German, respectively. There were no age × gender interactions on any of these variables (all Fs<1.91). In summary, gender differences and similarities in ability self-concepts and grades corroborate that our sample is rather typical. Against this backdrop, implicit gender stereotypes will be analyzed.

Implicit Gender Stereotypes

Larger effects in the GNAT indicate stronger math-male and language-female stereotypes, GNAT scores of 0 are interpreted as no stereotypes as they indicate that reaction times in the math-male (language-female) task were comparable to those in the math-female (language-male) task. As shown in Fig. 1, only males revealed math-gender stereotypes, whereas females, on average, did not show stereotypes. Similarly, females showed language-gender stereotypes, whereas males showed language-male counterstereotypes. A joint 2×2×2×2×2 ANOVA of the math and language gender stereotypes with the repeated-measures factor math versus language GNAT and the between-subjects factors gender, age, GNAT order, and GNAT task order yielded a large interaction effect of participant gender with math vs. language stereotype, F (1,360)=78.13, $\eta_p^2 = .18$. Focused tests of our hypotheses were subsequently conducted separately for math and language gender stereotypes.

Implicit Math-gender Stereotypes

The 2×2×2×2 ANOVA corroborated that men held stronger math-gender stereotypes than women did, F (1, 360)=15.80, $\eta_p^2 = .04$, contrasting Hypothesis 1 that females hold stronger stereotypes than males. Additionally, a main effect of task order, F (1, 360)=29.59, $\eta_p^2 = .08$, indicated that GNAT effects were biased in the direction of the task done first, as is often the case (cf. Greenwald et al. 1998) (all other Fs<2.50). No interaction gender × age group was found, F (1, 360)=2.49, p>.11, indicating that gender differences were comparable for adolescents and university students. One sample t-tests against 0 separately for male and female participants (α=.025) revealed math stereotypes for males, t (161)=5.54, $\eta_p^2 = .16$, but not for females, t (213)<1. The former finding, but not the latter, is in line with Hypothesis 3 that we find implicit math-gender stereotyping in each gender group. In sum, in the absence of a gender difference in explicit math stereotypes, implicit math-gender stereotypes were found only in males.

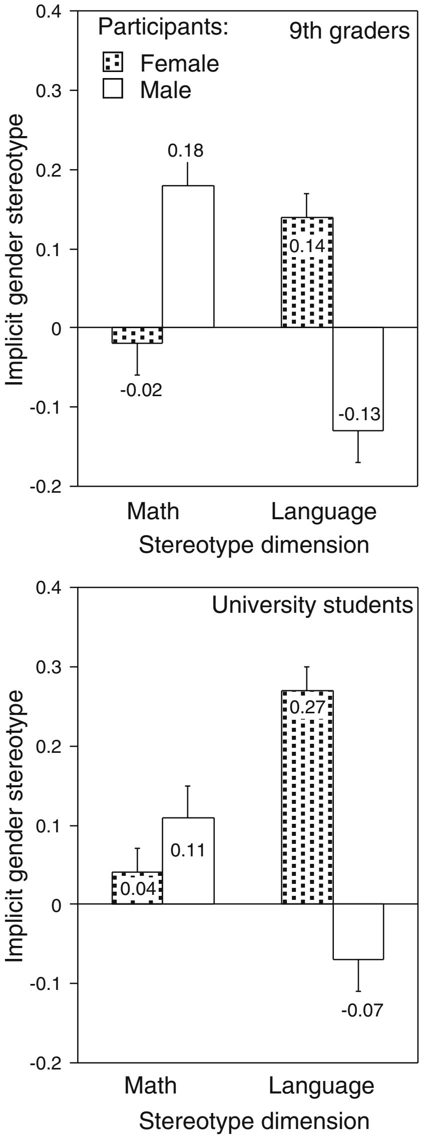

Fig. 1 Mean implicit gender stereotypes (D-like GNAT Effects) for math and language, separately for boys and girls (ninth graders, *upper panel*) and university students (*lower panel*). Positive effects indicate stereotypes math-boys and language-girls, negative effects indicate counterstereotypes. *Error bars* reflect standard errors of means. The gender difference is significant on each stereotype dimension, and each (counter)stereotype is significantly different from zero except females' math-gender stereotype

Implicit Language-gender Stereotypes

The same analysis as above on language-gender stereotypes revealed main effects of gender, F (1, 360)=69.17,

$\eta_p^2 = .16$, and task order, F (1, 360)=25.24, $\eta_p^2 = .07$. Females' stronger language-gender stereotypes than males' corroborate Hypothesis 2. Again, no interaction gender × age group was found, F (1, 360)<1. A main effect age group, with university students showing stronger stereotypes than adolescents, F (1, 360)=8.86, $\eta_p^2 = .02$ (all other Fs<2.90), cannot be interpreted as the university students' sample consisted of more women than men. One-sample t-tests against 0 separately for male and female participants (α=.025) revealed language-gender stereotypes in female participants, t (213)=8.54, $\eta_p^2 = .25$, and unpredicted language-male counterstereotypes for male participants, t (161)=−3.91, $\eta_p^2 = .09$. The former finding is in line with Hypothesis 4 of implicit language-gender stereotypes in each gender group, the latter contradicts this hypothesis. In summary, whereas all participants explicitly endorsed traditional gender stereotypes regarding language, only our female participants held implicit language-stereotypes favoring females, and male participants appeared to hold implicit language-male counterstereotypes.

Prediction of Outcomes by Implicit Gender Stereotypes

To the degree that implicit stereotypes are meaningfully related to individual differences in behavior, they should predict ability-related concepts. We therefore tested whether implicit stereotypes predict females' and males' ability self-concepts and achievement in a series of regression analyses. Age group (1, adolescents versus −1, university students) did not interact with gender (1, male, versus −1, female) and z-transformed *implicit math-gender stereotype* scores in predicting the math ability self-concept (|β|<.04). We therefore conducted a combined analysis of both age groups, predictors being gender, implicit math-gender stereotypes, and their interaction. The overall regression model was significant, F (3, 372)=4.43, R^2=.03. Neither gender nor implicit math-gender stereotypes explained math ability self-concepts (|βs|<.05), but their interaction did, β=−.17. Simple slopes tests showed that for males, stronger implicit math-gender stereotypes went along descriptively with higher math ability self-concepts (β=.12, p=.13). In contrast, for females, stronger implicit math-gender stereotypes went along with significantly lower math ability self-concepts (β=−.21).

With regard to achievement, we computed the difference score (math grade minus German grade) that controls for general achievement level and turned out more sensitive than each single grade. With this achievement score as the dependent variable, the same analysis as above again showed no interaction with age group (|β|<.04). In the combined analysis of both age groups with the predictors gender, implicit math-gender stereotypes, and their interaction, the overall regression model was significant, F (3, 364)=3.72, R^2=.03. Neither gender nor implicit math-gender stereotypes explained differences in achievement (|βs|<.06), but their interaction did, β=−.18. Simple slopes tests showed that for males, stronger implicit math-gender stereotypes went along with significantly higher relative math achievement (β=.23). In contrast, for females, stronger implicit math-gender stereotypes went along with lower relative math achievement (β=−.13, p=.06).

In a nutshell, the regression analyses showed that stronger math-gender stereotypes went along with lower math ability self-concepts and achievement of females and higher math self-concepts and achievement of males. These findings indicate that implicit math-gender stereotypes are meaningfully related to self-reported outcomes and thus attest to the validity of our measures. *Implicit language-gender stereotypes* did not interact with participant gender in predicting verbal ability self-concepts or grade differences between math and German (both |βs|<.06).

Discussion

We separately investigated implicit *math-male* stereotypes and implicit *language-female* stereotypes in Germany. Participants were adolescents attending grade 9 and university students. On explicit measures of stereotyping and achievement-related concepts, our sample appears quite typical. Females were more aware of stereotypes linking math with males in their environment than males. Moreover, females endorsed stereotypes linking language with females more than males and were more aware of language-gender stereotypes than males, who nevertheless also indicated language-gender stereotypes. Girls and women reported comparable math self-concepts as male participants, along with higher verbal self-concepts than boys and men. Females earned both better math and German grades than males. Overall these findings indicate that our sample is rather typical.

Against this backdrop of females acknowledging stronger math stereotypes than males, males revealed stronger implicit math-gender stereotypes than females did, who, on average, did not show implicit stereotypes regarding math. Female participants showed stronger implicit stereotyping linking language with females than male participants did, which appears at first sight in line with the difference between genders in explicit language stereotyping. However, male participants revealed a counterstereotypic language-male stereotype while reporting a traditional language-female stereotype.

Females' math ability self-concepts and their relative math achievement were lower the stronger their implicit math-gender stereotypes were. In contrast, males' math ability self-concepts and their relative achievement were higher the stronger their implicit math-gender stereotypes. The latter findings provide evidence for the validity of our implicit measures and corroborate relations between implicit stereotypes and achievement.

Taken together, our study shows first, rather expected findings on explicit stereotypes and ability-related concepts; second, relations with implicit stereotypes corroborate the validity of the latter measures; and third, these very implicit stereotyping measures revealed unexpected gender differences. The current findings offer a plausible explanation for German girls showing stronger implicit gender stereotypes than boys when IATs are used (see Steffens et al. 2010). In the IAT effect, math-boys and language-girls stereotypes are combined in a joint stereotype score. Girls' language-girls stereotypes, combined with no math-gender stereotypes, would result in implicit stereotyping in the IAT due to the language stereotype. This is what was observed in that study. In contrast, for boys math-boys stereotypes and language-boys counterstereotypes should cancel each other out, resulting in a small if any IAT effect, again mirroring the previous findings.

The current findings do not easily map onto those Nosek et al. found (2002a, b), that is, no gender differences in gender stereotypes measured with IATs in adults. They used the concepts *math* versus (*liberal*) *arts*, whereas *math* versus *language* was used in the present studies. It is mere speculation to wonder whether the concept (*liberal*) *arts* is more strongly linked to the concept *female* than the concept *language*. If so, missing counterstereotypes in the IATs administered by Nosek and colleagues would explain that men also demonstrated implicit math-gender stereotypes. A similar reasoning could apply to the study by Cvencek and colleagues (in press) where associations of *math* and *reading* were assessed in American elementary school children. Alternatively, the differences in findings could reflect cultural differences, with stronger implicit stereotypes endorsed in the U.S.A. An indicator for that could be that U.S. findings related to gender stereotypes have not been replicated in previous German studies (Steffens et al. 2005; Steffens et al. 2009). In line with the present findings, a math-gender GNAT administered to a sample of female students in the U.S.A. showed no significant implicit stereotyping in the presence of a female experimenter (see below) (Nosek and Banaji 2002). Future research is needed to compare implicit math-gender stereotypes separately from other stereotype dimensions in female participants in the U.S.A. to those of males.

Implicit math-gender stereotypes were related to achievement variables. Two limitations regarding this finding should be mentioned. First, our findings are cross-sectional; second, reported school grades may be somewhat distorted in a gender-stereotypic direction (e.g., Chatard et al. 2007). Nevertheless, we found that implicit math-gender stereotypes predicted women's low and men's high math achievement (see Steffens et al. 2010, for similar findings). Such findings are in line with the idea that stereotype endorsement plays a role in women's disengagement in math and science. No such effects were found with regard to implicit language stereotypes, which could be a hint that females' math achievement, but not males' language achievement, is hindered by ability-related implicit gender stereotypes. Further research is needed to test whether relations with males' achievement can be found if more "unmanly" domains are investigated (such as ballet?).

Along similar lines as our achievement-related findings, it has recently been shown that women with relatively strong stereotypes associating male with science were least likely to major in science (Smyth et al. 2009, unpublished). A tentative inclusion of the variable *math-intensive major* as an additional IV in our analysis of implicit math-gender stereotypes in university students showed that descriptively, men who studied math-intensive majors held stronger implicit math-gender stereotypes than those studying non-math majors (estimated marginal means = .16 vs. .08), whereas women with math-intensive majors held weaker math-gender stereotypes than those studying non-math majors (−.03 vs. .07). Even though this interaction was not statistically significant, $F\ (1,\ 173)=2.37$, $p<.13$, along with our regression findings, it confirms the potent link between implicit stereotyping and scientific self-concept Smyth and colleagues demonstrated with a much larger sample.

A limitation of the present study is that the internal consistency of the GNATs was low. If relations with other measures are in the focus of interest, IATs clearly appear more sensitive. The virtue of GNATs is precisely the reason why we used them here: To assess math-gender stereotypes separately from any other stereotype dimension. Future research should systematically test whether other implicit measures combine the virtues of GNATs with the better measurement quality of IATs (Steffens and Jonas 2010) or whether methodological improvements in the evaluation of IAT findings provide an alternative (Anselmi et al. in press).

The separate measurements of implicit math and language stereotypes yielded some findings that are unexpected at first sight, for example, the implicit counterstereotype regarding *language* in boys given their traditional explicit stereotypes that language is a girls' domain. Further, the lack of implicit stereotyping of math as male in girls and women seems to contradict numerous stereotype threat effects (e.g., Spencer et al. 1999). However we believe these findings can be reconciled. We assume that the implicit math-gender stereotypes in females have an exceptional position, as they do not reflect the typically found group-serving (and by extension, self-serving) associations

(Popa-Roch and Delmas 2010; van Ravenzwaaij et al. in press). Boys' implicit associations regarding math and language are consistent with their ingroup bias in gender ascriptions to persons who were good at math or spelling (i.e.: assuming these persons are male) (Steele 2003).

Whereas initially, it was assumed that implicit stereotypes essentially reflect cultural stereotypes and past socialization experiences (Greenwald and Banaji 1995), the pattern of implicit stereotypes we found extends the previous finding that "self and ingroup share desirable traits" (Rudman et al. 2001, p. 1164). Those authors demonstrated that people possess implicit gender stereotypes in self-favorable form because of the tendency to associate self with desirable traits. For example, men revealed stronger *men-powerful*/*women-weak* associations than women, particularly if power-related words were positive and weakness-related words negative in valence. If a talent for math and language is considered desirable and an implicit association of these subjects with a gender group is regarded as evidence for subjective perceptions of that talent, then our findings that girls associate girls with language and boys associate boys with both math and language are further evidence for self-serving implicit gender stereotypes. In our study the only exception to this pattern is that girls on average do not show self-serving math-girls stereotypes. This was the case in a situation with a female experimenter that avoided making gender salient.

We suspect that this lack of a self-serving association is an indicator for females' vulnerability to math-gender stereotypes. If females experience failure in a difficult math test and/or if math-gender stereotypes are made salient (e.g., by providing stereotypic test descriptions), implicit math-male stereotyping could increase rather easily and exert its detrimental influence as no self-serving implicit associations can act as a buffer. In fact, it has been demonstrated that the presence of a male rather than female experimenter is sufficient to activate females' math-gender stereotypes (Nosek and Banaji 2002). A related study has provided evidence that negative math-gender stereotypes weaken women's self-serving biases and make them less confident in their math abilities (Kiefer and Shih 2006).

According to our interpretation that a lack of a self-serving association is an indicator for females' vulnerability to math-gender stereotypes, boys or men should not be that much affected by language-gender stereotypes. Up to now, negative effects of language-gender stereotyping on men have been demonstrated only rarely (e.g., Keller 2007). Further, language-gender stereotypes may not be as threatening for men as math-gender stereotypes are for women. For example, bad performance in a task introduced as fitting women's abilities increased participants' perception of the male target person as masculine (Reinhard et al. 2008). It is quite probable that men would rather appreciate this consequence than fear it.

In summary, we found self-serving implicit gender stereotypes regarding math and language abilities, with the exception that females on average held no math-gender stereotypes. Individual differences in implicit math-gender stereotypes were related to individual differences in math achievement. In line with this, we assume that a lack of self-serving stereotypes is an indicator for females' vulnerability to math-gender stereotypes. Experimental research is needed to test whether activating implicit stereotypes linking math with male in female participants negatively affects math test performance and thus, such stereotypes contribute to the leakage of women from the math-science pipeline.

Acknowledgement This research is part of the doctoral dissertation of the second author and was funded by a grant from the DFG (Ste 938/9-1). We want to thank Janette Schult for valuable comments on an earlier draft of this paper, and Lars Buder, Ulrike Zieger, and Ulrike Leich for their help with regard to planning and running the studies.

Appendix: Explicit German Questions Asked, with Translations

Explicit self-concept

1–2 Mathe (Deutsch) macht mir Spaß./I like math (German).
3–4 Ich bin gut in Mathe (Deutsch)./I am good at math (German)
5–6 Ich lerne schnell in Mathe (Deutsch)./I learn things quickly in math (German).
7–8 Ich bin begabt für Mathe (Deutsch)./I am talented for doing math (German).

Grades

1–2 Was war Deine letzte Zeugnisnote in Mathe (Deutsch)?/What was your latest report grade in math (German)?

Explicit stereotypes[1]

1–4 Jungen (Mädchen) sind häufig begabt für Mathe (Deutsch)./Boys (girls) are often talented for doing math (German).

---

[1] For university students, Jungen (Mädchen) was replaced with Männer (Frauen)/men (women) throughout.

5–6 Mathe (Deutsch) ist eher ein typisches Jungenfach—Mädchenfach./Math is rather a typical boys' subject–girls' subject.

7–8 Was meinst Du? Wie würden die meisten Menschen Mathe einschätzen? Mathe (Deutsch) ist eher ein typisches Jungenfach—Mädchenfach./What do you think? How would most people judge math (German)? Math (German) is rather a typical boys' subject—girls' subject

References

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 12*, 385–390.

Anselmi, P., Vianello, M., & Robusto, E. (in press). Positive associations primacy in the IAT. A many-facet Rasch measurement analysis. *Experimental Psychology*.

Bortz, J. (1999). *Statistik für Sozialwissenschaftler (Statistics for the social sciences)*. Berlin: Springer.

Bos, W., Bonsen, M., Baumert, J., Prenzel, M., Selter, C., & Walther, G. (Eds.). (2008). *TIMSS 2007: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (TIMSS 2007: Primary school pupils' competencies in math and science in Germany in international comparison)*. Münster: Waxmann.

Chatard, A., Guimond, S., & Selimbegovic, L. (2007). "How good are you in math?" The eVect of gender stereotypes on students' recollection of their school marks. *Journal of Experimental Social Psychology, 43*, 1017–1024.

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (in press). Math-gender stereotypes in elementary-school children. *Child Development*.

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin, 28*, 1615–1628.

Ebert, I., Steffens, M. C., & Kroth, A. (2009, March–April). *Competent and cold men—competent and warm women: Implicit gender stereotypes correspond to changes in gender-role allocation*. Paper presented at the 51. Tagung experimentell arbeitender Psycholog|innen, Jena, Germany.

Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly, 18*, 585–609.

Eccles, J. S. (2005). Studying gender and ethnic differences in participation in math, physical science, and information technology. In J. E. Jacobs & S. D. Simpkins (Eds.), *Leaks in the pipeline to math, science, and technology careers* (pp. 7–14). San Francisco: Jossey-Bass.

Eckes, T. (1997). *Geschlechterstereotype: Frau und Mann in sozialpsychologischer Sicht (Gender stereotypes: Woman and men from a social psychology viewpoint)*. Pfaffenweiler: Centaurus.

Evans, E. M., Schweingruber, H., & Stevenson, H. W. (2002). Gender differences in interest and knowledge acquisition: The United States, Taiwan, and Japan. *Sex Roles, 47*, 153–167.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4–27.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464–1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197–216.

Gresky, D. M., Ten Eyck, L. L., Lord, C. G., & McIntyre, R. B. (2005). Effects of salient multiple identities on women's performance under mathematics stereotype threat. *Sex Roles, 53*, 703–716.

Hannover, B. (1991). Zur Unterrepräsentanz von Mädchen in Naturwissenschaften und Technik: Psychologische Prädiktoren der Fach- und Berufswahl (Underrepresentation of girls in the natural sciences and technology: psychological predictors of subject choice and for choice of occupation). *Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology, 5*, 169–186.

Hyde, S. J., & Kling, K. L. (2001). Women, motivation, and achievement. *Psychology of Women Quarterly, 25*, 364–378.

Hyde, S. J., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104*, 53–69.

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

Hyde, J. S., Fennema, E., Ryan, M., & Frost, L. A. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly, 14*, 299–324.

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

Keller, J. (2007). When negative stereotypic expectancies turn into challenge or threat: The moderating role of regulatory focus. *Swiss Journal of Psychology, 66*, 163–168.

Kiefer, A. K., & Sekaquaptewa, D. (2007a). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology, 43*, 825–833.

Kiefer, A. K., & Sekaquaptewa, D. (2007b). Implicit stereotypes, gender identification, and math-related outcomes: A prospective study of female college students. *Psychological Science, 18*, 13–18.

Kiefer, A., & Shih, M. (2006). Gender differences in persistence and attributions in stereotype relevant contexts. *Sex Roles, 54*, 859–868.

Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin, 105*, 198–214.

Köller, O., Daniels, Z., Schnabel, K., & Baumert, J. (2000). Kurswahlen von Mädchen und Jungen im Fach Mathematik: Die Rolle des fachspezifischen Selbstkonzepts und Interesses (Course selection of girls and boys in mathematics: the role of academic self-concept and interest). *Zeitschrift für Pädagogische Psychologie/German Journal of Educational Psychology, 14*, 26–37.

Lesko, A. C., & Henderlong Corpus, J. (2006). Discounting the difficult: How high math-identified women respond to stereotype threat. *Sex Roles, 54*, 113–125.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81*, 417–430.

Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology, 89*, 41–54.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2003). *Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill: TIMSS & PIRLS International Study Center, Boston College.

Muzzatti, B., & Agnoli, F. (2007). Gender and mathematics: Attitudes and stereotype threat susceptibility in Italian children. *Developmental Psychology, 43*, 747–759.

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-go association task. *Social Cognition, 19*, 625–666.

Nosek, B. A., & Banaji, M. R. (2002, February). *The power of the immediate situation: Gender differences in implicit math attitudes*. Paper presented at the Conference for the Society of Personality and Social Psychology, Savannah, GA. February 1, 2002.

Nosek, B. A., Banaji, M., & Greenwald, A. G. (2002a). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics, 6*, 101–115.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002b). Math = male, me = female, therefore math not = me. *Journal of Personality and Social Psychology, 83*, 44–59.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test II: Method variables and construct validity. *Personality and Social Psychology Bulletin, 31*, 166–180.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences, 106*, 10593–10597.

Popa-Roch, M., & Delmas, F. (2010). Prejudice IAT effects: The role of self-related heuristics. *Zeitschrift für Psychologie/Journal of Psychology, 218*, 44–50.

Ramm, M., & Bargel, T. (2005). *Frauen im Studium. Langzeitstudie 1983–2004 (Female students. Long-term study 1983–2004)*. Bonn: Bundesministerium für Bildung und Forschung (BMBF)/ Federal Ministry of Education and Research.

Reinhard, M. A., Stahlberg, D., & Messner, M. (2008). Failure as an asset for high-status persons—relative group performance and attributed occupational success. *Journal of Experimental Social Psychology, 44*, 501–518.

Rudman, L. A., Greenwald, A. G., & McGhee, D. E. (2001). Implicit self-concept and evaluative implicit gender stereotypes: Self and ingroup share desirable traits. *Personality and Social Psychology Bulletin, 27*, 1164–1178.

Rustemeyer, R. (1999). Geschlechtstypische Erwartungen zukünftiger Lehrkräfte bezügl. des Unterrichtsfaches Mathematik und korrespondierende (Selbst-)Einschätzungen von Schülerinnen und Schülern (Teachers' expectations concerning gender-specific performance in mathematics and the corresponding self-evaluation of male and female students). *Psychologie in Erziehung und Unterricht/Psychology in Education and Teaching, 46*, 187–200.

Schmader, T., Johns, M., & 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.

Smith, J. L. (2006). The interplay among stereotypes, performance-avoidance goals, and women's math performance expectations. *Sex Roles, 54*, 287–296.

Spencer, S. J., Steele, C. M., & Quinn, D. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4–28.

Steele, J. (2003). Children's gender stereotypes about math: The role of stereotype stratification. *Journal of Applied Social Psychology, 33*, 2587–2606.

Steffens, M. C., & Jonas, K. J. (2010). Implicit attitude measures [Editorial]. *Zeitschrift für Psychologie/Journal of Psychology, 218*, 1–3.

Steffens, M. C., Buchner, A., & Mecklenbräuker, S. (2005). Gender bias in fame judgments: Implicit gender stereotyping or matching study phase fame? *Psychonomic Bulletin & Review, 12*, 495–501.

Steffens, M. C., Kirschbaum, M., & Glados, P. (2008). Avoiding stimulus confounds in Implicit Association Tests by using the concepts as stimuli. *British Journal of Social Psychology, 47*, 217–243.

Steffens, M. C., Schult, J. C., & Ebert, I. D. (2009). Feminization of management leads to backlash against agentic applicants: Lack of social skills, not gender, determines low hireability judgments in student samples. *Psychology Science Quarterly, 51*, 16–46.

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology.*

van Ravenzwaaij, D., van der Maas, H. L. J., & Wagenmakers, E. J. (in press). Does the name–race Implicit Association Test measure racial prejudice? *Experimental Psychology.* doi:10.1027/1618-3169/a000093

Wilde, A., & Diekman, A. B. (2005). Cross-cultural similarities and differences in dynamic stereotypes: A comparison between Germany and the United States. *Psychology of Women Quarterly, 29*, 188–196.