Gender Stereotypes and Expected Backlash for Female STEM Students in Germany and Japan

Laura Froehlich[1]*, Saori Tsukamoto[2], Yasuko Morinaga[3], Kiriko Sakata[3], Yukiko Uchida[4], Melanie M. Keller[5], Stefan Stürmer[1], Sarah E. Martiny[6] and Gisela Trommsdorff[7]

[1]Research Cluster D²L², FernUniversität in Hagen, Hagen, Germany, [2]Division of Liberal Arts and Sciences, Aichi Gakuin University, Nisshin, Japan, [3]Graduate School of Humanities and Social Sciences, University of Hiroshima, Hiroshima, Japan, [4]Kokoro Research Center, Kyoto University, Kyoto, Japan, [5]IPN - Leibniz Institute for Science and Mathematics Education, Kiel, Germany, [6]Department of Psychology, UiT the Arctic University of Norway, Tromsø, Norway, [7]Department of Psychology, University of Konstanz, Konstanz, Germany

Although Germany and Japan are top-ranking in STEM, women are underrepresented in the STEM fields of physics, engineering, and computer science in both countries. The current research investigated widespread gender-science stereotypes in STEM in the two countries (Studies 1 and 2) and negative consequences of expected backlash (i.e., imagining negative reactions and lower ascribed communion in scenarios) for women's emotions and motivation in STEM due to role incongruity and lack-of-fit (Study 3). Studies 1 ($N = 87$) and 2 ($N = 22,556$) showed that explicit and implicit gender-science stereotypes are widespread and comparable in Germany and Japan. Study 3 ($N = 628$) showed that lower ascribed communion was related to less positive emotions, more negative emotions and anxiety emotions, and less study motivation for STEM students (from the fields of physics, engineering, and computer science) from Germany and Japan. Results point to more subtle expected backlash effects for women in STEM than hypothesized. Theoretical and practical implications for gender equality in STEM are discussed.

Keywords: backlash, cross-cultural psychology, gender stereotypes, social role theory, science technology engineering mathematics

INTRODUCTION

Around the world, women are underrepresented in Science, Technology, Engineering, and Mathematics (STEM) fields. Across the member states of the Organization of Economic Co-operation and Development (OECD), 72% of engineering and 80% of information technology degrees are awarded to men (OECD, 2015). However, gender distributions differ between STEM fields. Whereas women's representation in biology, chemistry and mathematics is equal or even higher than men's, women are clearly underrepresented in physics, engineering, and computer science (e.g., Cheryan et al., 2017). Women's underrepresentation in these fields is unlikely to be explained by gender differences in mathematical ability, as numerous studies found that men and women show equal math performance (e.g., Else-Quest et al., 2010; Lindberg et al., 2010). The topic of gender differences in STEM has been investigated in numerous disciplines. Social-psychological research highlights how gender stereotypes and their consequences for women's emotion, motivation, and behavior contribute to their underrepresentation in STEM (e.g., Eagly and Karau, 2002; Eagly and Wood, 2012).

Whereas a large amount of social-psychological work on women's underrepresentation in STEM has focused on the United States (e.g., Cheryan et al., 2017; Diekman et al., 2017), the gender gap in STEM varies around the world. It is of increasing importance to investigate factors that contribute to cross-cultural differences and similarities in women's underrepresentation in STEM (e.g., Yalcinkaya and Adams, 2020). Therefore, the current research focuses on Germany and Japan, two top-ranking countries in STEM, which for example, are among the top 5 countries in natural-science research (Nature Index, 2020) and technological expertise (U. S. News and World Report LP, 2020). Despite their success in STEM, in both countries women are underrepresented in physics, engineering, and computer science. In these fields, less than one third of undergraduate students were female (Germany: physics: 30%, engineering: 24%, computer science: 21%; Japan: science: 27%, engineering: 14%; Destatis, 2019; Gender Equality Bureau Cabinet Office, 2017).

Social psychological research has shown that gender stereotypes associate STEM with males (e.g., Nosek et al., 2009) and that women entering counter-stereotypic fields can experience social repercussions in form of backlash effects (e.g., Rudman and Glick, 2001). The current research investigates how gender stereotypes and expected backlash effects contribute to the gender gap in STEM in Germany and Japan, two different cultural contexts in which group membership is of varying relevance to the individuals (Markus and Kitayama, 1991) and which have received less scholarly attention than the cultural context of the United States.

Gender Stereotypes in STEM

Despite gender similarities in performance, women's STEM abilities and motivation are stereotyped as low in many countries (e.g., Miller et al., 2015; Nosek et al., 2009). Stereotypes are "beliefs and associations that link a whole group of people with certain traits or characteristics" (Kassin et al., 2011, p. 148) and can be described on the dimensions of agency and communion (e.g., Williams and Best, 1991). Agency consists of competence ("capable") and assertiveness ("ambitious"), whereas communion consists of warmth ("friendly") and morality ("honest"; Abele et al., 2016). Men are stereotyped as agentic and women as communal (e.g., Williams and Best, 1991). As STEM is stereotypically associated with traits that are more valued in men than in women (Cheryan et al., 2015), negative stereotypes about women's agency likely have detrimental consequences for women in STEM. They are associated with lower domain identification, career intentions (e.g., Cundiff et al., 2013), interest, sense of belonging (e.g., Cheryan et al., 2009), and lower enrollment in STEM classes (e.g., Stout et al., 2016). Thus, it can be assumed that gender stereotypes contribute to women's underrepresentation in STEM.

Research conducted separately in Germany and Japan showed that women are negatively stereotyped in STEM in both countries (e.g., Adachi, 2014; Ikkatai et al., 2020; Steffens and Jelenec, 2011). The current research conducts a joint investigation of gender stereotypes in these two countries to gain knowledge about potential similarities and differences in gender stereotypes and their psychological consequences for (female) STEM students. Further, we aim to study whether the psychological processes that are related to widespread gender stereotypes and women's underrepresentation in STEM are generalizable in these two countries representing different world regions: Whereas Germany can be categorized as a WEIRD (i.e., Western, Educated, Industrialized, Rich, Democratic), the East Asian country of Japan, although rich and industrialized, is commonly classified as non-Western (Henrich et al., 2010).

A recent model of cross-national variation in gender gaps in STEM participation (Yalcinkaya and Adams, 2020) proposed that individualistic, post-materialistic WEIRD countries show higher underrepresentation of women in STEM than collectivistic, materialistic countries. The model explains national differences in STEM gender gaps by differences in values emphasizing individual choice vs. financial security and relational expectations. However, there are deviations from this proposed dichotomy, embodied by Germany and Japan. Germany is more individualistic than Japan (e.g., Varnum et al., 2010), but the structural and economic factors are similar: Both countries are industrialized and affluent (e.g., Credit Suisse Research Institute, 2019), and in both countries the gender gap in STEM is large. As gender stereotypes arise from the gendered division of labor (e.g., Eagly and Wood, 2012) and women are underrepresented in STEM in Germany and Japan, we thus expect STEM ability to be stereotypically associated with men rather than women in both countries (Hypothesis 1; Studies 1 and 2). Focusing on Germany and Japan, we aim at investigating which aspects of the consequences of gender stereotypes for women in STEM are generalizable across countries and whether they are related to cultural variables reflecting the relevance of social group membership and associated stereotypes for the self (Study 3).

Backlash and Lack of Fit for Women in STEM

Social-psychological theories describe negative consequences of gender stereotypes for women in male-dominated domains (e.g., leadership, STEM). Social role theory (e.g., Eagly and Wood, 2012) posits that gender stereotypes arise because men and women occupy different social roles. The observation of gender-segregated social roles leads to stereotypes, which subsequently influence motivation, emotion, and behavior. Higher role segregation and stronger stereotypes lead to gender differences in behavior. Women (men) are expected to behave communal (agentic). However, women pursuing a STEM career behave counter-stereotypically, which can lead to negative social consequences like being perceived as unlikable.

The lack-of-fit framework describes that social roles stereotyped to require agentic traits (e.g., leadership positions) are perceived as incongruent with the female stereotype, resulting in a perceived lack-of-fit of women with these roles (e.g., Heilman, 1983). According to role congruity theory (Eagly and Karau, 2002) men's roles, but not women's, overlap with leadership roles. When women enter a field stereotyped as agentic or display agentic behavior–thereby violating

prescriptive gender stereotypes (i.e., how women should behave; e.g., Eagly and Karau, 2002), they likely experience a backlash effect (i.e., social repercussions for counter-stereotypical behavior). Agentic women receive negative social reactions in that they are evaluated as socially deficient and unlikable (low in communion) by others (Rudman and Phelan, 2008).

Based on social role theory and role congruity theory, we investigate how women in STEM expect backlash as a consequence of gender stereotypes. STEM fields, especially physics, engineering and computer science, are stereotypically associated with men (e.g., Cheryan et al., 2015) and work in STEM fields is not perceived as people-oriented (e.g., Gino et al., 2015), representing communal work goals (Cheryan et al., 2017). Therefore, we expect women in these STEM fields to expect backlash (Hypothesis 2; Study 3). The current research focuses on expected rather than experienced backlash for several reasons. First, investigating actual backlash behavior (repercussions for counter-stereotypical behavior from other people) would require an observational or experimental methodology, which was beyond the scope of the survey conducted in Study 3. Second, the focus of the current research was on how female students expect backlash due to their study major and how this subjective perception of potential backlash influences their subsequent emotions and motivation. We believe that this focus on subjective expectations of backlash is highly relevant, as these subjective expectations are likely to be a proximal predictor of emotions and motivation.

Negative social reactions can in turn influence women's emotions and motivation in STEM. Morinaga et al. (2017) investigated how benevolent sexism affects women's emotions and motivation in mathematics in two scenario experiments with Japanese female (junior) high-school students. When students imagined their math teacher to comment a good performance with "well done, although you are a girl!" (stereotype activation condition), they experienced more negative and less positive emotions than in a control condition ("well done!"). Stereotype activation lead to lower motivation mediated by emotions. In line with this, we expect that for women, but not for men, the expectation of more negative reactions to studying a STEM subject (expected backlash effect) is related to negative emotions (Hypothesis 3). In turn, these emotions predict lower motivation to study for STEM (Hypothesis 4).

A Cross-Cultural Approach to Expected Backlash for Female STEM Students

If STEM is stereotypically associated with men in both Germany and Japan, it is likely that these stereotypes have negative psychological consequences for female STEM students in both cultural contexts. The negative consequences of backlash effects have been predominantly investigated in the United States (e.g., Rudman and Glick, 2001; Rudman and Phelan, 2008; Eaton et al., 2020). It remains unclear whether the expected negative reactions for counter-stereotypical behavior are related to women's emotion and motivation in a similar way and intensity in cultural contexts in which membership in social groups is of varying relevance to the self. To fill this gap in the literature, the present research investigates gender stereotypes and their psychological consequences for women in stereotype-incongruent STEM fields in Germany and Japan and examines whether the psychological variable of self-construal, which reflects how central social group membership is for the self, is associated with the extent of expected backlash effects.

In the Japanese culture individuals tend to endorse an interdependent self-construal, a cultural orientation for which social group membership is central to the self. In the German culture individuals tend to endorse an independent self-construal, for which group membership is less central (Markus and Kitayama, 1991; Varnum et al., 2010). In cultures where individuals tend to endorse an interdependent self-construal, social networks are relatively stable (i.e., low relational mobility; Thomson et al., 2018) and people are highly sensitive to social rejection (e.g., Sato et al., 2014). We thus argue that self-construal is relevant when investigating expected backlash effects of female STEM students across cultures, as individuals endorsing an interdependent self-construal should be more prone to expecting negative social repercussions for their counter-stereotypical behavior (studying a STEM subject) than individuals endorsing an independent self-construal. We thus expect that the kind of self-construal moderates the effects of expected backlash on female STEM students' emotions and motivation (Hypothesis 5). We explore whether these relationships depend on the relational mobility afforded by the social situation. Associations between variables should be stronger in a low relational mobility situation (new relationships are likely to become stable) compared to a high relational mobility situation (relationships are flexible and formed by personal choice).

The Present Research

As a basis for the investigation of the consequences of gender stereotypes for female STEM students in Germany and Japan, in a first step (Studies 1 and 2) we aim at substantiating that in both countries gender stereotypes associate men more with STEM than women. Study 1 investigates explicit gender stereotypes about mathematical and general academic abilities. Because explicit measurement of stereotypes can be prone to response biases (e.g., Smith, 2014; Kemmelmeier, 2016), Study 2 investigates explicit and implicit gender-science stereotypes using samples from Project Implicit. In a second step, we investigate expected backlash effects for female students of physics, engineering and computer science for the first time jointly in Germany and Japan. Study 3 (pre-registered) investigates the consequences of gender stereotypes for German and Japanese STEM students. In two scenarios, participants were asked to imagine a conversation with a previously unknown person of the opposite gender who is asking about their field of study. The participants indicated how they expected their conversation partner to react and perceive them on communion. We hypothesize that women expect more negative reactions and lower communion ratings than men (expected backlash). Furthermore, expected backlash should have negative consequences for women's emotions and motivation in STEM and should be stronger for individuals

strongly endorsing an interdependent self-construal. Materials (Studies 1 and 3), data and analysis scripts (all studies), and the pre-registration (Study 3) are available on the OSF (https://osf.io/4awqe/).

STUDY 1: EXPLICIT GENDER-MATH STEREOTYPES

To replicate the basic premise that men are more strongly associated with STEM and high STEM ability than women (e.g., Steffens and Jelenec, 2011; Ikkatai et al., 2019) in both countries, in a questionnaire study we assessed participants' perceptions of widespread gender stereotypes about math and general academic abilities.

Methods

Data were collected in December 2013 (Japan) and September 2015 (Germany). University students were recruited as participants via e-mail, a virtual laboratory and in class. Participants did not receive compensation for participation. The sample consisted of 28 Japanese (age: $M = 26.15$ years, $SD = 7.34$, 42.9% female) and 59 German university students (age: $M = 33.25$ years, $SD = 10.18$, 74.6% female). Participants answered a questionnaire assessing gender stereotypes about math and general academic abilities and their valence. Materials were translated and back-translated by the research team. Participants listed stereotypical statements about women's and men's general academic and math abilities and rated the statements' valence (from -3 = very negative to $+3$ = very positive). Participants were asked not to provide their personal opinion, but indicate socially shared stereotypes in Germany or Japan. Finally, they provided demographic information (age, gender, nationality) and were debriefed.

Results
Stereotype Content

Japanese participants made 221 statements (women/math: 61, women/general: 55, men/math: 55, men/general: 50), and German participants made 924 statements (women/math: 218, women/general: 239, men/math: 228, men/general: 239). In both samples, most statements about women's math ability indicated a negative conception. For example, participants indicated "slow in doing mental arithmetic," "bad at logical thinking/algebra." In contrast, for women's general academic ability, participants mostly indicated that they are good at languages and humanities, for example, "good at languages" or "good at arts and music." Men's math ability was described with positive statements, e.g., "good at math/logical thinking," "good comprehension of mathematical formulas." In turn, men's general academic ability was characterized as "good at math and natural science" or "bad at languages." The statements reflected the widespread stereotype that women have high abilities in languages and humanities but low abilities in math and science, and vice versa for men (e.g., Steffens and Jelenec, 2011).

Stereotype Valence

Valence ratings were averaged for each category. Ratings were nested within participants, we therefore computed a linear mixed model. To do so, we transformed the data from wide format (1

row per participant) to wide format (4 rows per participant, reflecting repeated measures of Domain and Gender). Because many participants listed less than the maximum number of five statements per category, we used restricted maximum likelihood (REML) estimation as it can produce unbiased estimates of variance and covariance parameters in the presence of missing data and uses the full data set; in contrast to full maximum likelihood estimation with listwise deletion. The dependent variable was valence ratings, predictors were Gender (male vs. female, within participants), Domain (general academic vs. math, within participants), and Country (Germany vs. Japan, between participants). Main and interaction effects were entered as fixed effects, the covariance type was compound symmetry. The main effect of Country was non-significant, $F(1, 87.17) = 0.20$, $p = 0.657$. There were significant main effects of Gender, $F(1, 231.95) = 54.51$, $p < 0.001$, and Domain, $F(1, 253.04) = 5.49$, $p = 0.020$. The interaction of Domain and Gender was also significant, $F(1, 231.95) = 29.59$, $p < 0.001$. The interactions with country were non-significant, $Fs < 0.87$, $ps > 0.351$. Bonferroni-adjusted post-hoc comparisons for the interaction of Gender and Domain across countries revealed that women's math ability was rated significantly more negatively than men's [$M_{women} = -0.86$, 95% CI $(-1.16; -0.55)$, $SE = 0.15$, $M_{men} = 1.04$ $(0.73; 1.34)$, $SE = 0.15$; $t(223.63) = 9.37$, $SE = 0.20$, $p < 0.001$, Cohen's $d = 1.58$]. Valence of women and men's general academic abilities did not differ significantly [$M_{women} = 0.30$, 95% CI $(-0.01; 0.60)$, $SE = 0.16$; $M_{men} = 0.58$ $(0.26; 0.91)$, $SE = 0.16$; $t(239.34) = 1.33$, $SE = 0.22$, $p = 0.183$, $d = 0.23$]. Across countries women's math ability was rated more negatively than their general academic ability [$t(238.86) = 5.59$, $SE = 0.21$, $p < 0.001$, $d = 0.90$]. Men's math ability was rated more positively than their general academic ability [$t(246.25) = 2.12$, $SE = 0.21$, $p = 0.035$, $d = 0.41$].

Discussion

In line with previous studies conducted separately in Germany and Japan (Ikkatai et al., 2019; Steffens and Jelenec, 2011), findings indicate the presence of negative stereotypes about women's math ability in Japanese and German society (Hypothesis 1). Participants indicated that women's math ability is stereotyped more negatively than men's, and also more negatively than women's general academic ability. These effects can be considered large (Cohen, 1988). There were no country differences between stereotype content and valence ratings. However, samples were small and stereotypes were measured only explicitly. To rule out response bias in explicit stereotype measurement (e.g., Smith, 2014; Kemmelmeier, 2016), in Study 2, we investigated gender-science stereotypes with data from Project Implicit.

STUDY 2: EXPLICIT AND IMPLICIT GENDER-SCIENCE STEREOTYPES

Study 2 investigated explicit and implicit gender-science stereotypes in Germany and Japan by Project Implicit (https://implicit.harvard.edu), which provides different Implicit Association Tests (IAT; Greenwald et al., 1998) to the public

in various languages. The gender-science IAT is a behavioral task measuring the implicit association between the categories male/female and science/liberal arts. Participants from 34 countries who completed gender-science IATs on the Project Implicit website associated male with science and female with liberal arts more easily than the reverse category combination (Nosek et al., 2009).

Method

Data provided by Project Implicit contained responses from 72,094 participants. Participants with missing values on the measure of implicit gender-science association (n = 44,010), missing values on gender (n = 4,017), or an age below 18 years (n = 1,159) were excluded. The final sample (N = 22,556) consisted of 9,875 Japanese (age: 18–88 years, M = 28.46, SD = 10.23; 50% female) and 12,681 German participants (age: 18–87 years, M = 29.54, SD = 10.22, 54% female).

Participants completed the gender-science IAT between 2006 and 2017. They categorized words into four categories by pressing two keys. In a stereotype-congruent condition the categories male/science were paired on one key and female/liberal arts on the other; in the stereotype-incongruent condition the pairings were reversed. Faster responses in the stereotype-congruent condition compared to the stereotype-incongruent condition indicate a stronger male-science association. Details on Project Implicit's gender-science IAT procedure can be found in Nosek et al. (2009). In addition, participants responded to the item "How much do you associate science with males or females" (1 = strongly male to 7 = strongly female) as a measure of explicit gender-science stereotypes, and provided demographics.

Statistical Analyses

In contrast to Study 1, which included a mixed model with between- and within-participants factors, Study 2 predicted implicit and explicit stereotypes by the between-participants factors Gender and Country. To do so, we used factorial Analysis of Variance (ANOVA). As only participants who completed the IAT were included in the sample, there were no missing values in the analysis of implicit stereotypes. For explicit stereotypes, a subsample of 51% of participants who completed the IAT also completed the explicit stereotype measure. Again, analyses were conducted with the subsample that completed the respective measure. In additional ANCOVAs, we controlled for year of data collection.

Results

Implicit Stereotypes

Project Implicit computed D scores as a measure of the implicit gender-science association for each participant by dividing the difference in mean response latency between the two conditions by the participant's latency standard deviation inclusive of the two conditions using the improved scoring algorithm (Nosek et al., 2009). Participants from Germany as well as from Japan showed positive overall D scores, indicating a stronger implicit association of male/science and female/liberal arts than the reverse combination [$M_{Germany}$ = 0.43, 95% CI (0.42; 0.44), SE

= 0.01, M_{Japan} = 0.38 (0.37; 0.39), SE = 0.01]. We subjected the D scores to a 2 × 2 ANOVA with the between-participants factors Gender (men vs. women) and Country (Germany vs. Japan). Results showed significant main effects of Gender, F (1, 22,552) = 493.27, $p < 0.001$, Country, F (1, 22,552) = 70.49, $p < 0.001$, and a significant interaction, F (1, 22,552) = 145.89, $p < 0.001$. Bonferroni-adjusted post-hoc comparisons showed that in both countries, women showed stronger implicit associations of male/science and female/liberal arts than men [Germany: M_{Women} = 0.45 (0.44; 0.46), SE = 0.01, M_{Men} = 0.40 (0.39; 0.41), SE = 0.01), t (22,552) = 7.64, SE = 0.01, $p < 0.001$, d = 0.14; Japan: M_{Women} = 0.47 (0.46; 0.48), SE = 0.01, M_{Men} = 0.29 (0.27; 0.30), SE = 0.01, t (22,552) = 22.84, SE = 0.01, $p < 0.001$, d = 0.46]. For men, German participants showed a stronger implicit association than Japanese [t (22,552) = 14.25, SE = 0.01, $p < 0.001$, d = 0.14]. This difference was also significant for women, but with a small effect size [t (22,552) = 2.60, SE = 0.01, $p < 0.001$, d = 0.05]. Results were mainly robust when controlling for year of data collection (albeit the last comparison was no longer significant).

Explicit Stereotypes

The explicit gender-science stereotype item was completed by 11,601 participants (51% of the total sample). Means were above the scale midpoint, indicating that science was stereotyped to be male. A 2 × 2 ANOVA with Gender and Country as between-participants factors and explicit stereotypes as the dependent variable showed significant main effects of Gender, F (1, 11,597) = 89.02, $p < 0.001$, Country, F (1, 11,597) = 276.59, $p < 0.001$, and a significant interaction, F (1, 11,597) = 74.72, $p < 0.001$. Bonferroni-adjusted post-hoc comparisons showed that in Germany, men showed stronger endorsement of explicit stereotypes than women, this difference was non-significant in Japan [Germany: M_{Women} = 4.88 (4.85; 4.91), SE = 0.02, M_{Men} = 5.22 (5.18; 5.26), SE = 0.02, t (11,597) = 11.86, SE = 0.02, $p < 0.001$, d = 0.37; Japan: M_{Women} = 5.36 (5.32; 5.39), SE = 0.02, M_{Men} = 5.37 (5.33; 5.41), SE = 0.02, t (11,597) = 0.61, SE = 0.02, p = 0.541, d = 0.01]. Both men and women from Japan showed stronger stereotype endorsement than men and women from Germany [men: t (11,597) = 5.33, SE = 0.03, $p < 0.001$, d = 0.15, women: t (11,597) = 19.09, SE = 0.03, $p < 0.001$, d = 0.50]. Results were robust when year of data collection was controlled.

Discussion

In line with Study 1, Study 2 supported Hypothesis 1, showing that in large samples and with implicit and explicit stereotype measures, men were more strongly associated with science than women. It is prudent to note that significant country and gender differences should be interpreted with caution due to large sample sizes, effect sizes for country and gender differences were small to medium (0.01 < Cohen's d < 0.50). Study 2 replicated and extended findings from Study 1 and previous research (Ikkatai et al., 2019; Steffens and Jelenec, 2011), as it included much larger samples and explicit as well as implicit measures of gender-science stereotypes, whereas Study 1 focused on gender-math stereotypes. Taken together, Studies 1 and 2 take multi-faceted angles and present a comprehensive picture of gender stereotypes in the STEM domain. Based on the combined results, we

conclude that negative gender stereotypes about women's STEM ability are widespread in both countries. Study 3 thus focused on the consequences of these stereotypes and investigated to what extent female STEM students expect backlash for their stereotype-incongruent study major.

STUDY 3: EXPECTED BACKLASH FOR FEMALE STEM STUDENTS

Study 3 was a scenario study with German and Japanese university students of physics, engineering, and computer science as participants. In an online questionnaire, participants imagined being asked about their study major in a conversation with an unknown person of the opposite gender. They completed items on the expected reactions of the conversation partner, their emotions and study motivation. We hypothesized expected backlash (i.e., expected negative reactions of the conversation partner and lower ascribed communion) for women, but not for men (Hypothesis 2). This expected backlash should predict more negative/less positive emotions and lower study motivation (Hypotheses 3 and 4). Moreover, we expected these relationships to be stronger for women endorsing an interdependent self-construal (Hypothesis 5). Hypotheses were pre-registered (https://osf.io/afqxb/).

Participants and Procedure

Data were collected between January and September 2019. After registering their e-mail address in an online form, participants were invited to participate in two parts of an online questionnaire *via* personalized emails. Data from Part 1 and 2 (2-days interval between measurements) were matched with participant-generated codes. E-mail addresses could not be connected to questionnaire data. Participants provided written consent in accordance with EU General Data Protection Law. The study was approved by the ethics committee of the first author's institution.

Participants were recruited via university classes and Facebook groups/mailing lists of student associations of physics, mathematics, computer science, and engineering. The questionnaire (both parts) was completed by 656 participants. We excluded participants who were not university students or indicated non-STEM majors ($n = 24$), entered non-corresponding gender information at the two parts ($n = 2$), or indicated "other" as their gender ($n = 2$). The final sample consisted of 628 participants (Japanese: $n = 432$, 101 female, age: 18–33 years, $M = 19.73$, $SD = 1.59$; German: $n = 196$, 87 female, age: 18–57 years, $M = 26.88$, $SD = 8.37$).

A sample size of 100 female students per country was determined based on an a-priori power analysis for a repeated-measures ANOVA (Hypothesis 2) with a within-between interaction (medium effect size of $f = 0.15$, $\alpha = 0.05$, power = 0.80, 2 groups, 2 measurements), which resulted in a sample size of $N = 90$. As Hypotheses 3–5 required path modeling, sample size was increased to 100 female students per country (and at least as many male students), resulting in

a total minimum sample size of $N = 400$. The pre-registered sample size of female students was reached for the Japanese but not the German sample ($n = 87$). Data collection was terminated after 9 months of contacting Facebook groups and student councils of the STEM majors of all German universities, and 152 German university instructors. A sensitivity analysis showed that with the current sample small effects ($f = 0.06$) could be detected.

Materials

Materials were translated by the project team and back-translated by a professional translator. Moderators and demographics were assessed in Part 1, scenarios and outcomes in Part 2.

Part 1

Participants indicated whether they were university students, their field of study and gender. Independent/interdependent self-construal was measured with 10 items each (e.g., "I always try to have my own opinion," "I will sacrifice my self-interest for the benefit of the group I am in," 1 = do not agree, 7 = completely agree; Park and Kitayama, 2014).

Part 2

Participants were asked to imagine a conversation with an unknown person of the opposite gender in two scenarios. Female participants imagined a male conversation partner, whereas male participants imagined a female conversation partner. The wedding party scenario should represent high relational mobility (a flexible social network and opportunities to form relationships by choice), whereas the choir scenario should represent low relational mobility (a fixed network and long-term relationships due to circumstance; Thomson et al., 2018).

Wedding Party Scenario
"Please imagine you are attending a friend's wedding reception. You are introduced to a male/female person whom you have not met before. You start chatting with him/her and you feel like you are getting along well. During your conversation, he/she asks you about your university major. You tell him/her that you study (subject entered by participant displayed). Please take some time to imagine yourself in this situation."

Choir Scenario
"Please imagine that you recently decided to participate in your university's choir. Therefore, you attend the first choir meeting of the new semester. You are very motivated to join the choir and go to rehearsals regularly because you like singing and want to start a new extra-curricular activity for the next year. During the first meeting, a choir member asks you about your university major. You tell him/her that you study (subject). Please take some time to imagine yourself in this situation."

Following each scenario, participants described how they imagined the conversation partner's reaction ["How do you think would your conversation partner react to hearing that you study (subject)? Please write down his/her imagined reaction as detailed as possible. Keep in mind that reactions can either be verbal (what he/she says) or non-verbal (facial

TABLE 1 | Descriptive statistics.

		Reaction	Communion	Positive emotions	Negative emotions	Anxiety emotions	Motivation	Independent self-construal
Japan								
Men	M	4.42	3.97	3.71	2.27	3.32	4.43	4.41
	[95% CI]	[4.31; 4.52]	[3.84; 4.08]	[3.58; 3.83]	[2.15; 2.39]	[3.19; 3.45]	[4.32; 4.53]	[4.32; 4.50]
(n = 331)	α	0.74	0.89	0.92	0.82	0.76	0.86	0.68
Women	M	4.48	3.41	3.64	2.09	2.97	4.29	4.43
	[95% CI]	[4.30; 4.68]	[3.20; 3.60]	[3.44; 3.83]	[1.87; 2.33]	[2.71; 3.23]	[4.11; 4.47]	[4.28; 4.58]
(n = 101)	α	0.85	0.88	0.88	0.87	0.81	0.85	0.68
Germany								
Men	M	4.69	4.03	4.24	1.85	2.18	4.22	4.74
	[95% CI]	[4.53; 4.85]	[3.78; 4.28]	[4.01; 4.46]	[1.67; 2.04]	[1.96; 2.41]	[4.12; 4.34]	[4.60; 4.86]
(n = 109)	α	0.72	0.92	0.92	0.89	0.80	0.76	0.63
Women	M	4.87	4.24	4.67	2.13	2.14	4.49	4.89
	[95% CI]	[4.68; 5.07]	[3.97; 4.50]	[4.47; 4.89]	[1.91; 2.38]	[1.91; 2.38]	[4.29; 4.70]	[4.74; 5.04]
(n = 87)	α	0.59	0.94	0.89	0.92	0.83	0.92	0.67
Total								
Men	M	4.49	3.98	3.84	2.17	3.04	4.38	4.49
	[95% CI]	[4.40; 4.58]	[3.88; 4.09]	[3.72; 3.95]	[2.06; 2.27]	[2.92; 3.16]	[4.29; 4.46]	[4.42; 4.56]
(n = 440)	α	0.74	0.90	0.92	0.88	0.80	0.84	0.67
Women	M	4.66	3.79	4.12	2.11	2.58	4.38	4.64
	[95% CI]	[4.53; 4.81]	[3.62; 3.97]	[3.95; 4.29]	[1.95; 2.28]	[2.40; 2.77]	[4.26; 4.51]	[4.53; 4.75]
(n = 188)	α	0.71	0.92	0.91	0.90	0.83	0.88	0.68

Note: For scales with more than two items, Chronbach's α is displayed, for reaction and motivation Spearman's rho is displayed.

expression, body language etc.).”], rated the reaction valence (“How positive or negative is this reaction?” 1 = very negative, 7 = very positive) and impression (“How positive or negative do you think is your conversation partner's impression of you?” 1 = very negative, 7 = very positive).

Furthermore, they rated expected communion (“Please indicate how much your conversation partner thinks you possess the following traits,” 4 items, gentle, affectionate, supportive, sympathetic; Steinmetz et al., 2014; 1 = not at all, 7 = completely), emotions (“How would you feel in the scenario?” 13 items; Morinaga et al., 2017, 1 = do not agree, 7 = completely agree), and motivation [“In the scenario, how would you intend to work hard for (subject) from now on? Please indicate whether your motivation is stronger or weaker compared to before.” 1 = much weaker than before, 7 = much stronger than before; and “In the scenario, how has your motivation to study hard for (subject indicated above) changed?” 1 = completely lost motivation; 7 = motivation got much stronger; Morinaga et al., 2017]. Demographics included field of study, gender, birth year, and nationality. Further measures not reported in this paper were implicit theories of intelligence, gender identity, implicit gender-science attitudes, and benevolent sexism (Part 1), perceived agency, general motivation, career and research intentions, goals, perceived stereotype threat, future work domain and importance of digitalization for STEM (Part 2).

Statistical Analyses

Because all questions were programmed as mandatory in the online questionnaire, there was no missing data. Measurement invariance was tested with exploratory factor analysis (conducted in SPSS version 25) and confirmatory factor analysis (conducted in Mplus Version 8.6). Cutoff criteria for goodness of model fit in CFA were CFI/TLI ≥0.90, SRMR ≤0.06, RMSEA ≤0.08. Reaction valence and communion stereotypes were investigated with linear mixed models with REML estimation. Open-ended answers on reactions were categorized and subjected to frequency analysis (cross tabulation and χ^2 tests). Consequences of reactions and communion for emotions and motivation were investigated with path analysis in Mplus.

Results

Measurement Invariance and Descriptive Statistics

We investigated measurement invariance between national subsamples for multi-item measures. Multiple-group confirmatory factor analysis (CFA) with the national groups after model modifications showed partial metric invariance for all scales. For emotions, exploratory factor analysis (EFA) with promax rotation yielded three factors: positive (happy, proud, feeling good, satisfied, relieved, relaxed), negative (disappointed, angry, feeling bad, dissatisfied), and anxiety (anxious, nervous, embarrassed). In a CFA configural model (no equality constraints), three items (satisfied, relieved, feeling bad) were excluded due to low factor loadings and high cross-loadings. Loadings of item “angry” on negative emotions and item “embarrassed” on anxiety emotions were freed due to non-equivalence. The model showing partial metric invariance (i.e., factor structure and at least two loadings per factor constrained to be equal across groups) had acceptable model

TABLE 2 | Bivariate correlations [r, (95% CI)].

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Germany							
(1) Reaction	—	0.40***	0.69***	−0.46***	−0.35***	0.37***	0.09
		[0.23; 0.55]	[0.57; 0.78]	[−0.61; −0.31]	[−0.52; −0.14]	[0.19; 0.52]	[−0.09; 0.26]
(2) Communion	0.36**	—	0.51***	−0.13	−0.00	0.19	0.11
	[0.12; 0.53]		[0.32; 0.67]	[−0.29; 0.02]	[−0.20; 0.18]	[0.02; 0.35]	[−0.08; 0.29]
(3) Positive emotions	0.67***	0.35**	—	−0.25**	−0.23*	0.36***	0.14
	[0.49; 0.78]	[0.14; 0.52]		[−0.46; −0.05]	[−0.42; 0.00]	[0.21; 0.48]	[−0.06; 0.33]
(4) Negative emotions	−0.48***	−0.23*	−0.48***	—	0.60***	−0.22*	0.04
	[−0.67; −0.22]	[−0.44; 0.00]	[−0.64; −0.27]		[0.44; 0.73]	[−0.42; −0.02]	[−0.18; 0.23]
(5) Anxiety emotions	−0.24*	0.04	−0.45***	0.58***	—	−0.15	−0.10
	[−0.42; −0.05]	[−0.16; 0.22]	[−0.57; −0.30]	[0.41; 0.73]		[−0.35; 0.06]	[−0.27; 0.08]
(6) Motivation	0.14	0.01	0.26*	0.01*	−0.15	—	0.00
	[−0.19; 0.45]	[−0.26; 0.31]	[−0.06; 0.52]	[−0.24; 0.29]	[−0.30; 0.03]		[−0.20; 0.21]
(7) Independent self-construal	0.09	0.08	0.24*	−0.09	−0.37***	0.31**	—
	[−0.13; 0.31]	[−0.18; 0.32]	[−0.01; 0.46]	[−33; 0.15]	[−0.55; −0.17]	[0.08; 0.50]	
Japan							
(1) Reaction	—	0.45***	0.61***	−0.33***	−0.11	0.60***	−0.01
		[0.33; 0.55]	[0.51; 0.69]	[−0.44; −0.22]	[−0.23; 0.02]	[0.50; 0.69]	[−0.12; 0.08]
(2) Communion	0.23*	—	0.63***	0.05	0.20***	0.42***	−0.07
	[0.02; 0.42]		[0.53; 0.71]	[−0.06; 0.17]	[0.07; 0.33]	[0.30; 0.53]	[−0.17; 0.04]
(3) Positive emotions	0.51***	0.61***	—	−0.01	0.14*	0.48***	−0.07
	[0.33; 0.66]	[0.47; 0.74]		[−0.11; 0.12]	[0.01; 0.28]	[0.35; 0.59]	[−0.18; 0.04]
(4) Negative emotions	−0.61***	−0.01	−0.27**	—	0.51***	−0.25***	−0.11
	[−0.73; −0.44]	[−0.21; 0.26]	[−0.47; 0.03]		[0.43; 0.58]	[−0.36; −0.13]	[−0.22; 0.03]
(5) Anxiety emotions	−0.25*	0.20*	0.15	0.51***	—	−0.01	−0.24***
	[−0.41; −0.05]	[−0.03; 0.41]	[−0.08; 0.36]	[0.38; 0.64]		[−0.14; 0.11]	[−0.35; −0.13]
(6) Motivation	0.64***	0.36***	0.58***	−0.48***	−0.05	—	−0.05
	[0.45; 0.76]	[0.17; 0.52]	[0.44; 0.70]	[−0.66; −0.24]	[−0.24; 0.16]		[−0.06; 0.17]
(7) Independent self-construal	−0.06	0.01	0.02	−0.09	−0.20	0.09	—
	[−0.24; 0.13]	[−0.23; 0.21]	[−0.21; 0.23]	[−0.34; 0.13]	[−0.39; 0.01]	[−0.12; 0.30]	

Notes: Correlations for women (men) below (above) the diagonal.
*p < 0.05, **p < 0.01, ***p < 0.001.

fit [χ^2 (72) = 218.58, $p < 0.001$, RMSEA = 0.08, CFI = 0.96, TLI = 0.95, SRMR = 0.06] which did not significantly differ from the configural model [$\Delta \chi^2$ (8) = 9.31, $p = 0.317$].

EFA and CFA showed that the four communion items loaded on a single factor in both samples. A model showing partial metric invariance showed good fit when error terms of the items "supportive" and "sympathetic" were allowed to correlate, and the loading of the item "compassionate" was freed due to non-invariance, [χ^2 (3) = 3.82, $p = 0.282$, RMSEA = 0.03, CFI = 1.00, TLI = 1.00, SRMR = 0.07]. The fit of this model did not significantly differ from the configural model [$\Delta \chi^2$ (1) = 0.73, $p = 0.393$].

Internal consistency of the interdependent self-construal scale was not acceptable (0.55 < Cronbach's α < 0.72) with German men showing a value below 0.60, therefore, we used the independent self-construal scale, which had acceptable consistency (0.63 < α < 0.68). Issues with low reliability of this and other self-construal scales have been reported and discussed in earlier research (e.g., Gudykunst and Lee, 2003; Park and Kitayama, 2014). Multi-group CFA showed partial metric invariance in a single-factor model when error terms of "I always try to have my own opinions" and "I always express my opinions clearly" were allowed to correlate, and the loadings of the items "I always try to have my own opinions" and "It does

not concern me when my opinions or behavior differs from that of other people" were freed due to non-invariance [χ^2 (74) = 136.08, $p < 0.001$, RMSEA = 0.05, CFI = 0.92, TLI = 0.90, SRMR = 0.05]. The fit of this model did not significantly differ from the configural model [$\Delta \chi^2$ (6) = 11.87, $p = 0.065$]. Independent and interdependent self-construal were negatively correlated in all subgroups ($-0.12 < r < -0.54$). Descriptive statistics are displayed in **Table 1**, bivariate correlations in **Table 2**. Outputs of CFAs to investigate measurement invariance can be found on the OSF.

Reaction Valence and Communion Stereotypes

To test Hypothesis 2, data were transformed into long format (1,256 observations, 628 participants) due to the repeated measurements for the scenarios (in long format, one data row represented one observation instead of one participant). We computed linear mixed models with Gender (male vs. female, between-participants) and Scenario (high vs. low relational mobility, within-participants) as factors and valence of imagined reactions and communion as dependent variables. For reaction valence, there was a main effect of Gender, F (1, 1,251.97) = 6.22, $p = 0.013$. Women expected more positive reactions [$M = 4.66$ (4.55; 4.78), $SE = 0.04$] than men [$M = 4.49$ (4.41; 4.56), $SE = 0.06$), t (1,251.97) = 2.50, $SE = 0.07$, $p = 0.013$, $d = 0.15$].

TABLE 3 | Frequencies of categories of reactions to scenarios.

	Positive	Negative	Surprised	Neutral	Total
High relational mobility					
male	191 (43%)	69 (16%)	94 (21%)	85 (20%)	439 (100%)
female	39 (21%)	10 (5%)	123 (65%)	16 (9%)	188 (100%)
Low relational mobility					
male	159 (36%)	72 (17%)	80 (18%)	128 (29%)	439 (100%)
female	45 (24%)	16 (9%)	80 (43%)	47 (25%)	188 (100%)

The main effect of Scenario was also significant, F (1, 1,251.97) = 15.32, $p < 0.001$. Valence of reactions was more positive in the high relational mobility scenario [M = 4.71 (4.62; 4.81), SE = 0.05] compared to the low relational mobility scenario [M = 4.44 (4.34; 4.53), SE = 0.05, t (1,251.97) = 3.92, SE = 0.07, $p < 0.001$, d = 0.22]. The interaction was non-significant [t (1,251.97) = 0.81, p = 0.368].

For communion, there was a main effect of Gender, F (1, 1,247.91) = 6.36, p = 0.012. Women expected lower communion ratings [M = 3.79 (3.67; 3.92), SE = 0.06] than men [M = 3.98 (3.90; 4.07), SE = 0.04, t (1,247.91) = 2.53, SE = 0.08, p = 0.012, d = 0.22]. The main effect of Scenario and the interaction were non-significant (Fs < 1.46, ps > 0.227).

In an exploratory analysis, we categorized the open-ended responses on imagined reactions into four categories: positive, negative, surprised, and interested/neutral. Reactions were coded as positive when containing positive aspects (e.g., impressed, admiring, interested) and negative when containing negative aspects (e.g., rejection, disinterest, distancing, negative comments about STEM). Reactions were coded as surprised when cues for surprise were mentioned (e.g., surprised, amazed, perplexed). When surprise was mentioned in combination with other aspects, responses were coded as surprised. Reactions were coded as neutral when they did not contain positive, negative or surprised aspects (e.g., no apparent reaction, neutral) or if a combination of positive and negative reactions was mentioned (e.g., "it could be one of two possibilities, a positive reaction or a negative one"). Reactions were coded by two independent raters each who were fluent in the respective language. Interrater reliabilities were excellent (Cohen's Kappa ≥ 0.89), indicating high agreement between raters. There were gender differences in distributions across categories: men more frequently described positive expected reactions, whereas women more frequently described surprised reactions [**Table 3**, high relational mobility: χ^2 (3) = 113.19, $p < 0.001$; low relational mobility: χ^2 (3) = 43.29, $p < 0.001$]. Further descriptive analyses showed that none of the men, but 34% of women imagined their conversation partner to make a reference to their gender [e.g., "Really? Women are rare in (subject)," "理系女" (Rikejo, female scientist), "you are studying (subject) although you are a woman?"].

Consequences of Perceived Reactions and Communion

To test Hypotheses 3 and 4, we computed path models with multiple-group comparison for men/women including reaction valence and communion ratings as predictors of emotions. In

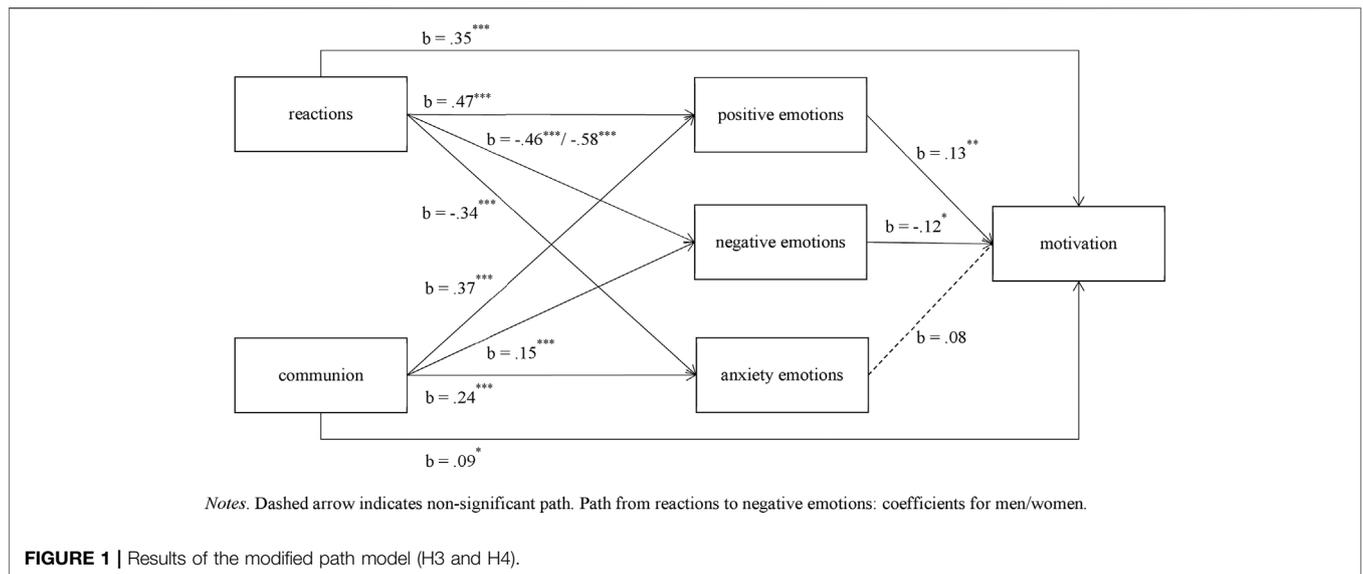**TABLE 4 |** Direct effects in modified path model.

	b	[LLCI; ULCI]	SE	p
Reaction				
country	−0.34	[−0.51; −0.17]	0.09	<0.001
Communion				
country	−0.03	[−0.20; 0.14]	0.09	0.753
Positive emotions				
reaction	0.47	[0.41; 0.53]	0.03	<0.001
communion	0.37	[0.31; 0.43]	0.03	<0.001
country	0.15	[0.03; 0.27]	0.06	0.011
Negative emotions				
reaction (men)	−0.46	[−0.54; −0.38]	0.03	<0.001
reaction (women)	−0.58	[−0.68; −0.47]	0.05	<0.001
communion	0.15	[0.07; 0.22]	0.04	<0.001
country	−0.17	[−0.32; −0.02]	0.08	0.028
Anxiety emotions				
reaction	−0.34	[−0.42; −0.27]	0.04	<0.001
communion	0.24	[0.16; 0.31]	0.04	<0.001
country	−0.14	[−0.29; 0.01]	0.08	0.069
Motivation				
reaction	0.35	[0.26; 0.44]	0.04	<0.001
communion	0.09	[0.01; 0.16]	0.04	0.031
positive emotions	0.13	[0.04; 0.22]	0.05	0.003
negative emotions	−0.12	[−0.22; −0.02]	0.05	0.022
anxiety emotions	0.08	[−0.02; 0.18]	0.05	0.125
country	0.19	[0.06; 0.33]	0.07	0.006

TABLE 5 | Indirect effects in modified path model.

	ab	[LLCI; ULCI]	SE	p
Reaction → motivation				
positive emotions	0.06	[0.02; 0.10]	0.02	0.004
negative emotions (men)	0.06	[0.01; 0.10]	0.02	0.025
negative emotions (women)	0.07	[0.01; 0.13]	0.03	0.025
anxiety emotions	−0.03	[−0.06; 0.01]	0.02	0.130
Communion → motivation				
positive emotions	0.05	[0.02; 0.08]	0.02	0.004
negative emotions	−0.02	[−0.04; 0.00]	0.01	0.049
anxiety emotions	0.02	[−0.01; 0.04]	0.01	0.137

turn, reactions and emotions predicted study motivation. For constructs measured with more than two items (i.e., communion, emotions, self-construal) we used factor scores that were generated under the assumption of partial metric invariance as manifest variables in the model (please note that we deviate from the pre-registration which included latent variable modeling to account for the fact that partial metric invariance but not full scalar invariance could be established). Scenarios were combined, but separate analyses showed similar results with larger effect sizes for the low relational mobility scenario. Results for separate analyses can be found on the OSF. Although the pre-registration stated that we would use scenario as a covariate, we opted for presenting the results for the scenarios separately to better reflect potential differences between scenarios. We controlled for country of data collection in all analyses.

In a first model, all paths were constrained to be equal for men and women. Country of data collection was entered as a control

Notes. Dashed arrow indicates non-significant path. Path from reactions to negative emotions: coefficients for men/women.

FIGURE 1 | Results of the modified path model (H3 and H4).

variable. Model fit was good [χ^2 (17) = 38.20, p = 0.002, RMSEA = 0.064, CFI = 0.99, TLI = 0.97, SRMR = 0.05]. Modification indices showed that fit could be further improved by relaxing the constraint for the path from reactions to negative emotions. Constraints were relaxed for the direct and indirect effects. Fit of the modified model was good [χ^2 (16) = 33.51, p = 0.006, RMSEA = 0.06, CFI = 0.99, TLI = 0.97, SRMR = 0.05]. A χ^2 difference test showed that the fit of this modified model was significantly better than that of the fully constrained model [$\Delta\chi^2$ (1) = 4.69, p = 0.030]. Results are displayed in **Table 4** (direct effects), **Table 5** (indirect effects), and **Figure 1**. All direct and indirect effects were in the expected direction and significant, except for a non-significant path from anxiety emotions to motivation and the indirect effects *via* anxiety emotions, which were both non-significant. The path from reactions to negative emotions and the indirect effect of reactions to motivation *via* negative emotions were stronger for female than male students. Unexpectedly, the paths from reactions and communion to positive and anxiety emotions were equal for female and male students.

Moderation by Self-Construal

To investigate whether self-construal moderated the relationships from reactions and communion to emotions and motivation (Hypothesis 5), we introduced independent self-construal and its interactions with reactions, communion, and emotions as additional predictors, controlling for country. Predictors involved in the interactions were centered. The additional paths were unconstrained. Model fit was not acceptable [χ^2 (64) = 141.94, p < 0.001, RMSEA = 0.06, CFI = 0.95, TLI = 0.92, SRMR = 0.07]. Inspection of results showed that independent self-construal interacted with reactions and communion to predict emotions, but did not interact with emotions to predict motivation. Thus, we computed a modified model excluding the interactions of emotions and self-construal on motivation. This modified model had

acceptable fit [χ^2 (34) = 71.58, p < 0.001, RMSEA = 0.06, CFI = 0.98, TLI = 0.95, SRMR = 0.06]. For male participants, there was an interaction of communion and independent self-construal to predict negative emotions as well as anxiety emotions. Simple slopes analyses showed that for male participants weakly endorsing an independent self-construal, higher communion ratings predicted higher negative emotions (b = 0.28, *SE* = 0.06, p < 0.001) and higher anxiety emotions (b = 0.38, *SE* = 0.06, p < 0.001), whereas these relationships were non-significant for male participants strongly endorsing and independent self-construal (negative emotions: b = 0.03, *SE* = 0.06, p = 0.594; anxiety emotions: b = 0.10, *SE* = 0.06, p = 0.121) Moreover, self-construal interacted with reactions to predict positive emotions. For male participants weakly endorsing an independent self-construal, the relationship of positive reactions and positive emotions was stronger (b = 0.38, *SE* = 0.05, p < 0.001) than for male participants strongly endorsing an independent self-construal (b = 0.28, *SE* = 0.05, p < 0.001). In turn, for female participants there were no direct or moderated effects of self-construal on positive or negative emotions, but self-construal negatively predicted anxiety emotions and positively predicted motivation. Results are depicted in **Table 6**.

Discussion

Study 3 showed mixed evidence for expected backlash for women in STEM (Hypothesis 2). Women expected their conversation partner to react more positively than men (contrary to expectations), but expected to be rated lower in communion than men (in line with expectations). These results might indicate a subtle expected backlash effect in that female students did not imagine blatant negative reactions to disclosing their STEM major to the conversation partner, but they expected lower communion ratings. This latter result is consistent with lack-of-fit models indicating that women in agentic fields (in the United States) are rated lower in communion for disconfirming the female stereotype (Rudman and Phelan, 2008).

TABLE 6 | Results of path model with moderation by independent self-construal.

	b	[LLCI; ULCI]	SE	p
Reaction				
country	−0.34	[−0.51; −0.17]	0.09	<0.001
Communion				
country	−0.03	[−0.20; 0.14]	0.09	0.754
Independent self-construal				
country	0.01	[−0.14; 0.16]	0.08	0.906
Positive emotions				
reaction	0.47	[0.41; 0.53]	0.03	<0.001
communion	0.37	[0.31; 0.43]	0.03	<0.001
self-construal (men)	−0.03	[−0.10; 0.05]	0.04	0.515
reaction*self-construal (men)	−0.10	[−0.19; −0.02]	0.04	0.020
communion*self-construal (men)	0.07	[−0.02; 0.16]	0.05	0.122
self-construal (women)	0.05	[−0.06; 0.16]	0.06	0.403
reaction*self-construal (women)	0.02	[−0.10; 0.13]	0.06	0.802
communion*self-construal (women)	0.02	[−0.10; 0.14]	0.06	0.769
country	0.14	[0.04; 0.28]	0.06	0.015
Negative emotions				
reaction (men)	−0.46	[−0.54; −0.38]	0.04	<0.001
reaction (women)	−0.55	[−0.66; −0.44]	0.06	<0.001
communion	0.15	[0.08; 0.23]	0.04	<0.001
self-construal (men)	−0.11	[−0.20; −0.01]	0.05	0.023
reaction*self-construal (men)	0.06	[−0.05; 0.16]	0.05	0.300
communion*self-construal (men)	−0.14	[−0.25; −0.03]	0.06	0.012
self-construal (women)	−0.12	[−0.26; 0.02]	0.07	0.102
reaction*self-construal (women)	−0.12	[−0.27; 0.04]	0.08	0.149
communion*self-construal (women)	−0.02	[−0.18; 0.13]	0.08	0.755
country	−0.17	[−0.32; −0.02]	0.08	0.023
Anxiety emotions				
reaction	−0.34	[−0.42; −0.27]	0.04	<0.001
communion	0.24	[0.16; 0.31]	0.04	<0.001
self-construal (men)	−0.22	[−0.31; −0.13]	0.05	<0.001
reaction*self-construal (men)	0.04	[−0.06; 0.15]	0.05	0.416
communion*self-construal (men)	−0.16	[−0.27; −0.05]	0.06	0.004
self-construal (women)	−0.26	[−0.40; −0.11]	0.07	0.001
reaction*self-construal (women)	−0.05	[−0.21; 0.10]	0.08	0.509
communion*self-construal (women)	−0.03	[−0.18; 0.12]	0.08	0.701
country	−0.14	[−0.30; 01]	0.08	0.058
Motivation				
reaction	0.35	[0.27; 0.44]	0.04	<0.001
communion	0.08	[0.00; 0.16]	0.04	0.038
self-construal (men)	0.06	[−0.02; 0.14]	0.04	0.171
self-construal (women)	0.19	[0.05; 0.32]	0.07	0.008
positive emotions	0.13	[0.04; 0.22]	0.05	0.004
negative emotions	−0.13	[−0.23; −0.03]	0.05	0.012
anxiety emotions	0.11	[0.01; 0.21]	0.05	0.040
country	0.19	[0.06; 0.33]	0.07	0.005

Contrary to Hypotheses 3 and 4, the path models largely showed gender similarities. However, results might also point to some negative consequences of the (subtle) expected backlash effect: Less positive reactions were related to more negative emotions (and consequently lower motivation) more strongly for female than male students. This might indicate that (some) women are sensitive to disconfirming the female stereotype and

consequently suffer negative consequences of backlash. Concerning Hypothesis 5, results indicated that self-construal played a moderating role for male, but not for female participants. Lower independent self-construal was associated with stronger relationships of reactions and communion to emotions than higher independent self-construal. This result might indicate that men who see themselves as less independent from social others are more susceptible to possible positive and negative effects of social reactions on their emotions. Whereas positive reactions were related to positive emotions, higher communion was related positive, as well as to negative and anxiety emotions. This pattern of results might represent a double-edged sword of communion for men who place less value on being independent of social others: on the one hand, it is in general socially desirable to be rated high on communion (Steinmetz et al., 2014), on the other hand being perceived as highly communal might induce masculinity threat due to precarious manhood beliefs (e.g., Bosson et al., 2021; Vandello and Bosson, 2013). Unexpectedly, self-construal did not play a moderating role for female participants, indicating that for women, self-construal did not relate to a higher or lower susceptibility to expected backlash effects.

GENERAL DISCUSSION

Women are underrepresented in STEM fields like physics, engineering and computer science around the world, including Germany and Japan, which are top-ranking in STEM (e.g., Nature Index, 2020; U. S. News and World Report LP, 2020). Whereas there are no gender differences in math ability (e.g., Else-Quest et al., 2010; Lindberg et al., 2010), stereotypes play a role in gender segregation in STEM (e.g., Nosek et al., 2009). Based on social role theory (Eagly and Wood, 2012) and role congruity theory (Eagly and Karau, 2002), we expected gender-science stereotypes to be associated with expected backlash (Rudman and Phelan, 2008) for female STEM students, which negatively affects their emotions and motivation.

Studies 1 and 2 showed that widespread gender stereotypes in Germany and Japan associated men with math and science and women with liberal arts (Steffens and Jelenec, 2011). Results were consistent when using mixed methods including open-ended questions, Likert-scale explicit stereotype measurement as well as the Implicit Association Test, spanning multiple years of measurement (2006–2017). Replicating previous research that was conducted in each country separately (Ikkatai et al., 2020; Steffens and Jelenec, 2011), the current research showed in a joint investigation of both countries that in line with Hypothesis 1, negative stereotypes about women's STEM ability were endorsed in both Germany and Japan. We thus conclude that these stereotypes likely contribute to women's underrepresentation in STEM in these countries.

Study 3 investigated expected backlash as a potential consequence of gender-science stereotypes. A scenario study with students of physics, engineering, and computer science from Germany and Japan as participants showed tentative evidence for expected backlash for female STEM students.

Concerning Hypotheses 2–4, gender differences were not as clear and pronounced as expected. Associations between expected reactions and communion to emotions and motivation were largely similar for male and female STEM students. Nevertheless, results point to subtle expected backlash effects for female students: They expected their conversation partner to rate them lower on communion (but not to react more negatively) than male participants. Furthermore, they more frequently expected surprised reactions than men, and 34% of women (0% of men) imagined their conversation partner to refer to their gender in the reactions. We take this as evidence that for women, gender is more salient in the imagined conversation about their study major. This salience might indicate that studying a STEM subject is seen as counter-stereotypical behavior violating prescriptive gender stereotypes. Female STEM students might thus expect that others perceive a lack of fit of women to STEM (Heilman, 1983; Eagly and Karau, 2002).

Results of Study 3 imply expected backlash for women in STEM, but participants did not expect this backlash to be blatantly negative. Backlash can manifest itself in subtle emotional responses like frowning or derisive smiling, which are discussed as possible indicators of implicit social punishment for disconfirming gender stereotypes (Rudman and Phelan, 2008). Such subtle responses can also be conceptualized as micro-aggressions (i.e., "brief, everyday exchanges that send denigrating messages to individuals because of their group membership," Sue, 2010, p. xvii). Gender-based micro-aggressions in STEM contexts have recently gained attention (e.g., Sekaquaptewa, 2019) and might have contributed to subtle expected backlash. Moreover, the associations from less positive reactions to motivation via negative emotions were significantly stronger for female than for male students. This result indicates that even a subtle expected backlash might have negative consequences for female STEM students.

The cross-cultural approach showed that gender-science stereotypes were endorsed in Germany and Japan to a similar extent, corresponding to the comparable underrepresentation of women in STEM in these countries. Furthermore, associations between expected backlash, emotions and motivation remained consistent when country of data collection was statistically controlled for. A model investigating independent self-construal as a moderator showed two noteworthy patterns of results. First, self-construal moderated the paths from reactions/communion to emotions, but not the paths from emotions to motivation. This might indicate that self-construal is more relevant for how social reactions are perceived and which emotions are elicited by these perceptions. In turn, these emotions were associated with motivation to study irrespective of the level of self-construal endorsed, speaking for effects of emotions on motivation that are generalizable across participants' cultural orientations. Second, self-construal moderated paths from reactions/communion to emotions only for male, but not for female participants. Thus, Hypothesis 5, that individuals endorsing an independent self-construal are less prone to expecting negative social repercussions for counter-stereotypical behavior of studying a STEM subject, was only supported for male participants. In contrast, female participants were susceptible to consequences of expected reactions to studying a STEM subject irrespective of the relevance of the group for their self. However, it should be noted that reliabilities of self-construal were at the lower end and unsatisfactory for some groups, calling for a replication with more reliable measures of self-construal.

Results were similar across scenarios, but stronger for the scenario representing low relational mobility. This indicates that the experienced negative consequences of stereotypes might be stronger for women in STEM in contexts in which the social network is more stable and less based on personal choice, as social rejection has more severe consequences in these contexts (Sato et al., 2014). Future studies should substantiate this preliminary evidence that the intensity of consequences of expected backlash for stereotype-incongruent behavior of women in STEM might be aggravated by the cultural factor of relational mobility.

Theoretical and Practical Implications

The present research showed that in the STEM fields of physics, engineering and computer science, similar social-psychological mechanisms as in other male-dominated domains (e.g., leadership) might impede gender equality in Germany and Japan. In accordance with social role theory (Eagly and Wood, 2012), the observation of gender segregation in male-dominated STEM fields is associated with the stereotype that in Germany and Japan, men are stereotypically perceived as better-suited for STEM than women. We applied role congruity theory (Eagly and Karau, 2002) to the STEM context to shed light on the psychological processes contributing to women's underrepresentation in STEM in Germany and Japan. Like in the leadership domain, women might experience backlash effects in gender-segregated STEM fields. Because STEM is incongruent with the traditional female social role, women in STEM might experience social rejection in cultural contexts like Germany and Japan, where negative gender-science stereotypes are widespread (Study 2, but see also Ikkatai et al., 2020; Steffens and Jelenec, 2011).

A recent model advancing role congruity theory describes the interplay of social roles and motivational causes for gender inequality in STEM. Goal congruity theory (Diekman et al., 2017) posits that gender roles build an opportunity structure to fulfill individual (stereotype-congruent) goals. Women tend to strive for communal goals (e.g., helping other people), whereas men tend to strive for agentic goals (e.g., gaining power). By valuing different goals, women and men select into stereotype-congruent roles (study fields and careers). STEM fields are not perceived as affording communal goals. Thus, pursuing STEM creates goal incongruity for women, which can lead to lower motivation and opting out of STEM. The current research showed that women's motivation is impaired by expected negative social reactions to studying a STEM subject. These negative reactions as a signal of lack-of-fit might communicate to women that STEM is perceived as incongruent with their gender role, thereby creating or aggravating goal incongruity. Importantly, participants in Study 3 had already successfully entered STEM majors, which means that they had sufficiently positive initial beliefs about STEM to enroll in this kind of major.

The current research shows that even for these participants, who are highly invested in pursuing a STEM major and potential career, expected negative social reactions can have detrimental consequences for their emotions and motivation. Paired with the widespread gender stereotypes, these consequences can further aggravate gender segregation in STEM, as the leaky pipeline shows that female STEM students are often less inclined to pursue a STEM career than their male counterparts (e.g., Diekman et al., 2017; Jasko et al., 2020).

Goal congruity theory might also explain why we found positive relationships of communion to positive, negative, and anxiety emotions for male as well as female students. Communal goals represent the basic need of relatedness and are thus important to everyone (Diekman et al., 2017). Moreover, communion is socially valued (e.g., Abele et al., 2008). Therefore, being perceived as communal in the scenarios was associated with more positive emotions for both genders. However, because STEM is perceived as incongruent with communal goals, goal conflict is likely to arise. Goal conflict can elicit anxiety and negative emotions (Gray and McNaughton, 2003), potentially explaining why communion was related to higher anxiety and negative emotions for both genders.

Results open up pathways to reduce women's underrepresentation in these STEM fields. Studies 1 and 2 showed that gender-science stereotypes are pervasive in Germany and Japan. There have been efforts to develop educational programs to reduce gender stereotypes and their effects, for example, focusing on teaching students a growth mindset or motivational and strategic trainings (e.g., Law et al., 2021; Moè, 2021). As changing stereotypes has been shown to be quite difficult (Heilman and Caleo, 2018), another fruitful road to gender equality in STEM in Germany and Japan is to reduce role and goal incongruity. An intervention to change communal goal affordances (i.e., the opportunities for goal pursuit) in STEM (Belanger et al., 2020) showed that perceiving communal goal affordances (e.g., collaborative lab activities) in STEM increased social belonging and interest, especially for women. Highlighting STEM's potential to afford communal goals might therefore alleviate goal incongruity and reduce gender-science stereotypes and backlash effects for women, because STEM is perceived as less incongruent with the female gender role.

Limitations and Future Directions

A first limitation of the current research is the measurement of social reactions in Study 3. Although we used a combination of open-ended and Likert-scale questions and participants were asked to imagine both verbal and non-verbal reactions, the items captured rather blatant than subtle reactions. Future research should investigate a broader variety of reactions to disconfirming stereotypes in STEM. Second, Study 3 measured backlash and its consequences only cross-sectionally. As goal incongruity might be anticipated and repeatedly experienced before it has detrimental consequences for women's STEM motivation (Diekman et al., 2017), future research should investigate consequences of backlash and incongruity in longitudinal studies. In addition, we compared two countries

with similar gender segregation in STEM (e.g., Destatis, 2019; Gender Equality Bureau Cabinet Office, 2017), but different cultural orientations (self-construal, relational mobility; e.g., Markus and Kitayama, 1991; Thomson et al., 2018). To fully unfold the possible interplay of social roles, gender, and culture for women in STEM, future research should investigate a larger sample of countries with varying positions on the individualistic/post-materialistic vs. collectivistic/materialistic continuum (Yalcinkaya and Adams, 2020) as well as varying levels of gender inequality, as previous cross-cultural research conducted in different European countries (e.g., Italy, Norway, Poland, Spain, United Kingdom) has shown that the extent and consequences of gender stereotypes may be in part shaped by a country's gender inequality (e.g., Castaño et al., 2020; Bedyńska et al., 2021; Moè et al., 2021).

Third, internal consistency of self-construal was low and relational mobility was not measured on the individual level, but varied in two scenarios. The inclusion of further scales to measure these cultural variables and further moderators and mediators (e.g., perceived goal conflict) could illuminate individual factors increasing women's susceptibility to backlash and role/goal incongruity. Another possible mediator could be rejection sensitivity (Sato et al., 2014), which might explain why gender was more salient in the scenarios for some of the female participants.

Fourth, similar to many other cross-cultural studies, we were not able to establish full scalar invariance of the multi-item measures used in Study 3. The level of partial metric invariance was reached in that factor loadings of at least two items per construct were equal. This enabled us to test relationships between variables in path analysis. However, we note that intercepts were not equal between national subsamples and we therefore refrained from estimating latent variables in structural equation modeling.

Finally, the scenarios might have activated occupational stereotypes along with gender stereotypes. People–particularly men–in STEM are stereotyped to be "socially awkward" (Cheryan et al., 2013). Therefore, in addition to gender stereotypes, occupational stereotypes about social skills might have been activated. These stereotypes might have caused men to also expect backlash to their study major. Future research should therefore disentangle backlash due to gender and occupational stereotypes. As in the current study scenarios were limited to social interactions outside of STEM, it might be also worthwhile to investigate expected backlash effects in further scenarios that are related to the academic/work domain.

CONCLUSION

Factors explaining gender inequality in STEM are manifold. The present research adds to the literature by investigating social-psychological and cultural mechanisms to relatively low STEM motivation for women in Germany and Japan. A mixed-methods

investigation of gender-science stereotypes confirmed negative stereotypes about women in STEM in both countries. Even though Germany and Japan differ in cultural orientations, the impact of stereotypes on gender segregation in STEM seems to be pervasive in both countries. Recent promising measures to reduce gender inequality do not focus on changing women's individual predictors of STEM success, but rather investigate how STEM is stereotyped. As stereotypes are socially and culturally shared, cross-cultural research may further illuminate the social context of gender inequality in STEM.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Open Science Framework, https://osf.io/4awqe/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the research ethics commissioner of the

FernUniversität in Hagen. The patients/participants provided their written informed consent to participate in this study.

REFERENCES

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., and Duan, Y. (2016). Facets of the Fundamental Content Dimensions: Agency with Competence and Assertiveness-Communion with Warmth and Morality. *Front. Psychol.* 7, 1810. doi:10.3389/fpsyg.2016.01810

Abele, A. E., Uchronski, M., Suitner, C., and Wojciszke, B. (2008). Towards an Operationalization of the Fundamental Dimensions of agency and Communion: Trait Content Ratings in Five Countries Considering Valence and Frequency of Word Occurrence. *Eur. J. Soc. Psychol.* 38 (7), 1202–1217. doi:10.1002/ejsp.575

Adachi, T. (2014). Occupational Gender Stereotypes Among university Students: Their Relationships with Self-Efficacy and Gender Role Attitudes. *Jpn. Assoc. Industrial/Organizational Psychol. J.* 27 (2), 87–100.

Bedyńska, S., Rycielski, P., and Jabłońska, M. (2021). Measuring Stereotype Threat at Math and Language Arts in Secondary School: Validation of a Questionnaire. *Front. Psychol.* 12, 553964. doi:10.3389/fpsyg.2021.553964

Belanger, A. L., Joshi, M. P., Fuesting, M. A., Weisgram, E. S., Claypool, H. M., and Diekman, A. B. (2020). Putting Belonging in Context: Communal Affordances Signal Belonging in STEM. *Pers Soc. Psychol. Bull.* 46 (8), 1186–1204. doi:10.1177/0146167219897181

Bosson, J. K., Jurek, P., Vandello, J. A., Kosakowska-Berezecka, N., Olech, M., Besta, T., et al. (2021). Psychometric Properties and Correlates of Precarious Manhood Beliefs in 62 Nations. *J. Cross-Cultural Psychol.* 52 (3), 231–258. doi:10.1177/0022022121997997

Castaño, A. M., Lubiano, M. A., and García-Izquierdo, A. L. (2020). Gendered Beliefs in STEM Undergraduates: A Comparative Analysis of Fuzzy Rating versus Likert Scales. *Sustainability* 12 (15), 6227. doi:10.3390/su12156227

Cheryan, S., Master, A., and Meltzoff, A. N. (2015). Cultural Stereotypes as Gatekeepers: Increasing Girls' Interest in Computer Science and Engineering by Diversifying Stereotypes. *Front. Psychol.* 6, 49. doi:10.3389/fpsyg.2015.00049

Cheryan, S., Plaut, V. C., Davies, P. G., and Steele, C. M. (2009). Ambient Belonging: How Stereotypical Cues Impact Gender Participation in Computer Science. *J. Pers Soc. Psychol.* 97 (6), 1045–1060. doi:10.1037/a0016239

Cheryan, S., Ziegler, S. A., Montoya, A. K., and Jiang, L. (2017). Why Are Some STEM fields More Gender Balanced than Others? *Psychol. Bull.* 143 (1), 1–35. doi:10.1037/bul0000052

Cheryan, S., Plaut, V. C., Handron, C., and Hudson, L. (2013). The Stereotypical Computer Scientist: Gendered media Representations as a Barrier to Inclusion for Women. *Sex Roles* 69 (1-2), 58–71. doi:10.1007/s11199-013-0296-x

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. Hillsdale, N.J.: L. Erlbaum Associates.

Credit Suisse Research Institute (2019). Global Wealth Report 2019. Retrieved from https://www.credit-suisse.com/about-us/en/reports-research/global-wealth-report.html.

Cundiff, J. L., Vescio, T. K., Loken, E., and Lo, L. (2013). Do gender-science Stereotypes Predict Science Identification and Science Career Aspirations Among Undergraduate Science Majors? *Soc. Psychol. Educ.* 16 (4), 541–554. doi:10.1007/s11218-013-9232-8

Destatis (2019). *Studierende in Mathematik, Informatik, Naturwissenschaft (MINT) und Technik-Fächern.* Wiesbaden, Germany: Students in mathematics, computer science, science and techology (STEM) fields. Retrieved from https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Tabellen/studierende-mint-faechern.html.

Diekman, A. B., Steinberg, M., Brown, E. R., Belanger, A. L., and Clark, E. K. (2017). A Goal Congruity Model of Role Entry, Engagement, and Exit: Understanding Communal Goal Processes in STEM Gender Gaps. *Pers Soc. Psychol. Rev.* 21 (2), 142–175. doi:10.1177/1088868316642141

Eagly, A. H., and Karau, S. J. (2002). Role Congruity Theory of Prejudice toward Female Leaders. *Psychol. Rev.* 109 (3), 573–598. doi:10.1037/0033-295x.109.3.573

Eagly, A. H., and Wood, W. (2012). "Social Role Theory," in *Handbook of Theories of Social Psychology, Volume II.* Editors P. van Lange, A. Kruglanski, and E. Higgins (Los Angeles, [Calif.]: SAGE), 458–476. doi:10.4135/9781446249222.n49

Eaton, A. A., Saunders, J. F., Jacobson, R. K., and West, K. (2020). How Gender and Race Stereotypes Impact the Advancement of Scholars in STEM: Professors' Biased Evaluations of Physics and Biology Post-Doctoral Candidates. *Sex Roles* 82 (3-4), 127–141. doi:10.1007/s11199-019-01052-w

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 (1), 103–127. doi:10.1037/a0018053

Gender Equality Bureau Cabinet Office (2017). *Gender Equality White Paper (Summary Version) 2017 Edition*. Tokyo, Japan: White paper on gender equality 2017. Retrieved from http://www.gender.go.jp/about_danjo/whitepaper/h29/gaiyou/html/honpen/b1_s05.html.

Gino, F., Wilmuth, C. A., and Brooks, A. W. (2015). Compared to Men, Women View Professional Advancement as Equally Attainable, but Less Desirable. *Proc. Natl. Acad. Sci. U S A.* 112 (40), 12354–12359. doi:10.1073/pnas.1502567112

Gray, J. A., and McNaughton, N. (2003). *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*. Oxford: Oxford University Press.

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998). Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *J. Pers Soc. Psychol.* 74, 1464–1480. doi:10.1037//0022-3514.74.6.1464

Gudykunst, W. B., and Lee, C. M. (2003). Assessing the Validity of Self Construal Scales. *Hum. Comm Res* 29 (2), 253–274. doi:10.1111/j.1468-2958.2003.tb00838.x

Heilman, M. E., and Caleo, S. (2018). Combatting Gender Discrimination: A Lack of Fit Framework. *Group Process. Intergroup Relations* 21 (5), 725–744. doi:10.1177/1368430218761587

Heilman, M. E. (1983). Sex Bias in Work Settings: The Lack of Fit Model. *Res. Organizational Behav.* 5, 269–298.

Henrich, J., Heine, S. J., and Norenzayan, A. (2010). Most People Are Not WEIRD. *Nature* 466 (7302), 29. doi:10.1038/466029a

Ikkatai, Y., Inoue, A., Kano, K., Minamizaki, A., McKay, E., and Yokoyama, H. M. (2019). Parental Egalitarian Attitudes towards Gender Roles Affect Agreement on Girls Taking STEM fields at university in Japan. *Int. J. Sci. Edu.* 41 (16), 2254–2270. doi:10.1080/09500693.2019.1671635

Ikkatai, Y., Minamizaki, A., Kano, K., Inoue, A., McKay, E., and Yokoyama, H. M. (2020). Gender-biased Public Perception of STEM fields, Focusing on the Influence of Egalitarian Attitudes toward Gender Roles. *Jcom* 19 (01), A08. doi:10.22323/2.19010208

Jasko, K., Pyrkosz-Pacyna, J., Czarnek, G., Dukała, K., and Szastok, M. (2020). The STEM Graduate: Immediately after Graduation, Men and Women Already Differ in Job Outcomes, Attributions for success, and Desired Job Characteristics. *J. Soc. Issues* 76 (3), 512–542. doi:10.1111/josi.12392

Kassin, S. M., Fein, S., and Markus, H. (2011). *Social Psychology*. 7th ed. Belmont, Calif., United Kingdom: Wadsworth.

Kemmelmeier, M. (2016). Cultural Differences in Survey Responding: Issues and Insights in the Study of Response Biases. *Int. J. Psychol.* 51 (6), 439–444. doi:10.1002/ijop.12386

Law, F., McGuire, L., Winterbottom, M., and Rutland, A. (2021). Children's Gender Stereotypes in STEM Following a One-Shot Growth Mindset Intervention in a Science Museum. *Front. Psychol.* 12, 641695. doi:10.3389/fpsyg.2021.641695

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 (6), 1123–1135. doi:10.1037/a0021276

Markus, H. R., and Kitayama, S. (1991). Culture and the Self: Implications for Cognition, Emotion, and Motivation. *Psychol. Rev.* 98 (2), 224–253. doi:10.1037/0033-295X.98.2.224

Miller, D. I., Eagly, A. H., and Linn, M. C. (2015). Women's Representation in Science Predicts National Gender-Science Stereotypes: Evidence from 66 Nations. *J. Educ. Psychol.* 107, 631–644. doi:10.1037/edu0000005

Moè, A., Hausmann, M., and Hirnstein, M. (2021). Gender Stereotypes and Incremental Beliefs in STEM and Non-STEM Students in Three Countries: Relationships with Performance in Cognitive Tasks. *Psychol. Res.* 85 (2), 554–567. doi:10.1007/s00426-019-01285-0

Moè, A. (2021). Doubling Mental Rotation Scores in High School Students: Effects of Motivational and Strategic Trainings. *Learn. Instruction* 74, 101461. doi:10.1016/j.learninstruc.2021.101461

Morinaga, Y., Sakata, K., Furukawa, Y., and Fukudome, K. (2017). Mathematics Motivation and Gender Stereotypes of Junior and Senior High School Girls. *Kyoiku shinrigaku kenkyu* 65 (3), 375–387. doi:10.5926/jjep.65.375

Nature Index (2020). The Ten Leading Countries in Natural-Sciences Research. *Nature.* 580, S44.

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. *Proc. Natl. Acad. Sci.* 106 (26), 10593–10597. doi:10.1073/pnas.0809921106

OECD (2015). Can the Performance gap between Immigrant and Non-immigrant Students Be Closed? *PISA in Focus* 53, 1–4. doi:10.1787/22260919

Park, J., and Kitayama, S. (2014). Interdependent Selves Show Face-Induced Facilitation of Error Processing: Cultural Neuroscience of Self-Threat. *Soc. Cogn. Affective Neurosci.* 9 (2), 201–208. doi:10.1093/scan/nss125

Rudman, L. A., and Glick, P. (2001). Prescriptive Gender Stereotypes and Backlash toward Agentic Women. *J. Soc. Issues* 57 (4), 743–762. doi:10.1111/0022-4537.00239

Rudman, L. A., and Phelan, J. E. (2008). Backlash Effects for Disconfirming Gender Stereotypes in Organizations. *Res. Organizational Behav.* 28, 61–79. doi:10.1016/j.riob.2008.04.003

Sato, K., Yuki, M., and Norasakkunkit, V. (2014). A Socio-Ecological Approach to Cross-Cultural Differences in the Sensitivity to Social Rejection. *J. Cross-Cultural Psychol.* 45 (10), 1549–1560. doi:10.1177/0022022114544320

Sekaquaptewa, D. (2019). Gender-based Microaggressions in STEM Settings. *NCID Currents* 1 (1), 1. doi:10.3998/currents.17387731.0001.101

Smith, P. B. (2014). "Response Bias(es)," in *Encyclopedia of Quality of Life and Well-Being Research*. Editor A. C. Michalos (Dordrecht: Springer Reference), 5539–5540. doi:10.1007/978-94-007-0753-5_2503

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 (5-6), 324–335. doi:10.1007/s11199-010-9924-x

Steinmetz, J., Bosak, J., Sczesny, S., and Eagly, A. H. (2014). Social Role Effects on Gender Stereotyping in Germany and Japan. *Asian J. Soc. Psychol.* 17 (1), 52–60. doi:10.1111/ajsp.12044

Stout, J. G., Grunberg, V. A., and Ito, T. A. (2016). Gender Roles and Stereotypes about Science Careers Help Explain Women and Men's Science Pursuits. *Sex Roles* 75 (9-10), 490–499. doi:10.1007/s11199-016-0647-5

Sue, D. W. (2010). *Microaggressions in Everyday Life: Race, Gender, and Sexual Orientation*. Hoboken, N.J.: Wiley.

Thomson, R., Yuki, M., Talhelm, T., Schug, J., Kito, M., Ayanian, A. H., et al. (2018). Relational Mobility Predicts Social Behaviors in 39 Countries and Is Tied to Historical Farming and Threat. *Proc. Natl. Acad. Sci.* 115 (29), 7521–7526. doi:10.1073/pnas.1713191115

U. S. News and World Report LP (2020). Best Countries 2020. Global Rankings, International News and Data Insights. Retrieved from https://www.usnews.com/media/best-countries/overall-rankings-2020.pdf.

Vandello, J. A., and Bosson, J. K. (2013). Hard Won and Easily Lost: A Review and Synthesis of Theory and Research on Precarious Manhood. *Psychol. Men Masculinity* 14 (2), 101–113. doi:10.1037/a0029826

Varnum, M., Grossmann, I., Kitayama, S., and Nisbett, R. E. (2010). The Origin of Cultural Differences in Cognition: Evidence for the Social Orientation Hypothesis. *Curr. Dir. Psychol. Sci.* 19 (1), 9–13. doi:10.1177/0963721409359301

Williams, J. E., and Best, D. L. (1991). *Measuring Sex Steretoypes: A Multination Study*. New York: SAGE.

Yalcinkaya, N. S., and Adams, G. (2020). A Cultural Psychological Model of Cross-National Variation in Gender Gaps in STEM Participation. *Personal. Soc. Psychol. Rev.* 24, 345–370. doi:10.1177/1088868320947005