Gender bias in the classroom: A network study on self and peer ability attribution

Elena De Gioannis* Federico Bianchi Flaminio Squazzoni

University of Milan University of Milan University of Milan

This is the manuscript accepted for publication by Social Networks (Published version here). Please cite as follows:

De Gioannis, E., Bianchi, F., Squazzoni, F. (2023). Gender bias in the classroom: A network study on self and peer ability attribution. *Social Networks*, 72: 44-51. doi: https://doi.org/10.1016/j.socnet.2022.09.001

Abstract

Stereotypes can contribute to the gender gap in STEM by shaping people's expectations on their own and others' performance. When gender is salient, expectations on task performance might reflect gender constructs even when information on individual abilities is available. We tested this hypothesis in a network study on students from ten high school classes in Milan, Italy. We asked the students to choose the four best candidates from their classmates for three hypothetical inter-class competitions in reading, math, and science. Results showed that females were more likely to be nominated for the reading competition but less likely for science. We did not find any statistically significant results for the math competition. We also found that female students were less likely to nominate themselves for any competition, regardless of the subject, even controlling for their own performance and self-concept.

I. Introduction

In the United States, women make up only less than one-third of the workforce in STEM (Science, Technology, Engineering, and Mathematics) sectors, and men still vastly outnumber women when majoring in these fields [Bureau of Census, 2019]. In Europe in 2017, 41% of scientists and engineers in Europe were women [Eurostat, 2019]. Despite significant recent progress, enhanced by diversity and inclusion policies targeting women in various sectors, gender stereotypes still seem to be deeply ingrained in perceptions and practices across all spectrums of society [Ellemers, 2018].

While the uneven distribution of men and women in STEM occupations reinforces gender associations [Eagly and Steffen, 1984], stereotypes are also rooted in the corresponding belief that women and men differ in logical and reading skills [Kersey et al., 2019]. These symmetrical beliefs consolidate "widely shared, hegemonic cultural beliefs about gender and their effects" and

*elena.degioannis@unimi.it

shape (re-shape) "social relational contexts", which in turn could reproduce these gender patterns [Ridgeway and Correll, 2004, p. 511].

Despite important initiatives in many countries aimed at reducing these gender patterns [OECD, 2015], together with the changing nature of collective beliefs, gender differences in attitudes and roles in STEM are still persistent. For instance, recent research in Italy found that the endorsement of stereotypical beliefs on math and reading abilities was prominent among children and their parents [Galdi et al., 2017, Passolunghi et al., 2014, Tomasetto et al., 2015, 2012], as well as among teachers [Carlana, 2019].

The endorsement of these gender stereotypes often leads to faulty assessments, which in turn influence expectations on people's performance, thus biasing opportunities and work outcomes for both men and women [Hentschel et al., 2019]. For instance, in the context of education, where abilities have a prominent role, peers influence choices and intentions, especially of female students in STEM [Dasgupta et al., 2015, Riegle-Crumb and Morton, 2017, Robnett and Leaper, 2013, van der Vleuten et al., 2018]. Unfortunately, previous studies on gender stereotypes have mostly focused on participants' beliefs about the performance of certain social categories (i.e., "men", "women"), e.g., by posing questions such as: "who is better at science and math?". This implies neglecting opinion on the performance of specific, familiar individuals, such as classmates in the school context, which would be captured by questions such as: "Who is better at science and math *in your class*?". While it may be questionable as to whether we can expect an influence of gender in peers' assessment of performance, since ability in science and math is known, the answer is not straightforward.

As theorized by the 'status characteristic theory' [Berger et al., 1972], when gender is salient, i.e., when gender stereotypes may be activated, gendered beliefs on performance could play a relevant role in determining expected differences between men and women in the outcome, even when information on ability is available. Understanding this link between gender and performance expectations when ability information is available, requires reconstructing a context-specific, social construction process involving information, expectations, and social group pressures.

Unfortunately, it is hard to find suitable data to understand such complexity. On the one hand, we should identify a context where a small social group engages in potential gender-biased behavior. On the other hand, data should include information on both personal and relational characteristics of all social group members. As reported in previous studies on peer assessment [Kench et al., 2009], when asked to evaluate peers, people tend to show loyalty to friends and award them a higher evaluation compared to non-friends. Controlling for friendship bias whenever testing this association would be necessary. However, this type of data is often difficult to access and this is why only a few studies have tested this hypothesis [Grow et al., 2016, Kisfalusi et al., 2019].

In order to fill this gap, we ran a network study on a set of mixed-gender high school classes in Italy (Level 3, ISCED 2011; 12th grade). Following the assumptions of the 'status characteristic theory', we tested whether gender could affect the formation of expectations on classmates' performance, even in contexts where information on students' abilities is available. By designing a hypothetical academic competition where students were required to nominate members of a team representing the classroom, we could elicit students' assessments of their own as well as their classmates' expected performance.

Our objective here was twofold. On the one hand, we aimed to understand whether gender influences students' expectations on their classmates' performance in disciplines typically considered either feminine or masculine. By modeling peers' expectations as social networks [Grow et al., 2016, Kisfalusi et al., 2019, Paluck, 2011, Shepherd and Paluck, 2015], we wanted to verify the effect of gender on peers' expectations on performance, yet controlling for the potential effect of

students abilities and friendship relationships [Grow et al., 2016]. On the other hand, we wanted to test the effect of gender in the self-evaluation of these skills.

Our contribution on the study of gender bias in STEM and humanities is manyfold. First, we provide new insights on the application of the status characteristic theory and its assumptions on the combination of multiple salient status characteristics in an 'out-of- lab' setting. By means of a network study performed in a school context, we had a population target similar to a lab-experiment design, i.e., young students, but we took our observations from a 'natural' social environment. Second, we propose a new and indirect method to test gender bias on reading/mathematical skills, which allowed us to avoid social desirability bias (details in the methodology section). We then provided insights on gender differences in the propensity of be involved in competitive tasks, especially in a context in which gender stereotypes may be activated. Finally, we expanded this recent line of research on bias in the case of stereotypes on boys and reading, often neglected in previous research.

II. Research background

Status attribution

In situations in which gender is salient, inequalities arise because women and men are attributed a status of either inferiority or superiority purely based on their gender. Status attribution tends to generate and reinforce gender inequalities by determining the power, prestige, and influence that actors exert while interacting with others [Berger et al., 1972, Wagner and Berger, 1997]. According to the "status characteristic theory" (SCT), two or more interacting actors evaluate each other based on known salient characteristics, "the states of which are differentially evaluated", and are called "status characteristics" [Berger et al., 1972, p. 242]. These characteristics have two states, which reflect a socially organized hierarchy of meanings related to individual capacities (i.e., high/positive or low/negative), and so "provide the basis for inferring differences" in power, prestige, influence, participation, performance evaluation, and expectations [Berger et al., 1972, Wagner and Berger, 1997]. For a certain task, individuals in the lower state "receive less attention, they are given lower evaluations and they exert less influence" [Foddy and Smithson, 1999, p. 308], while "those with a status advantage will adopt a repertoire of attitudes and behaviors that is associated with their status superiority" [Wagner et al., p. 48]. In contexts where gender is a salient characteristic, individuals may form gender-based expectations of others' and personal performance in a task.

Berger et al. [1972] argued that this gender attribution may persist even when other information is available, e.g., abilities. When status characteristics are multiple, subjects tend to combine status information rather than selecting it. This implies that ability would magnify gender-based expectations if information was consistent. Otherwise, these expectations would change slightly or even drastically, according to new information [Wagner and Berger, 1997]. For instance, Pugh and Wahrman [1983] grouped participants in a laboratory experiment into mixed-sex couples and asked them to perform a task individually to then decide which of the two performances they wanted to submit to researchers. They tested three combinations of information on gender and ability, i.e., no information on ability while gender was said to be irrelevant to the task, ability relevant to the task and equal performance of men and women, and finally, ability relevant to the task and higher performance of women. They found that women were less willing to conform to their counterpart's answers, whereas men were less inclined to impose their idea only when information on women's superiority in the task was available. In the other two cases, neither men nor women changed their behavior compared to the control group (no information).

Ability attribution in a school setting

Schools are an ideal setting to examine this type of gender bias in 'outside the lab', real-world social settings. Unfortunately, only a few studies have analyzed the association between gender and ability attribution in school settings. Grow et al. [2016] analyzed the effect of gender and ethnicity on ability attribution in a sample of Hungarian secondary-school classes. Results revealed that perceived ethnicity affected ability attribution, while gender did not. Nonetheless, the effect of gender was found in a more recent study by Kisfalusi et al. [2019], who tried to disentangle the effect of status generalization and social identity on ability attribution. Results showed a tendency among primary school pupils to nominate classmates who preferably shared their own gender and ethnicity as 'clever students'.

However, these studies focused on estimating gender effects on general academic skills rather than skills in specific school subjects. While children tend to consider their own gender as the cleverest, adolescents are less sensitive toward in-group favoritism but are more sensitive to societal stereotypes [Passolunghi et al., 2014, p. 4]. This implies that the salience of gender in a group interaction of young adults depends on the type of ability required by a task. Societal stereotypes describe men as good at math with women having higher reading abilities [Steffens and Jelenec, 2011]. Thus, the gender -attributed high status would be male if the task requires mathematical skills, female if the task requires reading skills. Therefore, when studying ability attribution within a school setting with young adults, testing the role of gender requires distinguishing among school subjects.

Competitive context and expectations

In a competitive task setting, other individual characteristics could influence expectations on performance and consequently nominations. Previous studies found a general tendency for women to be less competitive than men [Buser et al., 2012], suffering more from anxiety triggered by performance and competition [Baraskar and Shinde, 2018], and being more risk-averse [Fisk, 2018]. This tendency is further exacerbated when women compete in what is conceived as a masculine field.

More specifically concerning science, research has indicated that women would be more reluctant to be involved in mathematics and science competitions [see Steegh et al., 2019]. Gender differences emerged in both participation and achievement in math and science Olympiads, but not in events which, unlike Olympiads, are non-competitive problem-solving occasions. More generally, Günther et al. [2010] found that women were less likely to compete with men in situations in which they believed they could lose, regardless of the realism of such an expectation.

Since information on classmates' attitudes toward competition is rarely available to students, it is probable that the competitive nature of the task can influence self-nominations but not peers' nominations.

III. Hypotheses

According to the status characteristic theory, whenever gender is salient, people would form gender-based expectations on performance in a task even when information on abilities is available. More specifically, in our case, a higher status would be attributed to women for the reading competition and men for the math/science competition, even controlling for the potential effect of abilities in these subjects. Therefore, we formulated the following hypotheses:

H1a: Female students will be more likely to be nominated than male students for the reading competition, even when controlling for grades of nominated students and existing friendship relationships.

H1b: Male students will be more likely to be nominated than female students for the competition in mathematics and science, even when controlling for grades of nominated students and existing friendship relationships.

Moreover, given the task in our setting, we expected young women to be affected by their attitude toward competition whenever deciding whether to candidate themselves or not. This led us to formulate the following hypothesis on self-nominations.

H2: Female students will be less likely to nominate themselves than male students for all three competitions.

IV. Methodology

Data

Data were collected from a sample of high school classes in Milan (Italy) between January-February 2020, by means of a computer-assisted survey. School principals were contacted in November 2019 by a postal letter inviting them to participate in the study. Due to the COVID-19 outbreak in March 2020, only five schools eventually agreed to participate, with a total of ten age-homogenous classes (195 students, 56% female). Subjects were 18-19 years old and were attending the last year of high school (Level 3, ISCEED 2011). All students from each selected class provided their written informed consent. As a 'floating teacher' system, Italian secondary schools provide a particularly useful environment to study long-term mixed-gender peer relationships with shared information on school grades. In this system, students are assigned to one class over the whole school course (5 years) with the same set of classmates. Each class is assigned a classroom where all lectures are delivered by different teachers, who are requested to move between different classrooms. Therefore, students share the same classroom with the same set of classmates for a considerable amount of hours during weekdays for a 5-year long period, on average.

We, therefore, assumed that social relationships within classes were relatively stable, and students had well-established and clear perceptions of their classmates' abilities. All classes belonged to *liceo* secondary schools, a type of secondary school whose aim is to prepare students for tertiary education. While there are six types of *liceo*, all have certain core subjects in common, including reading, mathematics, and science. We chose to target students in Milan because it is one of the largest metropolitan areas in Italy, and so there was a wide choice of high schools and a relatively comparable student population in terms of socio-economic background.

Indirect measurement of gender bias

Students were confronted with a hypothetical, yet realistic, situation, i.e., an inter-class student competition, similar to those regularly involving Italian schools (e.g., *Olimpiadi della Matematica* for a math equivalent of the AMC - American Math Competition in the U.S.). To provide an incentive to choose the best candidates, students were told that in the case of victory, the prize would have been a collective trip to a European city for the whole class. Students were asked to choose four

classmates (including themselves if they wanted), who in their opinion would have been the best candidate for three hypothetical school competitions, one in reading, one in math, and the last in science. By eliciting nominations for a fictional competition and integrating this information with data on the nominees' gender, we obtained an indirect measurement of gender bias.

Compared to traditional approaches, our instrument allowed us to avoid social desirability bias and the influence of other self-presentational factors [Greenwald et al., 1998]. Social desirability bias refers to the tendency of respondents in a survey to answer in a way which reflects what they believe is socially acceptable rather than what they really think. Stereotypes and gender bias are typically subject to this issue. In social surveys, in particular, respondents could try either to hide their true beliefs from the interviewer intentionally, in order to avoid being considered stereotyped (self-presentation to others), or adapt their immediate answer to one associated with their perceived best version of themselves (self-presentation to self) [Greenwald et al., 1998].

Furthermore, since the aim of the study was to measure participants' expectations on their peers' performance, designing an instrument that stimulated students to report their opinion rather than that of their teachers was necessary. For instance, if we had asked students to nominate the best classmates in a subject, they would have probably immediately nominated those with the highest grades. We would have, thus, ended up with an instrument measuring teachers' evaluations rather than the classmates' evaluations.

Nevertheless, school grades here do play an important role. On the one hand, they are used as a proxy for ability information available to students. On the other hand, in evaluating gender bias in nominations, we assumed that school grades measured the student's real ability in the subject. However, it is worth noting that this instrument would have two potential weaknesses whenever (1) students ignore their classmates' grades because of missing any information about their respective abilities, and (2) school grades do not necessarily measure actual ability in a subject.

As regard the first issue, Italian students have constant access to information on all class members' school grades. Indeed, the Italian law on data privacy in schools establishes that the final term grades must be public [GPDP, 2010]. Furthermore, students are continuously evaluated through multiple tests during school terms. In order to ensure the transparency of teachers' evaluation, any oral test must be conducted in front of all classmates and their evaluation must be communicated to each student "promptly and publicly" [DPR, 2019, p. 2].

As regard teachers' evaluations, grades may also reflect other factors that are not strictly related to abilities. Previous studies found that grades are influenced by gender differences in resistance to schooling [Geven et al., 2017], e.g., skipping class and effort in school, as well as teachers' bias; a problem known as the "gender grading gap" [Protivínský and Münich, 2018]. Evidence shows that teachers generally tend to favor female students and give them higher grades [Voyer and Voyer, 2014], except for math-related subjects [Hofer, 2015, Spear, 1984]. Yet with these limitations, teachers' evaluations can be considered the cornerstone of the school system, especially in Italy where test standardization is rare. While partially biased, grades generally are a good measure of students' actual ability in school subjects. Finally, we decided to add science as a third context as, despite being related to the same sector, the gender gap in math and science differs [Cheryan et al., 2017]. In Italy in 2019, the percentage of female undergraduates was 75% in biology, 32% in physics, and 49% in mathematics [AlmaLaurea, 2019]. Furthermore, given that research mostly focuses on the math domain, unfortunately there is scant knowledge on other scientific sectors.

Instruments

Following Grow et al. [2016] and Kisfalusi et al. [2019], we collected relational data on the subjects' expectations on classmates' performance, together with individual data relating to the subjects' socio-demographic characteristics and school grades. Data were collected through a computer-assisted questionnaire administered to each class during regular school time. English translations of the questions are provided in the Supplementary Material.

Relational data

Relational data were collected by means of sociometric questions formatted according to the conventional repeated roster method for name-generators [Kilduff and Krackhardt, 2008].

Nominations: We asked each student to nominate the four best candidates from their classmates for three hypothetical inter-class competitions in reading, math, and science. We then built three tie-variables, so that for each tie-variable X, a dyad $x_{ij} = 1$ if subject i nominated subject j as a candidate for the competition in the specified subject, otherwise $x_{ij} = 0$. Students were allowed to nominate other classmates ($i \neq j$) or themselves ($i = j$). They were confronted with a hypothetical, yet realistic situation, i.e., an inter-class student competition, similar to those regularly involving Italian schools (e.g., *Olimpiadi della Matematica* for a math equivalent of the AMC - American Math Competition in the U.S.). In order to provide an incentive to choose the best candidates, students were told that in the case of victory, the prize would have been a collective trip to a European city for the whole class.

Friendship: We asked students to nominate those classmates whom they considered to be 'friends', so that $x_{ij} = 1$ if i considered j to be a 'friend', while $x_{ij} = 1$ otherwise, with ($i \neq j$).

Individual data

Gender: We asked students to report their gender by selecting one of the following categories: "male", "female", and "other". Given that only one student selected the latter option, we decided to exclude the record from the analysis.

Grades: We asked students to report their grades obtained in reading, math, and science, as they were published as the final evaluation of the previous school year. In Italy, final grades range between six and ten. Students with a final grade below six are not allowed to the following academic year.

Self-concept: Following Gilbert [2015], we asked students to rate on a 5-point Likert scale the extent to which they identified themselves with reading, math, and science (detail in the Supplementary Material) on a 5-item battery. This included: relevance of the subject, relevance attributed to others' opinion about their ability in the subject, reaction to a failure in a school test in the subject, the relevance of the ability in the subject for future career and success in college (Cronbach alpha 0.79 in reading, 0.85 in math, 0.89 in science).

V. Analytical methods

We used two different analytical approaches to test the two hypotheses. Following Grow et al. [2016], Hypothesis 1 was tested through a meta-analysis of class-level Exponential Random Graph Models (ERGMs) [Lusher et al., 2012, Robins et al., 2007], which allowed us to model statistical dependencies of student nominations. Hypothesis 2 was tested using a logistic regression model with fixed effects.

Hypothesis 1

We estimated a set of ERGMs of nominations for each subject (reading, math, and science) on seven of our class samples. We decided to exclude three classes from the sample because of the high percentage of missing data on ties and individual characteristics. The use of ERGMs allowed us to estimate the net effect of students' gender on their likelihood of being nominated as a candidate for the competition while controlling for factors from both individual (i.e., abilities) or relational interdependent processes (i.e., friendship).

Following Snijders and Baerveldt [2003], we performed a meta-analysis of the estimated ERGM parameters for each subject. This allowed us to estimate macro-level parameters across all sampled classes as weighted least square (WLS) means which, unlike a simple mean, model heterogeneity in standard error. We then used the t-ratio to test the statistical significance of estimated mean effects and estimate the net effect of gender on nominations across the sampled classes. More details on the meta-analysis are included in the Supplementary Material.

Missing data from the included classes were estimated via multivariate imputation by chained equations (see the Supplementary Material), based on the observed values for each case in the dataset and the relations observed for the other participants [Krause et al., 2020, Robins et al., 2004, van Buuren and Groothuis-Oudshoorn, 2011].

All ERG models were specified for the same parameter set. To test Hypothesis 1, we included a *gender receiver* parameter, whose coefficient indicated the likelihood that a female rather than a male student was nominated. To control for students' grades, we specified a *school grade* parameter, whose coefficient indicated the likelihood that students with higher grades were nominated more frequently than those with lower grades.

We also controlled for possible endogeneity regarding nominations by including the following structural parameters:

a) *edges*, i.e., the baseline likelihood of nominating at least one class member;

b) *reciprocity*, i.e., the likelihood that two students nominated each other;

c) *geometrically weighted indegree* (GWIDEGREE), i.e., the likelihood that a student was nominated as a function of the number of previously received nominations;

d) *geometrically weighted edgewise shared partners* (GWESP), i.e., the likelihood that student i nominated a classmate j if i also nominated at least another classmate k who, in turn, nominated j [Robins et al., 2009].

Finally, to control for the likelihood that students simply nominated friends, we specified a *friendship* parameter as a dyadic covariate. Following Grow et al. [2016], in previous model versions, we also specified a *gender homophily* parameter, i.e., the tendency to nominate more frequently classmates of the same gender. However, given that the effect of this parameter was not statistically significant and did not improve our models' fit, we decided to exclude it from our final models. Since students were asked to nominate four classmates, we set nodal outdegree to a maximum of 4. More detail on the model specifications and estimation is provided in the Supplementary Material.

Hypothesis 2

We modeled self-nominations of students as a binary variable and specified a fixed-effects logistic regression model. This allowed us to account for the clustered nature of data while minimizing bias due to the low number of clusters [McNeish and Kelley, 2019, McNeish and

Stapleton, 2016]. We estimated the model for all ten classes (n = 195 students). Missing data were estimated using multivariate imputation by chained equations, as described above. The model included *gender* as the main predictor. To control for intervening factors, we also included *grades* and *self-concept* in the specified subjects.

VI. Results

Table 1 shows the results of the meta-analysis of parameter estimates of ERG models, which allowed us to disentangle the net effect of gender from other factors on ability attribution among classmates. To test Hypotheses 1, we looked at the mean coefficient estimates of the *receiver gender* parameters. As regards the competition in reading (H1a), the estimated positive and statistically significant coefficient shows that female students were more likely to be nominated than male students, even controlling for their abilities in reading as measured by school grades and the confounding effect of other model parameters, thus confirming hypothesis H1a. Conversely, regarding math and science (H1b), the estimated coefficients show that male students were more likely to be nominated than female students. However, results were statistically significant only for science. Therefore, H1b was only partially confirmed, as we did not find any difference between female and male students in the case of math.

Table 1: *Meta-analysis results of ERG models*

	Reading competition	Math competition	Science competition
Receiver – gender	1.49***	-.43	-.33*
(female vs. male)	(.19)	(.21)	(.06)
Friendship ties	.99***	1.41***	.61***
	(.08)	(.13)	(.15)
Edges	-9.37***	-9.91***	-5.19***
	(.99)	(1.79)	(.57)
Reciprocity	-.67***	-.98**	-.51**
	(.16)	(.37)	(.15)
GWIDEGREE	-.64*	-.76	-1.25***
	(.31)	(.81)	(.24)
GWESP	.73***	.83***	.87***
	(.13)	(.19)	(.09)

Note. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001; two-tailed t-tests for all mean estimators except for gender (right-tailed for reading, left-tailed for math and science) and grade (right-tailed). Results on 7 networks for reading and math competition, on 6 networks for science competition due to convergence issues.

Unsurprisingly, our results showed that higher school grades in the subject and friendship generally increased the probability of nominations. Students with higher grades were nominated more frequently than those with lower grades and students preferred to nominate friends over other classmates. The structural parameters showed a typical situation found in other studies on status-related networks, as students tended not to reciprocate nominations, whereas nominations tended to be hierarchical, i.e., students tended to nominate others who had already been nominated by their nominations [e.g., Krackhardt, 1994]. Our models' goodness of fit is reported in the Supplementary Material.

Concerning Hypothesis 2, Table 2 shows the results of fixed-effect logistic regression models regarding the probability of self-nomination. Estimated coefficients showed that female students were less likely to nominate themselves than male students regardless of subjects, even while controlling for the effects of their school grades and self-concept. Unsurprisingly, students with higher grades had a higher probability of nominating themselves compared to those with lower grades, the same being true for students with a higher self-concept in the subject. H2 was therefore confirmed, as the observed gender-based differences in self-nominations were more likely due to a net effect of gender, despite differences in school grades and self-concept. Note that gender differences were higher for the math and science competitions, compared to the reading competition.

Table 2: *Results of Fixed-effects logistic regression*

	Reading competition	Math competition	Science competition
Gender (Female)	-1.33**	-1.49**	-2.17***
	(.47)	(.45)	(.49)
School grade	1.07***	.92***	.64**
	(.25)	(.23)	(.21)
Self-concept	1.02**	.36	.78**
	(.32)	(.24)	(.23)
N	195	195	195

Note. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001

We also tested whether this gender gap varied according to the grade and self-concept levels, by adding two interaction terms in the model, i.e., gender and grade, and gender and self-concept. However, estimated coefficients of these interactions did not provide clear evidence of any effect in either case. This would suggest that our data does not allow us to conclude that the gender gap was heterogeneous across levels of grade and self-concept. Not only did female students generally tend to shy away from competitions, their skills did not encourage them to engage in situations where they could prove their abilities. Indeed, the gender gap persisted also among top-performing students, even if they were fully entitled to participate in competitions.

VII. Discussion

As stated, we hypothesized a realistic situation in an academic setting (i.e., a competition between high-school students) to test the role of gender in making expectations on student's own and others' abilities in a stereotypical context (i.e., reading and mathematical/scientific skills), when information on abilities is typically available. Students were asked to choose four classmates (including themselves if they wanted), who in their opinion would have been the best candidate for the school competition. Consistent with the "status characteristic theory", we hypothesized that students would have formed expectations based on gender stereotypes even in cases where classmates' abilities contradicted the stereotypical attribution of different STEM and reading skills to men and women. Using Exponential Random Graph models, we tested whether female students were nominated less frequently for math and science competitions, and more frequently for the reading subject. We also performed a multi-level logistic regression analysis to check whether female students were less likely to nominate themselves for these competitions.

Note that two characteristics of the school system of our sample allowed us to assume that

students had information on classmates' abilities. First, students attended the same class with the same peers for four years, thereby establishing long-term relationships with countless occasions to observe each others' abilities. Second, final grades were publicly available to students.

Our findings partially confirmed our initial hypotheses. In regard to peers' nominations, we found that female students were less likely to be nominated for science and math, while they were more likely to be nominated for the reading competition. However, the difference was statistically significant only for the reading and science competition, while it could not be confirmed for the math competition. Results on peers' nominations would suggest that students in our sample may have still endorsed gender stereotypes on reading skills, and this affected their expectations on others' performance, even when they were aware of their abilities. The difference in findings between math and science, if not related to the sample, could confirm previous research showing that the scientific domain has now partially lost its masculine image [Passolunghi et al., 2014, Plante et al., 2009, Vuletich et al., 2020]. The weakening of stereotypes associating men with math should not surprise us, as this could be an effect of campaigns and initiatives introduced in many educational, public, and private organizations to reduce the under-representation of girls in STEM. These include, for instance, awareness of role models [Olsson and Martiny, 2018], public debate and discussion of the gaps within schools (UN International day of women and girl in science), and initiatives targeted at young women to increase their knowledge and interest in science (e.g., "Shecodes", "sciencegirlslab" in the U.S.).

Even given that the small size of our sample prevents us from making any generalization, our findings on the difference between math and science suggest that these two subjects should be distinguished whenever studying gender stereotypes. While research on stereotypes in school subjects has generally focused on math and has not investigated the association between gender and other scientific subjects, the lively debate on the under-representation of women and the leaky pipeline has considered women over the whole spectrum of STEM [Diekman et al., 2015]. There is a need for further research to explore these differences. For example, in summarizing research on the under-representation of women in STEM, Kahn and Ginther [2017] underlined that the gap is extended to the physical sciences and geosciences, while Ertl et al. [2017] and Blažev et al. [2017] included STEM and not only math in the instruments used to measure stereotypes. We believe that improving our understanding of the gender gap in scientific areas different from math is necessary if we want to enhance women's empowerment in STEM.

While our network analysis suggests a tendency to favor female students in reading and male students in science, we did not find a similar difference in the formation of self-expectations, as originally hypothesized. Indeed, in all three competitions, female students were less likely to nominate themselves compared to male students, and the difference was statistically significant even when controlling for grades at school and self-concept. The gap was greater for science and math but was also present for reading.

This finding could be explained by the general tendency of women to avoid competition [Gneezy et al., 2003] and suffer more from anxiety than men [Baraskar and Shinde, 2018], already found in previous studies. As emphasized by Niederle and Vesterlund [2010], there is general consensus on the fact that men and women have different attitudes toward competition. The internalization of these social and cultural pressures could explain why women are less keen to enter and win competitions due to a lack of over-confidence [Niederle and Vesterlund, 2011]. This depends on different perceptions of stress and anxiety typically associated with competition, which would prevent women from engaging in competitions even when they are fully qualified to win [Deaner et al., 2020]. This tendency is clear if we look at STEM competitions targeted specifically at students. As noted by Steegh et al. [2019], while in the U.S. participation rates of male and female students have almost reached parity, a relevant gap still exists at an international

level, where young women are still a minority, e.g., from 0% to 14% in mathematics, chemistry and physics Olympiads.

Finally, our study provides an example of the application of the SCT's assumptions on the information-combining process for performance expectations in an 'out-of- lab' setting. Assuming that in a school setting, teachers' evaluations are a good proxy of classmates' information on others' abilities, even when ability information is available, gender may still influence individual expectations on their own and others' performance in contexts in which gender stereotypes can be activated. Our findings would suggest that this mechanism also applies in a familiar environment, such as school, where everyone has full information about other classmates. Furthermore, young women seem to be influenced by stereotypes and low self-esteem even when they compare their abilities to those of acquainted peers at their same level.

Having noted this, our research does have certain limitations. First, our sample was neither randomly selected nor representative. While this imposes caution when interpreting our findings, we must recognize that it is hard to perform fine-grained, network data collection with standard randomized, representative samples. Another limitation is related to the unequal representation of female and male students within classes. As previously mentioned, students in our sample came from different types of high schools, some male-dominated, others more female-dominated. While these differences did not affect our measurement of gender bias, as all students studied reading, math, and science, they may have affected student nominations. However, our models did allow us to control for this heterogeneity.

Finally, gender differences in other factors, such as self-confidence and stress resilience, necessary to deal with a competitive context, could influence the gendered pattern of nominations. Our method of measuring gender bias could indeed reflect not only stereotypes on gender and abilities but also gender and adaptation to competition. The results of the network analysis would suggest that this was not the case, as we found a strong and large effect of gender on nominations for the reading competition, indicating that students did not seem to associate women with a lower capacity to face competitive situations. Here, experimental research either in the lab or in the field that considers the link between information, abilities, and expected performance in competitive settings could help to disentangle possible causal links between these factors [Baldassarri and Abascal, 2017]. Furthermore, the lab setting provides more control over available information which is impossible in fieldwork.

In conclusion, our findings suggest that there is still a need to intervene in settings in which gender stereotypes about women's and men's STEM skills can be activated. Furthermore, initiatives targeted at reducing gender stereotypes on math should be extended also to other scientific areas. Furthermore, gaps in STEM reflect only one side of the coin. The other is the belief that women are more inclined toward the Arts and Humanities. Our findings suggest that this association still seems to be strong. However, further studies are necessary to understand how these stereotypical associations affect choices and behaviors of young adults during their life course. In particular, given the importance of STEM sectors and jobs in the future economies and societies worldwide, the low self-confidence of young women could demotivate them from investing in these academic careers, with detrimental implications on gendered labor markets and traditional work-family division of labor. The gendered expectations of their peers could even exacerbate this demotivation, and negatively affect their STEM career intentions [Riegle-Crumb and Morton, 2017, Robnett and Leaper, 2013, van der Vleuten et al., 2018].

References

AlmaLaurea. XXII indagine profilo dei laureati, 2019. URL https://www.almalaurea.it/sites/

almalaurea.it/files/docs/universita/profilo/profilo2020/almalaurea_profilo_rapporto2020.pdf.

Delia Baldassarri and Maria Abascal. Field experiments across the social sciences. 43(1):41–73, 2017. ISSN 0360-0572, 1545-2115. doi: 10.1146/annurev-soc-073014-112445. URL http://www.annualreviews.org/doi/10.1146/annurev-soc-073014-112445.

Vandana Jaiprakash Baraskar and Pradeep Parshuram Shinde. To study the gender difference in personality factor and anxiety in adolescents – sociability self-confidence ambition and anxiety. 5(18):1458–1461, 2018. ISSN 23492562, 23492570. doi: 10.18410/jebmh/2018/305. URL https://jebmh.com/assets/data_pdf/Pradeep_P_Shinde_-_2_-_FINAL.pdf.

Joseph Berger, Bernard P. Cohen, and Morris Zelditch. Status characteristics and social interaction. 37(3):241–255, 1972. ISSN 0003-1224. doi: 10.2307/2093465. URL https://www.jstor.org/stable/2093465.

Mirta Blažev, Mia Karabegović, Josip Burušić, and Leila Selimbegović. Predicting gender-STEM stereotyped beliefs among boys and girls from prior school achievement and interest in STEM school subjects. 20(4):831–847, 2017. ISSN 1381-2890, 1573-1928. doi: 10.1007/s11218-017-9397-7. URL http://link.springer.com/10.1007/s11218-017-9397-7.

Bureau of Census. Employed persons by detailed occupation, sex, race, and hispanic or latino ethnicity, 2019. URL https://www.bls.gov/cps/tables.htm.

Thomas Buser, Muriel Niederle, and Hessel Oosterbeek. Gender, competitiveness and career choices, 2012. URL http://www.nber.org/papers/w18576.

Michela Carlana. Implicit stereotypes: Evidence from teachers' gender bias*. 134(3):1163–1224, 2019. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjz008. URL https://academic.oup.com/qje/article/134/3/1163/5368349.

Sapna Cheryan, Sianna A. Ziegler, Amanda K. Montoya, and Lily Jiang. Why are some STEM fields more gender balanced than others? 143(1):1–35, 2017. ISSN 1939-1455, 0033-2909. doi: 10.1037/bul0000052. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/bul0000052.

Nilanjana Dasgupta, Melissa McManus Scircle, and Matthew Hunsinger. Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. 112(16):4988–4993, 2015. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1422822112. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1422822112.

Robert O. Deaner, Amanda McClellan, Virgil Zeigler-Hill, and Joyce F. Benenson. Sex differences in exclusion and aggression on single-sex sports teams. 2020. ISSN 2330-2933, 2330-2925. doi: 10.1037/ebs0000247. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/ebs0000247.

Amanda B. Diekman, Erica S. Weisgram, and Aimee L. Belanger. New routes to recruiting and retaining women in STEM: Policy implications of a communal goal congruity perspective: Communal processes in STEM. 9(1):52–88, 2015. ISSN 17512395. doi: 10.1111/sipr.12010. URL http://doi.wiley.com/10.1111/sipr.12010.

DPR. Regolamento recante coordinamento delle norme vigenti per la valutazione degli alunni e ulteriori modalita' applicative in materia, ai sensi degli articoli 2 e 3 del decreto-legge 1 settembre 2008, n. 137, convertito, con modificazioni, dalla legge 30 ottobre 2008, n. 169. (09g0130), 2019. URL https://www.istruzione.it/esame_di_stato/Primo_Ciclo/normativa/allegati/dpr122_2009.pdf.

Alice H. Eagly and Valerie J. Steffen. Gender stereotypes stem from the distribution of women and men into social roles. 46(4):735–754, 1984. ISSN 1939-1315(Electronic),0022-3514(Print). doi: 10.1037/0022-3514.46.4.735.

Naomi Ellemers. Gender stereotypes. 69(1):275–298, 2018. ISSN 0066-4308, 1545-2085. doi: 10.1146/annurev-psych-122216-011719. URL http://www.annualreviews.org/doi/10.1146/annurev-psych-122216-011719.

Bernhard Ertl, Silke Luttenberger, and Manuela Paechter. The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. 8: 703, 2017. ISSN 1664-1078. doi: 10.3389/fpsyg.2017.00703. URL http://journal.frontiersin.org/article/10.3389/fpsyg.2017.00703/full.

Eurostat. Women in science and technology, 2019. URL https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20190211-1.

Susan R. Fisk. Who's on top? gender differences in risk-taking produce unequal outcomes for high-ability women and men. 81(3):185–206, 2018. ISSN 0190-2725, 1939-8999. doi: 10.1177/0190272518796512. URL http://journals.sagepub.com/doi/10.1177/0190272518796512.

Margaret Foddy and Michael Smithson. Can gender inequalities be eliminated? 62(4):307–324, 1999. doi: 10.2307/2695831.

Silvia Galdi, Alberto Mirisola, and Carlo Tomasetto. On the relations between parents' and children's implicit and explicit academic gender stereotypes. (2):215–238, 2017. ISSN 1827-2517. doi: 10.1482/87248. URL https://doi.org/10.1482/87248.

Sara Geven, Jan O. Jonsson, and Frank van Tubergen. Gender differences in resistance to schooling: The role of dynamic peer-influence and selection processes. 46(12):2421–2445, 2017. ISSN 0047-2891, 1573-6601. doi: 10.1007/s10964-017-0696-2. URL http://link.springer.com/10.1007/s10964-017-0696-2.

Patricia N. Gilbert. The role of role models: How does identification with STEM role models impact women's implicit STEM stereotypes and STEM outcomes?, 2015. URL https://digitallibrary.tulane.edu/islandora/object/tulane%3A27945.

U. Gneezy, M. Niederle, and A. Rustichini. Performance in competitive environments: Gender differences. 118(3):1049–1074, 2003. ISSN 0033-5533, 1531-4650. doi: 10.1162/00335530360698496. URL https://academic.oup.com/qje/article-lookup/doi/10.1162/00335530360698496.

GPDP. La privacy tra i banchi di scuola, 2010. URL https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/2416449.

Anthony G. Greenwald, Debbie E. McGhee, and Jordan L. K. Schwartz. Measuring individual differences in implicit cognition: The implicit association test. 74(6):1464–1480, 1998. ISSN 1939-1315(Electronic),0022-3514(Print). doi: 10.1037/0022-3514.74.6.1464.

André Grow, Károly Takács, and Judit Pál. Status characteristics and ability attributions in hungarian school classes: An exponential random graph approach. 79(2):156–167, 2016. ISSN 0190-2725, 1939-8999. doi: 10.1177/0190272516643052. URL http://journals.sagepub.com/doi/10.1177/0190272516643052.

Christina Günther, Neslihan Arslan Ekinci, Christiane Schwieren, and Martin Strobel. Women can't jump?—an experiment on competitive attitudes and stereotype threat. 75(3):395–401, 2010. ISSN 01672681. doi: 10.1016/j.jebo.2010.05.003. URL https://linkinghub.elsevier.com/retrieve/pii/S0167268110000855.

Tanja Hentschel, Madeline E. Heilman, and Claudia V. Peus. The multiple dimensions of gender stereotypes: A current look at men's and women's characterizations of others and themselves. 10:11, 2019. ISSN 1664-1078. doi: 10.3389/fpsyg.2019.00011. URL https://www.frontiersin.org/article/10.3389/fpsyg.2019.00011/full.

Sarah I. Hofer. Studying gender bias in physics grading: The role of teaching experience and country. 37(17):2879–2905, 2015. ISSN 0950-0693, 1464-5289. doi: 10.1080/09500693.2015.1114190. URL http://www.tandfonline.com/doi/full/10.1080/09500693.2015.1114190.

Shulamit Kahn and Donna Ginther. Women and STEM, 2017. URL http://www.nber.org/papers/w23525.pdf.

Peter L. Kench, Nikki Field, Maila Agudera, and Margo Gill. Peer assessment of individual contributions to a group project: Student perceptions. 15(2):158–165, 2009. ISSN 10788174. doi: 10.1016/j.radi.2008.04.004. URL https://linkinghub.elsevier.com/retrieve/pii/S1078817408000321.

Alyssa J. Kersey, Kelsey D. Csumitta, and Jessica F. Cantlon. Gender similarities in the brain during mathematics development. 4(1):19, 2019. ISSN 2056-7936. doi: 10.1038/s41539-019-0057-x. URL http://www.nature.com/articles/s41539-019-0057-x.

Martin Kilduff and David Krackhardt. *Interpersonal networks in organizations: cognition, personality, dynamics, and culture*. Number 30 in Structural analysis in the social sciences. Cambridge University Press, 2008. ISBN 978-0-521-86660-6 978-0-521-68558-0. OCLC: ocn191206908.

Dorottya Kisfalusi, Béla Janky, and Károly Takács. Double standards or social identity? the role of gender and ethnicity in ability perceptions in the classroom. 39(5):745–780, 2019. ISSN 0272-4316, 1552-5449. doi: 10.1177/0272431618791278. URL http://journals.sagepub.com/doi/10.1177/0272431618791278.

David Krackhardt. Graph theoretical dimensions of informal organization. 89, 1994.

Robert W. Krause, Mark Huisman, Christian Steglich, and Tom A. B. Snijders. Missing data in cross-sectional networks - an extensive comparison of missing data treatment methods. 62: 99–112, 2020.

Dean Lusher, Johan Koskinen, and Garry Robins, editors. *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press, 2012. ISBN 978-0-511-89470-1. doi: 10.1017/CBO9780511894701. URL http://ebooks.cambridge.org/ref/id/CBO9780511894701.

Daniel McNeish and Ken Kelley. Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. 24(1):20–35, 2019. ISSN 1939-1463, 1082-989X. doi: 10.1037/met0000182. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/met0000182.

Daniel McNeish and Laura M. Stapleton. Modeling clustered data with very few clusters. 51 (4):495–518, 2016. ISSN 0027-3171, 1532-7906. doi: 10.1080/00273171.2016.1167008. URL https://www.tandfonline.com/doi/full/10.1080/00273171.2016.1167008.

Muriel Niederle and Lise Vesterlund. Explaining the gender gap in math test scores: The role of competition. 24(2):129–144, 2010. URL https://ideas.repec.org/a/aea/jecper/v24y2010i2p129-44.html.

Muriel Niederle and Lise Vesterlund. Gender and competition. 3(1):601–630, 2011. ISSN 1941-1383, 1941-1391. doi: 10.1146/annurev-economics-111809-125122. URL http://www.annualreviews.org/doi/10.1146/annurev-economics-111809-125122.

OECD. *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. PISA. OECD, 2015. ISBN 978-92-64-23002-6 978-92-64-22994-5. doi: 10.1787/9789264229945-en. URL https://www.oecd-ilibrary.org/education/the-abc-of-gender-equality-in-education_9789264229945-en.

Maria Olsson and Sarah E. Martiny. Does exposure to counterstereotypical role models influence girls' and women's gender stereotypes and career choices? a review of social psychological research. 9, 2018. ISSN 1664-1078. doi: 10.3389/fpsyg.2018.02264. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6292925/.

Elizabeth Levy Paluck. Peer pressure against prejudice: A high school field experiment examining social network change. 47(2):350–358, 2011. ISSN 1096-0465(Electronic),0022-1031(Print). doi: 10.1016/j.jesp.2010.11.017.

Maria Chiara Passolunghi, Tania Irene Rueda Ferreira, and Carlo Tomasetto. Math–gender stereotypes and math-related beliefs in childhood and early adolescence. 34:70–76, 2014. ISSN 10416080. doi: 10.1016/j.lindif.2014.05.005. URL https://linkinghub.elsevier.com/retrieve/pii/S1041608014000983.

Isabelle Plante, Manon Théorêt, and Olga Eizner Favreau. Student gender stereotypes: contrasting the perceived maleness and femaleness of mathematics and language. 29(4):385–405, 2009. ISSN 0144-3410, 1469-5820. doi: 10.1080/01443410902971500. URL http://www.tandfonline.com/doi/full/10.1080/01443410902971500.

Tomáš Protivínský and Daniel Münich. Gender bias in teachers' grading: What is in the grade. 59: 141–149, 2018. ISSN 0191491X. doi: 10.1016/j.stueduc.2018.07.006. URL https://linkinghub.elsevier.com/retrieve/pii/S0191491X17302584.

M. D. Pugh and Ralph Wahrman. Neutralizing sexism in mixed-sex groups: Do women have to be better than men? 88(4):746–762, 1983. ISSN 00029602, 15375390. URL http://www.jstor.org/stable/2779483. Publisher: University of Chicago Press.

Cecilia L. Ridgeway and Shelley J. Correll. Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. 18(4):510–531, 2004. ISSN 0891-2432, 1552-3977. doi: 10.1177/0891243204265269. URL http://journals.sagepub.com/doi/10.1177/0891243204265269.

Catherine Riegle-Crumb and Karisma Morton. Gendered expectations: Examining how peers shape female students' intent to pursue STEM fields. 8, 2017. ISSN 1664-1078. doi: 10.3389/fpsyg.2017.00329. URL http://journal.frontiersin.org/article/10.3389/fpsyg.2017.00329/full.

Garry Robins, Philippa Pattison, and Jodie Woolcock. Missing data in networks: exponential random graph (p) models for networks with non-respondents. 26(3):257–283, 2004. ISSN 03788733. doi: 10.1016/j.socnet.2004.05.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0378873304000280.

Garry Robins, Pip Pattison, Yuval Kalish, and Dean Lusher. An introduction to exponential random graph (p*) models for social networks. 29(2):173–191, 2007. ISSN 03788733. doi: 10.1016/j.socnet.2006.08.002. URL https://linkinghub.elsevier.com/retrieve/pii/S0378873306000372.

Garry Robins, Pip Pattison, and Peng Wang. Closure, connectivity and degree distributions: Exponential random graph (p*) models for directed social networks. 31(2):105–117, 2009. ISSN 03788733. doi: 10.1016/j.socnet.2008.10.006. URL https://linkinghub.elsevier.com/retrieve/pii/S0378873308000592.

Rachael D. Robnett and Campbell Leaper. Friendship groups, personal motivation, and gender in relation to high school students' STEM career interest. 23(4):652–664, 2013. ISSN 1532-7795. doi: 10.1111/jora.12013. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jora.12013.

Hana Shepherd and Elizabeth Levy Paluck. Stopping the drama: Gendered influence in a network field experiment. 78(2):173–193, 2015. ISSN 0190-2725, 1939-8999. doi: 10.1177/0190272515581212. URL http://journals.sagepub.com/doi/10.1177/0190272515581212.

T. A. B. Snijders and Chris Baerveldt. A multilevel network study of the effects of delinquent behavior on friendship evolution. 27(2):123–151, 2003. ISSN 0022-250X, 1545-5874. doi: 10.1080/00222500305892. URL http://www.tandfonline.com/doi/abs/10.1080/00222500305892.

Margaret Goddard Spear. Sex bias in science teachers' ratings of work and pupil characteristics‡. 6(4):369–377, 1984. ISSN 0140-5284. doi: 10.1080/0140528840060407. URL http://www.tandfonline.com/doi/abs/10.1080/0140528840060407.

Anneke M. Steegh, Tim N. Höffler, Melanie M. Keller, and Ilka Parchmann. Gender differences in mathematics and science competitions: A systematic review. 56(10):1431–1460, 2019. ISSN 0022-4308, 1098-2736. doi: 10.1002/tea.21580. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/tea.21580.

Melanie C. Steffens and Petra Jelenec. Separating implicit gender stereotypes regarding math and language: Implicit ability stereotypes are self-serving for boys and men, but not for girls and women. 64(5):324–335, 2011. ISSN 0360-0025, 1573-2762. doi: 10.1007/s11199-010-9924-x. URL http://link.springer.com/10.1007/s11199-010-9924-x.

Carlo Tomasetto, Silvia Galdi, and Mara Cadinu. Quando l'implicito precede l'esplicito: gli stereotipi di genere sulla matematica in bambine e bambini di 6 anni. (2):169–186, 2012. ISSN 1827-2517. doi: 10.1482/37693. URL https://doi.org/10.1482/37693.

Carlo Tomasetto, Alberto Mirisola, Silvia Galdi, and Mara Cadinu. Parents' math–gender stereotypes, children's self-perception of ability, and children's appraisal of parents' evaluations in 6-year-olds. 42:186–198, 2015. ISSN 0361476X. doi: 10.1016/j.cedpsych.2015.06.007. URL https://linkinghub.elsevier.com/retrieve/pii/S0361476X15000338.

Stef van Buuren and Karin Groothuis-Oudshoorn. Multivariate imputation by chained equations in r. 45(3):1–67, 2011. doi: https://doi.org/10.18637/jss.v045.i03.

Maaike van der Vleuten, Stephanie Steinmetz, and Herman van de Werfhorst. Gender norms and STEM: the importance of friends for stopping leakage from the STEM pipeline. 24(6):417–436, 2018. ISSN 1380-3611, 1744-4187. doi: 10.1080/13803611.2019.1589525. URL https://www.tandfonline.com/doi/full/10.1080/13803611.2019.1589525.

Daniel Voyer and Susan D. Voyer. Gender differences in scholastic achievement: A meta-analysis. 140(4):1174–1204, 2014. ISSN 1939-1455, 0033-2909. doi: 10.1037/a0036620. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/a0036620.

Heidi A. Vuletich, Beth Kurtz-Costes, Erin Cooley, and B. Keith Payne. Math and language gender stereotypes: Age and gender differences in implicit biases and explicit beliefs. 15(9):e0238230, 2020. ISSN 1932-6203. doi: 10.1371/journal.pone.0238230. URL https://dx.plos.org/10.1371/journal.pone.0238230.

David G. Wagner and Joseph Berger. Gender and interpersonal task behaviors: Status expectation accounts. 40(1):1–32, 1997. ISSN 0731-1214, 1533-8673. doi: 10.2307/1389491. URL http://journals.sagepub.com/doi/10.2307/1389491.

David G. Wagner, Rebecca Ford, and Thomas Ford. Can gender inequalities be reduced? 51(1): 47–61. doi: 10.2307/2095477.