Sex Roles (2022) 86:587–603
https://doi.org/10.1007/s11199-022-01289-y

ORIGINAL ARTICLE

Is Emma or Liam the Top Scorer in Math? The Effects of a Counter-Stereotypical Role Model on Math Achievement

Nadia Leroy[1] · **Sylvain Max**[2] · **Pascal Pansu**[1]

Accepted: 7 April 2022 / Published online: 21 April 2022

Abstract
From 2015 to 2018, the math gender gap decreased primarily due to a decline in boys' performances (OECD, 2015, 2016, 2018). However, there is ample evidence that girls continue to be negatively stereotyped in math. Using a longitudinal design, we examined whether prolonged exposure to a counter-stereotypical role model embodied by a female top math scorer may prevent other girls in the class from experiencing stereotype threat. Multilevel analyses were conducted among 1,043 6th graders nested in 46 math classes. There was a decline in math performance throughout the school year for all students, but being a girl had a buffering effect against this decline. The results failed to support the main effect hypothesis (H1) which anticipated that student gender and top math scorer gender would be associated with student math achievement when jointly considered. The results supported the cross-level interaction hypothesis (H2) which anticipated that the greatest benefits would emerge for girls exposed to a counter-stereotypical role model; that is, in a class whose top math scorer was a girl. These results offer new insights regarding the extent to which a counter-stereotypical role model embodied by the top math scorer may influence differences in math performances.

Keywords Gender differences · Mathematics achievement · Role models · Stereotypes · Multilevel analysis

The gender gap in math performance is a recurring problem for education professionals. This gender gap can have long-term consequences for girls' future lives. Indeed, there is a clear male dominance in the most highly paid positions in STEM fields (Betz & Sekaquaptewa, 2012; Lockwood, 2006; O'Brien et al., 2017). According to the literature, this math gender gap first appears in the early years of primary school, and then increases over the high school years (Fryer & Levitt, 2010; Hyde & Mertz, 2009; Hyde et al., 1990a, b; Lindberg et al., 2010; Spelke, 2005). However, over the past decade, many countries have made significant progress in reducing the gender achievement gap. Several studies show a decrease in the differences between boys' and girls' mathematics performance (Halpern et al., 2007; Hedges & Nowell, 1995; Hyde et al., 1990a, 2008; Lindberg et al., 2010). Meta-analyses reveal that the assumption of gender differences in math is not always true (Ghasemi et al., 2019). For instance, Else-Quest et al. (2010) analysis of TIMSS 2003 data indicates cross-national variations in the magnitude of the gender gap in math achievement, with an overall effect size (d) for eighth graders in 46 countries of –.01. The findings are also comparable to those of Hyde et al. (2008) who reported d values of between –.01 and –.02 for fourth and eighth graders respectively.

More recently, in 2015, the average gap between boys' and girls' math performances in OECD countries was reportedly 8 points. The PISA 2018 results showed that boys continue to outperform girls, but that the gap was down to 5 points. Whilst a gender gap in math achievement seems to persist in some nations but not in others, Hyde et al. (1990a, b) meta-analytic studies demonstrate that boys report significantly more self-confidence in math ($d = .25$). These findings are in line with the gender similarities hypothesis

✉ Nadia Leroy
nadia.leroy@univ-grenoble-alpes.fr

Sylvain Max
sylvain.max@bsb-education.com

Pascal Pansu
pascal.pansu@univ-grenoble-alpes.fr

[1] LaRAC - Laboratory of Research On Acquisition in Context, Université Grenoble Alpes, 1251 Avenue centrale, BP 47, 38040Cedex 9 Grenoble, France

[2] LESSAC & CEREN, EA 7477, Burgundy School of Business, Université Bourgogne Franche-Comté, 29 rue Sambin, 21000 Dijon, France

(Cantley & McAllister, 2021), which maintains that men and women are similar in most, but not all, psychological variables. Consistent with Steele's theory (Steele, 1997), the gender stereotype about gender roles and math may discourage girls from feeling confident or approaching achievement situations differently (Roberts, 1991). In other words, despite evidence of gender similarities in math achievement (Hyde et al., 2008), the negative stereotype that girls lack ability in math persists (Hedges & Nowell, 1995; Hyde et al., 1990a, b), suggesting that gender disparities would emerge from societal stereotypes.

Explanations for the Gender Gap in Math Performance

Research has linked the gender gap in math performance to socio-cultural environments, relating in particular to the ways boys and girls are socialized (e.g., Hadjar et al., 2014). Due to the way in which teachers and parents interact with them, young girls may also be discouraged from entering STEM fields (Beede et al., 2011; Else-Quest et al., 2010; Good et al., 2008; Guiso et al., 2008; Helwig et al., 2001; Hyde & Mertz, 2009; Inzlicht & Ben-Zeev, 2000, 2003; Jacobs & Eccles, 1992; Nollenberger et al., 2016; Rattan et al., 2012; Rodríguez-Planas & Nollenberger, 2018; Spencer et al., 1999; Spinath & Spinath, 2005; Steele, 1997; Tiedemann, 2000; Tomasetto & Appoloni, 2013). This phenomenon is known as stereotype threat (Steele, 1997; Steele & Aronson, 1995) and describes the risk of a person being reduced to a stereotyped trait associated with the group to which the person belongs. Stereotype threat has a negative impact on the ability of an individual to fully utilize their skills in an assessment relating to the negatively stereotyped field (Steele, 1997). For example, the stereotype may be triggered in memory (Beilock et al., 2007; Schmader & Johns, 2003). In this case, the fear of confirming the negative gender stereotype can interfere with the accomplishment of the task (Schmader et al., 2008). Furthermore, the negatively stereotyped members of a group fear that they may confirm the negative stereotype in question (Steele, 1997; Steele & Aronson, 1995). Numerous studies (e.g., Good et al., 2008; Spencer et al., 1999) have shown that this stereotype threat, which implicitly or explicitly activates gender identity, can ultimately affect girls' engagement in STEM-related courses, thereby limiting their choice of career (Marx & Roman, 2002; Rattan et al., 2012).

Here, girls' risk being judged as less capable than boys in a predominantly male field such as mathematics, which activates stereotype threat. Research has shown that the negative stereotype for girls in mathematics impacts upon their self-confidence and performance (Régner et al., 2014). Shih et al. (1999) showed that when doing a math test, young Asian women achieved lower grades when the fact that they were female was pointed out to them. However, their performance improved when their ethnic group was emphasized because the stereotype of Asian math ability is a positive one. This study is relevant because it shows that the girls' performance varies according to which aspect of their social identity is activated. A large body of stereotype threat research has confirmed that girls can be induced to think about the stereotypes that target them, and in some cases, they confirm these stereotypes (Régner et al., 2014; Spencer et al., 1999).

It has also been shown that stereotype threat influences girls during the learning process which precedes the test (Rydell & Boucher, 2017; Rydell et al., 2010; Taylor & Walton, 2011). This negative effect is no longer present when there is only one social group, as has been demonstrated in studies where only girls took a math test (Huguet & Régner, 2009; Inzlicht & Ben-Zeev, 2000). To counteract these negative stereotype threat effects, many studies have been conducted with the aim of encouraging women to access STEM fields, some of which have used female role models as a means to increase the sense of belonging to STEM fields (Bagès & Martinot, 2011; Bertrand & Duflo, 2017; Lockwood, 2006; Plant et al., 2009; Shin et al., 2016; Stout et al., 2011; Weisgram & Bigler, 2007).

Interventions to Reduce the Gender Gap and Stereotype Threat

Numerous interventions have been developed to counteract stereotype threat among women in mathematics (Kit et al., 2008). The simplicity of exposure to a role model renders this type of intervention more attractive than other complex interventions (e.g., growth mindset). There are two influential role model theories in the literature (Ahn et al., 2020): the stereotype inoculation model (Dasgupta, 2011) and the motivational theory of role modeling (Morgenroth et al., 2015). According to Dasgupta (2011), contact with in-group role models functions as a "social vaccine" that inoculates minority-group members against stereotype threat. According to Morgenroth et al. (2015), role models serve three distinct functions: behavioral models, representing what is achievable, and being inspirational. Theoretically, the role model acts as an upward identification target (Lockwood & Kunda, 1997). In other words, role models are defined as performers whose behavior conveys information to observers about the characteristics of the appropriate responses which should be displayed in a certain situation (Bandura, 1972).

The premise upon which these studies are based is that observers (or "role aspirants") internalize stereotypical knowledge about gender roles and act, accordingly, resulting in gender-conforming aspirations and behaviors. Because role models illustrate the specific goals, behaviors, and strategies that the role aspirants internalize and mimic, they can be usefully deployed in the field of education to strengthen student motivation. In addition, according to theories focusing on the development of gender roles (Eagly & Wood, 2011), people perceive certain roles as more or less in alignment with their gender, which is why observing men and women in gender roles promotes gender aspirations and behaviors.

Role models function equally well for stereotypes and counter-stereotypes. Observing or interacting with men and women in roles which are not commonly assigned to them (e.g., a female scientist, or a male preschool teacher) provides a counter-stereotypical role model. Frequent exposure to counter-stereotypical role models is reported to reduce gender stereotypes and promotes non-traditional behavior (Olsson & Martiny, 2018). Many interventions involving observation or interaction with counter-stereotypical gender role models have been implemented. Initiatives that aim to promote the entry of women into fields where they are under-represented and negatively stereotyped have moved in this direction and have been based on the idea that the exposure of girls to female counter-stereotypical role models can increase a sense of belonging to STEM fields (Dasgupta, 2011) as well as reinforcing the belief that hard work is the road to success in STEM fields (Shin et al., 2016).

These studies rely on methods of reducing negative stereotype threat effects in laboratory environments (e.g., Good et al., 2008; Spencer et al., 1999). Just as there are situations that promote the activation of negative stereotypes - such as women taking a math assessment in a room full of men - researchers have found that there are also situations that remove or reduce this type of threat and lead to improved performance. The literature on role models in adolescence and adulthood provides numerous examples of laboratory studies, which generally give female university students information about successful women in fields in which they are under-represented and negatively stereotyped. These studies involve the exposure of students to literature, biographies, or advertisements depicting women in counter-stereotypical roles, and their results show that exposure to counter-stereotypical role models influences girls' gender beliefs and their traditional attitudes towards women. Mcintyre et al. (2003) showed that women can overcome stereotypes and improve their performance in mathematics when reminded of the achievement of other women. Other field-based studies have assessed the effect of interacting with counter-stereotypical female role models and have shown that a role model who demonstrates skill in a negatively stereotyped area for her/his ingroup (e.g., math for girls) can influence other members of the group to attach importance to performance in the domain in question. The role model shows that it is possible for them to be successful, and this has a direct impact on their aspirations and choices (Wigfield & Eccles, 2000). A large body of work has established that science teachers can also serve as role models and help to improve girls' academic performance (Hoffmann & Oreopoulos, 2009; Paredes, 2014). Finally, a female student who is the top math scorer in a class could also serve as a role model, although as far as we are aware, this hypothesis is yet to be tested.

While most of these studies showed a buffering effect of in-group counter-stereotypical role models from stereotype threat, some studies nonetheless report no buffering effect, which has been linked to several factors that alter the effectiveness of the inspiration that role models provide to role aspirants (Lockwood & Kunda, 1997). First, an effective role model must demonstrate competence and achievable success in the desired or relevant area. This dimension is important because role models are people who demonstrate the skills and motivation that role aspirants lack (Marx & Ko, 2012; Marx & Roman, 2002). An effective role model is someone with whom role aspirants can identify (Lockwood & Kunda, 1997; Marx & Ko, 2012). Second, similarity is essential because upward identification is stronger when perceived shared similarities exist (Collins, 1996, 2000). Third, the achievements of the role model must be perceived as attainable (e.g., Hoyt & Simon, 2011; Lockwood & Kunda, 1997, 1999). Attainability influences the role aspirant's expectation of success.

Finally, brevity of exposure to the role model (Bigler & Liben, 1990; Frawley, 2008) has been suggested as an explanatory factor for its short-lived effect (Dasgupta & Asgari, 2004; Olsson & Martiny, 2018; Rosenberg-Kima et al., 2008; Savenye, 1990). Research protocols based on brief exposure can be criticized for being poorly representative of what children experience in their daily lives at school. It has been argued that longitudinal interventions and prolonged exposure to role models would be more effective in bringing about internalized changes. Indeed, Nhundu (2007) found that elementary school girls who had been exposed to counter-stereotypical material over a 3-year period expressed greater aspirations to pursue a non-traditional career than girls who had been exposed to traditional educational materials. It is particularly important to study and increase contact with role models in the early years of training, and in transition periods (Dasgupta, 2011).

Study Aims and Hypotheses

Based on the contradictory findings on gender similarities in math achievement and the potential effect of implicit and explicit cultural biases such as gender stereotypes, the present study aimed to explore the effects of top math scorer gender on the direction and magnitude of gender differences. To date, no study has used a longitudinal design to examine the impact of role models (i.e., a female or male top math scorer) in an ordinary class where the negative stereotype may be implicitly activated. In this paper, we specifically examined whether the role model of a female top math scorer, versus a male, reduced the gender gap favoring boys over time, assuming that girls are generally negatively stereotyped in mathematics. Examining the effect of top math scorer role model in a class is highly relevant for several reasons. First, the top math scorer is a peer with whom other students regularly interact with in the classroom, and personal contact has been demonstrated to be a powerful way to find role models (Dasgupta & Asgari, 2004). Second, similarity matters, and the dimensions of similarity can be as varied as "common academic or professional interests, similar life history, shared group membership, similar goal orientations, etc." (Dasgupta, 2011, p. 236). We assumed that belonging to the same class increases perceived similarity. Third, it seems likely that interactions with role models over a prolonged period of time will have greater effects than brief exposure. Thus, it appears that exposure over the course of a school year to role models has the potential to improve the effects reported by studies on short-term exposure. This is particularly relevant during the early stages of secondary school as there is a consensus that girls' gradual dropout from some STEM fields begins after the age of 12 (Sáinz & Eccles, 2012) given the predisposition of girls to underestimate their ability to succeed in those fields (Correll, 2001; Sáinz & Eccles, 2012). Furthermore, Bagès and Martinot (2011) demonstrated in classrooms that a slightly upward comparison target can be inspirational and lead students to academic progress when they feel like the role model. In the same paper, they answer the question "who is the best role model for children in math?" (p. 2). Same-gender role model effects on math performance are moderated by the explanation given to the role model's math success. When no explanation is given, students perform better with a female (versus a male) role model.

In line with these ideas, we made two hypotheses. The additive hypothesis (*H1*) proposes that the joint effect of student gender and top math scorer gender is equal to the sum of their unique contributions. Given the decline in math performance for high school students, especially for boys (Wijsman et al., 2016), we expected a buffering effect of student gender on math performance, especially for girls, regardless of whether they are members of a classroom in which the top math scorer is a girl or a boy. The additive model is one in which student gender at Level 1 is included as a main effect on math performance, and top math scorer gender is added as a Level 2 main effect without any interaction term. It assumes that the effect of top math scorer gender is not conditional on the effect of student gender. If this hypothesis is confirmed, then student gender (Level 1) and top math scorer gender (Level 2) predictors will both be significantly associated with math achievement when assessed in a multilevel model.

The synergistic hypothesis (*H2*) proposes that the joint effect of student gender and top math scorer gender goes above and beyond the sum of their individual effects. Student gender and top math scorer gender combine to have a multiplicative effect, such that the effect of the top math scorer gender on math achievement is contingent upon student gender. Testing this hypothesis implies that a cross-level interaction term will be entered to examine the different effects of top math scorer gender on math performance depending on student gender. If this hypothesis is confirmed, then math achievement will be the highest for girls in math classes where the top math scorer is a girl.

Method

This study is non-interventional research, and the approval of the French state authorities was not required. However, the research team obtained consent from the relevant school authorities and from parents.

Participants

A total of 1043 first-year junior high school students from 46 classes in 15 junior high schools in France participated in this study. The students were evenly distributed across gender (50.6% boys ($n=528$) and 49.4% girls ($n=515$), and their age ranged between 9 and 14 years old ($M=11.20$, $SD=.50$). Regarding their academic progression, 75.74% ($n=790$) of students were "on track," 17.96% had repeated a year ($n=187$), 3.8% ($n=40$) had skipped a year (data were not available for 2.5% ($n=26$) of the sample). The socioeconomic status of the students' families represented a diverse middle-class range from semiskilled workers through to professionals. The teachers of the classes included 6 men and 24 women with an average seniority of 22 years ($SD=11.20$, $Min=2$ years, $Max=41$ years). Among the 46 math classes, 10 were taught by a male teacher and 36 by a female teacher. All the schools were state schools and 33% ($n=5$) of them were in Priority Education Zones, which were created to restore equality of opportunity amongst students in urban areas with social difficulties, through financial assistance.

Measures

Math Scores

The students' baseline levels in math were obtained from their scores on the national math test which takes place at the beginning of the school year for all French students entering junior high school ($M=13.43$, $SD=3.47$). This is a standardized test that was developed by the French Ministry of Education. It is scored out of 20, and it covers geometry, algebra, and calculus. Since students do not take another standardized math test at the end of the first year of junior high school, we developed a test to assess their performance at this time. France has a national curriculum, so all the students had covered the same program. The test we developed with math teachers contained geometry, algebra, and calculus items, and was based on the topics covered during the school year (score out of 20; $M = 10.97$, $SD = 4.61$; $\alpha = .86$).

Identification of the Top Math Scorer

To identify the top math scorer in each class, we collected the three terms averages for each student. These three math scores were then used to establish an annual average to find out whether the top math scorer was a girl or a boy.

Top Math Scorer Gender and Student Gender

Top math scorer gender and student gender were coded 1 for girls and 0 for boys.

Procedure

The school principals, math teachers, and parents all received a letter informing them of the research project and had to return a consent form allowing participation. All the participants were assured that the data would remain anonymous and confidential. Students were invited to answer a questionnaire at the beginning of the year and took a mathematics test at the end of the school year. At each time point, the researcher went to each classroom to administer the questionnaire for one hour during a regular class session. Oral and written instructions were given to help students understand the questionnaire. Students' data were matched across time point by using a pseudo anonymized identification code generated from the date of birth and a code assigned to each class.

Data Analysis Plan

As these data have a multilevel structure where students (level-1) are nested within classes (level-2), we used the multilevel model technique (Bryk & Raudenbush, 2002; Goldstein, 2011) to separate within-group individual effects from between-group aggregate effects. First, we fitted the simplest possible multilevel model which is a random intercept unconditional model (empty model). Model A is equivalent to a one-way ANOVA with a random effect. It is useful as a preliminary step in multilevel data analysis as it provides information about math achievement variability at each level. We used this step to investigate if multilevel modeling was more appropriate than ordinary least squares regression. The intraclass correlation coefficient (*ICC*), which gives information about the proportion of interclass (i.e., between) variance *versus* the total variance, is useful in determining whether multilevel models were required.

Secondly, we fitted a series of conditional models including the random effect of the clustering variable in which we conditioned math achievement on predictors at level-1 and predictors at level-2. See Supplement A in the online supplement for a taxonomy of these multilevel models fitted to address each step of this analytical sequence. See Supplement B in the online supplement for definitions and interpretations of the parameters in the fitted multilevel models. In Model B, the independent effect of math performance at the beginning of the year on math performance at the end of the year was introduced as a control variable. Next, the independent effect of teacher gender, which is a level-2 characteristic of class, was investigated in Model C controlling for the effect of math performance at the beginning of the year. The independent effect of level-1 predictor student gender and level-2 predictor top math scorer gender on math performance at the end of the year were finally investigated through a series of models aiming to test the additive hypothesis (*H1*) and synergistic hypothesis (*H2*).

More specifically, in Model D we investigated the effect of student gender in which the intercept from the composite model (γ_{00}) is defined as the expected math performance of a student whose gender is zero (i.e., a boy) and whose initial performance in math is zero. γ_{10} represents the expected change in math achievement associated with 1-unit increase in math score at the beginning of the year. γ_{20} is the expected change in math achievement associated with student gender whose value is 1 (i.e., a girl). In Model E, we added a level-2 predictor to test the additive hypothesis (*H1*) of both student gender and top math scorer gender predictors on math performance at the end of the year. This model contains the same fixed effects and random effects as model D but includes a supplementary fixed effect (γ_{02}) representing the expected change in math performance at the end of the year associated with top math scorer gender whose value is 1 (i.e., a girl).

In Model F, we introduced a level-2 random effect on the slope of student gender to test whether its effect on math performance at the end of the year varies across classes. Finally, we fitted Model G in which we considered the effect of the level-2 predictor top math scorer gender on the slope of the level-1 predictor student gender. This model implies a cross-level interaction effect (γ_{22}) which tests whether the effect of student gender on math performance at the end of the year depends on top math scorer gender; that is, whether there is a synergistic effect (*H2*) between both variables to predict math performance at the end of the year. This synergistic effect was investigated for the following four (2×2) conditions: boys in a male top math scorer class (the reference condition); boys in a female top math scorer class; girls in a male top math scorer class; and girls in a female top math scorer class.

To quantify how much the math performance at the end of the year variation is explained by multilevel model's predictor, we used a *pseudo-*R^2 statistic. At each step of modeling, Model A (empty model) served as a yardstick for comparison to compute the proportional reduction in residual variance as we added predictors. As the taxonomy in Supplement A in the online supplement shows, some models are nested, and others are non-nested. A model is nested in a subsequent model if every parameter in the former also appears in the subsequent one. This feature must be considered in comparing competing models.

To determine if a significant difference between two models existed when adding or eliminating model parameters, we used the likelihood ratio test. This test, which consists of a comparison of two nested models, was used to select the right model for our purposes. This ratio test compares the deviance (*- 2 Log L*) of two competing models by subtracting the smaller deviance from the larger deviance. This difference (Δ_D), which conforms to a Chi-square distribution when its log is multiplied by -2, is a Chi-square test with the number of degrees of freedom equal to the difference in the number of parameters (i.e., fixed effects and random effects) estimated in each model. When the two models compared involved differences in fixed effects, the models were fitted with a full maximum likelihood estimator (FML). Otherwise, we used the default restricted maximum likelihood estimator (REML) which involves only a difference in random effects. For comparison and selection of a non-nested model, we used Akaike's (1973) information criterion (*AIC*) and Bayesian information criterion (*BIC*; Schwarz, 1978). In this case, when comparing two models, the model with the lowest *AIC* and *BIC* values was selected.

In this study, missing data occurred for the math performance at the beginning of the year (1.50%) and for the outcome variable, math achievement at the end of the year (12.80%). We used Little's MCAR test (Little, 1988) which revealed a significant p-value for the test, indicating that the missing data did not seem to be completely at random $\chi^2(23, N = 1{,}020) = 42.7$, $p = .007$). As illustrated by Rubin (1976), when the missingness is not random, but where it can be fully accounted for by variables with complete information, data would be considered as missing at random (MAR). As the missing data evaluation shows (See Supplement C in the online supplement for more detailed information on the missing data analysis), the probability of missing data in math test at the end of the year is higher for students who have repeated a year. As the probability that a value for a variable is missing is related to other observed values in the dataset but not to the variable itself, we suppose that data are MAR. In addition, MAR also assumes that within the category of student who have repeated a year, math test performance scores are MCAR, because math test performance scores are randomly missing for student who have repeated a year and those who did not. As a result, means and standard deviations do not differ between the observed and missing data for the math performance test. Thus, if the parameters are estimated with Full Maximum Likelihood (FML), MAR will provide asymptotically unbiased estimates (Little & Rubin, 2002). The multilevel model allowed us to solve the problem of missing values in the outcome as they could be fitted with FIML methods, which took care of the missing data in the dependent variable. Analysis of the data was done using the R package lme4 (v1.1–26; Bates et al., 2015).

Results

Descriptive Statistics: Who was the Top Math Scorer in the Class?

Table 1 shows the repartition of girls and boys conditioning on top math scorer gender and conditional percentages. The results showed that there were no statistical differences in the distribution of girls and boys according to the gender of the top math scorer in the class, $\chi^2(1, N = 1{,}043) = 1.33$, $p = .25$). With regard to math scores (see Table 2) in the national test at the beginning of the school year, girls ($M = 12.98$, $SD = 3.42$) had a lower math achievement than boys ($M = 13.89$, $SD = 3.46$), $t(1018) = 4.24$, $p < .001$, $d = .26$). However this difference disappeared over the course of the year, as girls' math achievement for girls ($M = 11.00$, $SD = 4.43$) at the end of the year was the same as that of boys ($M = 10.95$, $SD = 4.78$), $t(905) = -.16$, $p = .86$. This result indicates that by the end of the year, the performance gap between girls and boys had disappeared (see Fig. 1).

Hypothesis 1: Additive Effects of Student Gender and Top Math Scorer Gender

As the girls had managed to narrow the performance gap by the end of the school year, the second part of the analysis sought to find out whether class characteristics such as top

Table 1 Class Gender Split and Top Math Scorer by Gender

Student Gender	Top Math Scorer Gender		
	Boy	Girl	Total
Boys	320 (60.60%)	208 (39.40%)	528
Girls	293 (56.90%)	222 (43.10%)	515
Total	613	430	1043

Conditioning percentages are presented in parentheses

Table 2 Means, Standard Deviations and Standard Errors of Math Scores as a Function of Each Group

Variable	Student Gender	Top Math Scorer Gender	*n*	*M*	*SD*	*SE*
Math score at the beginning of the year	Girls	Total	504	12.98	3.42	.15
		Girl	214	13.51	3.62	.25
		Boy	290	12.58	3.21	.19
	Boys	Total	516	13.89	3.46	.15
		Girl	204	14.40	3.31	.23
		Boy	312	13.55	3.52	.20
	Total		1020	13.43	3.47	.11
Math score at the end of the year	Girls	Total	444	11.00	4.43	.21
		Girl	187	12.01	4.55	.33
		Boy	257	10.27	4.20	.26
	Boys	Total	463	10.95	4.78	.22
		Girl	179	11.25	4.94	.37
		Boy	284	10.76	4.68	.28
	Total		907	10.97	4.61	.15

math scorer gender had moderated the drop in math achievement observed in all the students.

We fitted an unconditional model (Model A, Table 3) which only contained the random effect of the clustering variable (i.e., class grouping variable) at level-2 as a determinant of the intercept in the model predicting math achievement at level-1. In this model, each students' math scores for each student were predicted by the intercept that varied across classes. It aimed to break down the variance of math achievement into two levels: intraclass (level-1; 1043 students) and interclass (level-2; 46 classes). As the *ICC* shows that 17.82% of the math achievement variance can be explained at level-2, we included the class grouping variable in the models, and used multilevel modeling to take the data hierarchy into account. After including the initial math scores as a control variable in Model B, the math achievement variance was reduced from just over half (51.42%) at level-1, and from 37.39% at level-2. This model shows a better fit than the unconditional model ($\Delta_D(1) = 716.9$, $p < .001$). This model indicates that a one-unit increase in the initial math score is associated with an expected increase in math achievement of .18 ($p < .001$).

To find out whether math teacher gender impacted upon math achievement, we tested the teacher gender effect in Model C (level-2 variable). This variable was found to have no statistically significant effect and was therefore not retained in later analyses ($\beta = .73$, $p = .66$). In Model D, we incorporated student gender as an additional level-1 predictor to test if student gender accounted for differences in math scores at the end of the year. This had the effect of shifting the mean math score up for girls ($\beta = 1.07$, $p < .001$), which showed that girls had better math scores than boys at the end of the year, after controlling for the effect of initial scores.

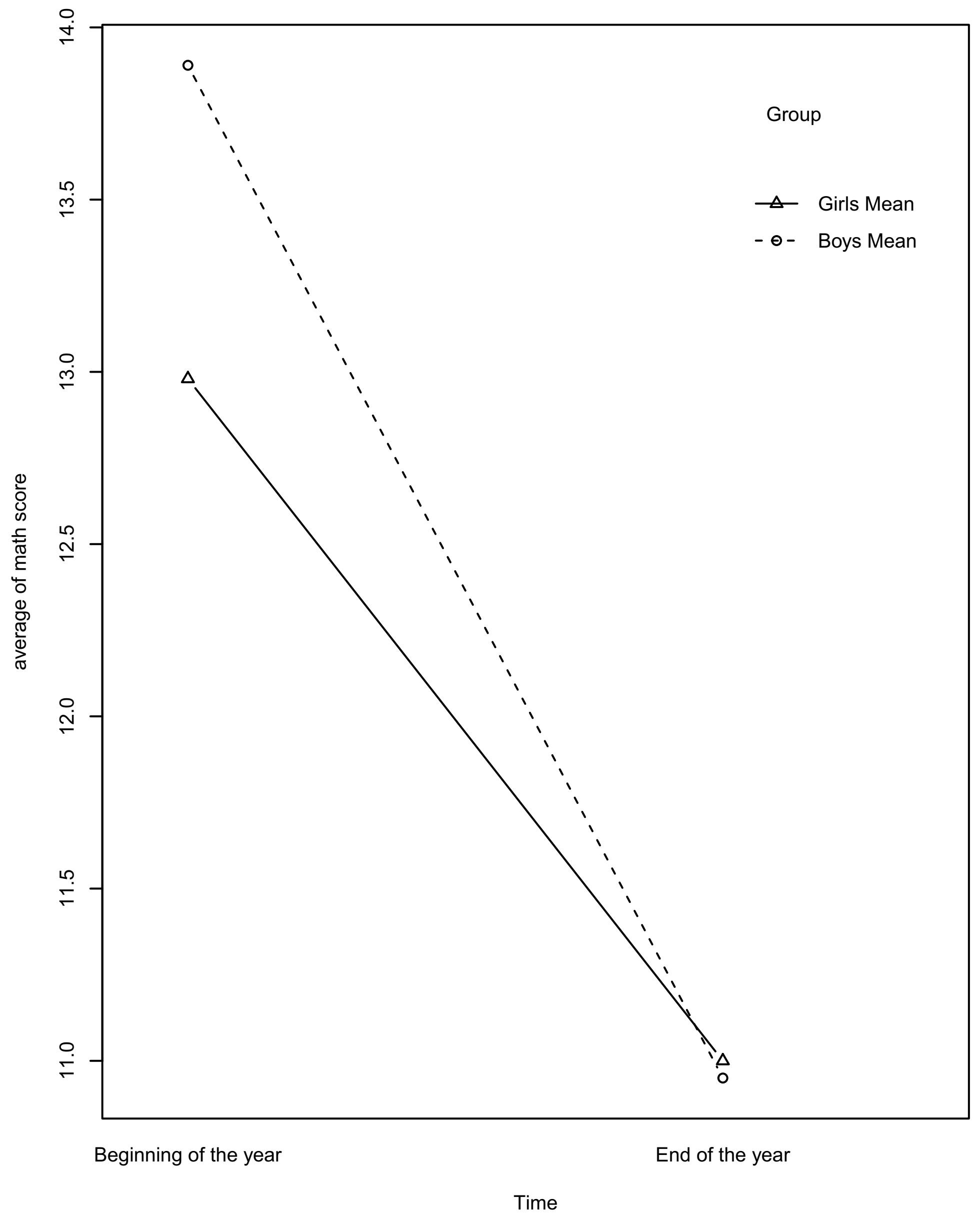
In Model E, the fixed effect of top math scorer gender slope (parameter γ_{02} in Table 1) was added to test the additive hypothesis (*H1*) by finding out if the top math scorer gender uniquely contributed to average math achievement, regardless of whether the student was a boy or a girl. This effect was non-significant ($\beta = .30$, $p = .56$), indicating that regardless of the gender of the top math scorer, there were non-significant differences in math achievement. In other words, being in a class where the top math scorer is either a girl or a boy did not influence the average score in math.

Hypothesis 2: Synergistic Effects of Student Gender and Top Math Scorer Gender

In Model F, we then added a level-2 random effect on the student gender slope (parameter u_{2j} in Table 1). In contrast to Model D, in which the student gender slope did not vary, in this random slope model, the student gender slopes varied across classes. Although the *pseudo* R^2 statistic was non-significant (Δ_D (1) = 3.17, $p = .075$), the introduction of this level-2 random effect resulted in a gain of 1.21 points in the explanation of residual variance at level-1 and a gain of 1.32 points at level-2 (comparison between Model D and Model F). In other words, this model suggests that the effect of student gender on math achievement may depend on level-2 predictors.

We then tested if student gender effect varied according to top math scorer gender, and we looked closely at girls' scores when the top math scorer was a girl. To do this, we added a cross-level interaction term (parameter γ_{22} in Supplement A in the online supplement) between student gender (level-1) and top math scorer gender (level-2) in Model G. The results indicate that the cross-level interaction term was statistically significant ($\beta = .93$, $p = .018$), which confirms

Fig. 1 Students' Math Performance Progression Over One School Year in Classes



that the positive fixed effect associated with being a girl is even stronger in a class where the top math scorer is a girl, (i.e., in a configuration where girls benefit from a female role model). Between Model E and Model G which are nested, a better fit was provided by Model G ($\Delta_D\,(1) = 5.53$, $p = .018$). The same conclusion was yielded by the comparison between Model F and G, which are non-nested with lower values for *AIC* and *BIC* for Model G. Consequently, whereas math scores decrease during the school year for both boys and girls, this decrease is smaller for girls, and especially for girls who are in a class where the top math scorer is a girl. This shows that girls benefit from the buffering effect of a female role model. In contrast, boys in a class where the top math scorer is a girl display the biggest drop in their math scores as can be seen in Fig. 2.

Discussion

The results from this study are unique in that no prior study has yet assessed the role of top math scorer gender on math achievement from a role model perspective. This study assessed two competing hypotheses regarding the effects of student gender and top math scorer gender on math achievement: *H1* considering that the effects of student gender and top math scorer gender are additive (i.e., a linear combination of their effects); and *H2* considering that the effects of student gender and top math scorer gender acting as a synergistic effect on each other (i.e., in a multiplicative combination of their effects). This study focused on math achievement as it is a school subject for which girls are particularly at risk of dropout due to stereotype threat. In this study, we

Table 3 Multilevel Multiple Linear Regressions with Math Score as Dependent Variable

			Model A	Model B	Model C	Model D	Model E	Model F	Model G
Fixed effects	Intercept_{ij}		10.87^{***}	-1.63^{***}	-2.24^{**}	-2.51^{***}	-2.62^{***}	-2.52^{***}	-2.38^{***}
(std. Error)			(.33)	(.48)	(.73)	(.51)	(.54)	(.51)	(.55)
	Math score at the beginning of the year		_	$.18^{***}$	$.18^{***}$	$.19^{***}$	$.19^{***}$	$.19^{***}$	$.19^{***}$
				(.01)	(.006)	(.01)	(.01)	(.01)	(.006)
	Student gender[a]		_	_	_	1.07^{***}	1.06^{***}	1.06^{**}	$.68^{**}$
						(.20)	(.20)	(.24)	(.25)
	Top math scorer gender		_	_	_	_	$.30^{ns}$	_	$-.15^{ns}$
							(.52)		(.55)
	Student gender x Top math scorer gender		_	_	_	_	_	_	$.93^{*}$
									(.39)
	Teacher gender[b]		_	_	$.73^{ns}$	_	_	_	_
					(.66)				
Random effects	Intercept	level-1	17.52	8.51	8.51	8.24	8.24	8.03	8.19
		level-2	3.80	2.38	2.36	2.42	2.47	2.37	2.47
	Student gender slope	level-2	_	_	_	_	_	.64	_
Fit indices	Pseudo R^{2c}	level-1	_	51.42%	51.42%	52.96%	52.96%	54.17%	53.25%
		level-2	_	37.37%	39.89%	36.31%	35%	37.63%	35%
	- 2 Log L		5242.6	4525.70	4524.43	4497.06	4496.71	4493.89	4491.18
	$\Delta_D{}^{d}$		_	(A—B)	(B—C)	(B—D)	(D—E)	(D—F)	(E—G)
				716.89^{***}	1.27^{ns}	28.65^{***}	$.35^{ns}$	3.17^{t}	5.53^{*}
	AIC		5248.60	4533.70	4534.43	4507.05	4508.71	4516.57	4505.18
	BIC		5263.03	4552.88	4558.40	4531.02	4537.47	4545.34	4538.74

$^{*}p<.05$; $^{**}p<.01$; $^{***}p<.001$; $t<.10$

[a]Student gender and top math scorer are coded 1 for girls, 0 for boys

[b]Teacher gender is coded 1 for woman, 0 for man

[c]Empty model is the reference to quantify how much of the outcome variance is explained by predictors

[d]A likelihood ratio test was performed between the two competing statistical models in brackets

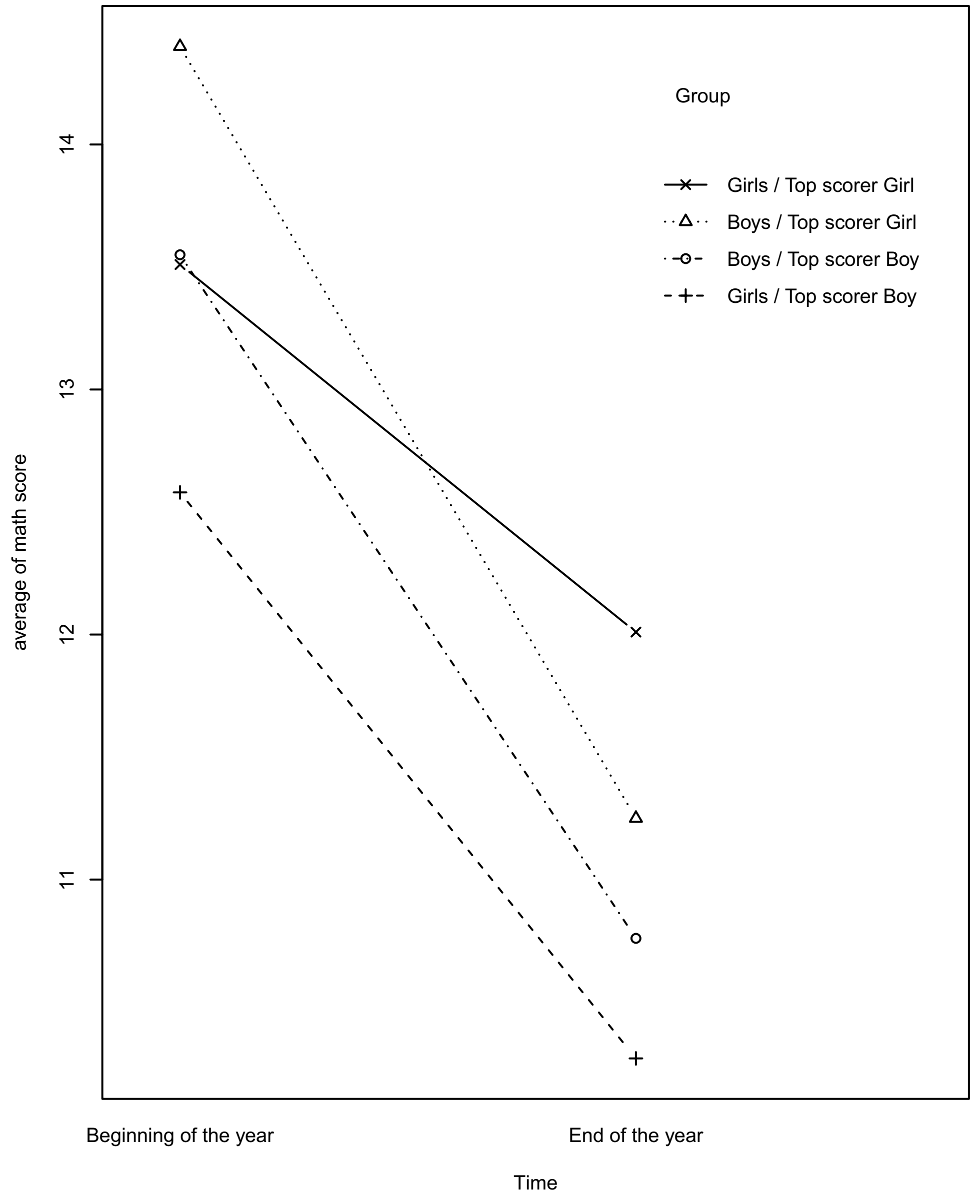
examined the influence of a counter-stereotypical role model in protecting girls from the gender gap in math performance.

To the best of our knowledge, no previous study has used a longitudinal design to examine the specific impact of the role model in an ordinary class where the math-gender stereotype threat is implicitly activated by the gender of the top math scorer in the class. Moreover, this study analyzed the effect of prolonged exposure to a counter-stereotypical role model involving real classmates who were physically present in the classroom throughout an entire school year. Previous studies on the impact of role models on girls' math achievement have been based on biographies of STEM role models, and exposure to senior undergraduate students, female teachers, or expert peers for students who had already specialized in STEM disciplines (Cheryan et al., 2011; Stout et al., 2011; Van Camp et al., 2019). Our study aimed to overcome these limitations by providing direct empirical evidence of the impact of counter-stereotypical role models for girls on their math achievement.

We observed that for all students (girls and boys), the math score decreased throughout the school year. This is not surprising, because the passage from elementary school to junior high school is an important transition period during which students must become accustomed to new teachers who are often perceived as more demanding and less supportive than those in elementary school (Cantin & Boivin, 2004). In junior high school, the frames of reference change and the class climate is different (e.g., more control, new standards of performance). Students must also cope with stricter grading, more complex learning, new notions, and more social comparison-based standards. They therefore experience more pressure and competition overall.

Despite this expected decline in math achievement for all students during the year, we expected to observe a buffering effect of being a girl. As PISA studies have shown, if the gender gap in math achievement decreases, then this is mainly due to a drop in boys' performance rather than an increase in girls' performance. Although girls are catching up (Hyde & Linn, 2006), this absence of a gender gap should not be taken to mean that stereotype threat is not operating at all (Huguet & Régner, 2009). Teachers and policy makers should not assume the absence of stereotype threat means there is no reason to worry about stereotype threat. There is ample evidence that girls continue to be negatively stereotyped in

Fig. 2 Students' Math Performance Progression Over One School Year in Classes Where the Top Math Scorer was a Girl or a Boy



mathematics (Good et al., 2008, 2012; Marx & Roman, 2002; Spencer et al., 1999; Steele, 1997). This fact justified our interest in the effect that counter-stereotypical role models embodied by the top math scorer could have, especially on girls' math achievement. Indeed, it was expected that math achievement would decline more for boys than for girls.

This relative decline was confirmed by our results which demonstrated the positive differential effect associated with student gender on math achievement. This main effect shows that the math achievement decline between the beginning and the end of the school year is less severe for girls than it is for boys, thus confirming our hypothesis relating to the buffering effect of girls' gender on math achievement decline. However, our results failed to support the additive hypothesis (*H1*) by showing that when considered together, only student gender was associated with math achievement; top math scorer gender was not. Observing this lack of additive contribution revealed that student gender is more important than top math scorer's gender in math achievement. Our study supports the synergistic effect hypothesis (*H2*) showing that when girls are in a class where a girl is the top math scorer throughout the school year, they perform better at the end of the school year, than they do in a class where a boy is the top math scorer. This cross-level interaction effect provides an optimal representation of the combined role played by student gender and the top math scorer gender, at least in the math domain.

Consequently, this study improves our understanding of the gender gap in math by showing that in addition to the buffering effect in favor of girls, student gender and top math scorer gender work together to produce a synergistic effect on math achievement. Support for this synergistic effect provides credible evidence that one way to protect the performance of girls in mathematics is to expose them to a real top math scorer counter-stereotypical role model. The direct observation of a counter-stereotypical role model boosts the buffering effect of being a girl and reduces the negative effect of the stereotype threat.

Although girls' math performance was lower than the one of boys at the beginning of the year, girls in classes where the top math scorer was a girl surpassed the boys at the end of the year. This was not the case for girls in a class where the top math scorer was a boy. Therefore, we can assert that the buffering effect of being a girl is necessary but needs to be reinforced by the effect of a counter-stereotypical role model, such as a girl top math scorer in order to see girls outperform boys.

Limitations and Future Research Directions

Though these findings advance understanding of the buffering and booster effect of a counter-stereotypical role model, these results should be viewed with caution for several reasons. First, our study was conducted only in schools that had volunteered for the research program. A future study could use a larger sample of schools to strengthen the statistical validity of the results. These correlational results also need to be tested by an experimental design, for instance, by asking participants to identify the top math scorer in their class before doing a math test, or by randomly assigning participants to a condition with a male top math scorer or a female top math scorer in mathematics. Another limitation concerns the role of identification with the model target, which was not measured in this study. If students prefer upward identification and downward contrast because these two mechanisms are not threatening to the self (Bouffard et al., 2014; Buunk & Dijkstra, 2017; Buunk et al., 1990), then it is also very likely that these mechanisms could moderate the effect of a counter-stereotypic role model on performance. Future research is thus needed to test the moderating role of identification on the protective effect of a counter-stereotypical role model for girls in math.

Second, although not a limitation per se, this study did not rely on students' ratings, which would have provided a complementary perspective in identifying the underlying mechanisms of this beneficial configuration for girls, but also those involved in the debilitating effect on boys' math achievement. Although the ratings obtained in this present study did not allow us to consider the psychological processes which may help both girls and boys to become more confident in their ability, it would be interesting for future studies to incorporate the impact of role models on perceived self-efficacy in math, and in turn in math achievement, as Bagès et al. (2016) did. Careful selection of the role models would be required to ensure a similarity between the role model and aspirant roles. The effects of a role model cannot be viewed unilaterally. Instead, its effect might depend on the role aspirants' learning processes (Morgenroth et al., 2015), and on their ambitions, expectations, and inclinations. This dynamic process is often overlooked in the current view of role models in education. Integrating socio-cognitive theories, in particular observational learning theory (Bandura, 1965, 1969, 1977) would then be a very promising avenue (Ahn et al., 2020). Further research is therefore needed to draw firmer conclusions about the impact of counter-stereotypical role models on role aspirants and to incorporate the other processes that may shape girls' aspirations and attitudes in STEM fields. A future study could also explore the students' perceptions of the top math scorer in their math class, and check whether this fits the measure we used.

Also, to better understand the effects of a role model, it would be relevant to study the conditions of the emergence of counter-stereotype role models. For instance, by looking at a macro level, at the impact of environmental factors such as national education policies and school climate. Sherblom et al. (2006) have shown that the school climate is associated with academic success both in math and reading. It can be hypothesized that some school climates may favor the emergence of female role models, and others of male role models. It is therefore possible that contextual moderators affect the effectiveness of role models on performance. Future research on this topic would be valuable. Motivational theories such as achievement goal theory (Ames, 1992) and self-determination theory (Deci & Ryan, 1985) could provide insight into the potential multiple processes involved in stereotype threat effects on performance, and thus could provide avenues for practical interventions to support the retention of girls in math.

More specifically, goal structure (Ryan & Ryan, 2005) and motivational style (Nadler & Komarraju, 2016) may alter the experience and impact of stereotype threat on achievement. For example, the focus of mastery goal structure on developing personal competence is positively related to task engagement (Harackiewicz et al., 2000, 2002; Hulleman et al., 2010), whereas the focus of competitive performance goal structure on demonstrating competence is a negative predictor of academic performance (Hulleman et al., 2010). A mastery goal structure may act as a buffer between the experience of stereotype threat and mathematic achievement. This may shield students from the experience of stereotype threat, whereas a performance goal structure may amplify the negative effect of stereotype threat, which would be more deleterious for those students who experience it.

Third, our results suggest that the direct observation of a counter-stereotypic pattern amplifies the protective effect in stigmatized (i.e., negatively stereotyped) students, and reduces the negative effect of the stereotype against them. Although this finding seems to apply to girls in our study, the question remains open for boys. It would be worth exploring the student factors that could interact with the role model effect to contribute to our understanding of boys' underachievement even in a class where the top scorer is a boy. As the effects of stereotype threat are likely to depend on individual awareness of cultural stereotypes (Hartley & Sutton, 2013), an examination of whether awareness of stereotypes acts as a moderator of the effect of the top scorer gender could help us better understand why boys may not succeed in a class where the top scorer is a boy. Our results found that boys as a group did not benefit from the role model effect. However, does this apply to all boys or only some? Previous research (e.g., Aronson et al., 1999; Nguyen & Ryan, 2008) has shown that identification with the domain moderates stereotype threat effects. Therefore, boys' susceptibility to benefit from a role model may be moderated by the extent to which they identify with, and value academic success. Whether or not a boy has identified with math may matter because students with differing levels of identification tend to have different concerns (Cheryan & Plaut, 2010; Easterbrook & Hadden, 2021; Steele, 1997). In addition, the ongoing experience of other cultural stereotype threats, such as the belief that boys are often academically inferior to girls (Hartley & Sutton, 2013), may lead to weaker identification and/or disengagement with the math domain for some boys.

Finally, the period investigated here was relatively short since our study was limited to one school year. To provide more support for these results, the influence of counter-stereotypical role models should be explored over a longer period of time. Future research could also assess gender stereotypes and subsequent behaviors longitudinally, to determine whether a change in stereotypes is internalized and implemented. As suggested by previous studies (e.g., González-Pérez et al., 2020; O'Brien et al., 2017) longitudinal studies with repeated exposure to role models are needed to develop our understanding of the process and the lasting effects of role model exposure.

Practice Implications

Because of its pivotal role in everyday life and in many vocations (Leder & Forgasz, 2018; Valero, 2017), gender differences in math achievement are an important topic for which multiple interpretations have been proposed. One possible interpretation is that such differences do not really exist in terms of academic competence, but that they emerge because of multiple moderating variables. Addressing questions about the relation between student gender, top scorer gender, and math achievement as well as questions about the nature of these relations (i.e., whether they are additive or synergistic) is important for accurate gender effect assessments. For instance, the combined effect of the student gender effect and top scorer gender could be substantially underestimated under the additive assumption. In addition, they are essential in terms of targeting an intervention; for instance, the benefits of an intervention (i.e., promoting girls' achievement in math) could be greater for girls in classes where the top scorer is a girl under the synergistic assumption. Furthermore, they would be informative in understanding the complex mechanisms of gender effects on math achievement. From an applied perspective, these findings suggest some useful ideas for teachers' practices. For example, one way to train teachers would be to teach them to recognize for whom, when, and in what situations a gender counter-stereotypical model may be useful to students (e.g., in our study, it is only useful for girls). We could also provide teachers with guidelines for action, such as the importance of care in selecting the role models, which is required to ensure the similarity between the role model and aspirant roles, as well as varying the role models to avoid the stigmatization of certain students.

Moreover, the results of the multilevel analysis suggest that the counter-stereotypical role model only played a role in boosting the buffering effect of being a girl in mathematics achievement, since boys did not actually succeed better in classes where the top math scorer was a boy. In this regard, it is noteworthy that manipulating these factors can alter gender differences by reducing the gap for girls, while it may increase it for boys. Whilst it is important to distinguish adverse impact from gender equity, finding a booster effect related to the top math scorer gender for girls should not be used to support gender-segregated education, because a boy in a class where the top scorer is a boy doesn't produce such a booster effect. Indeed, nations with a high percentage of gender-segregated math classes display considerable differences in math achievement (Wiseman, 2008). Investigating the mechanisms underlying these effects and analyzing a wide range of potential mediators and moderators, such as self-perceptions, professional aspirations, and motivations, would be interesting from a theoretical as well as practical point of view.

In practice, there are some issues to address because the gender of the top scorer in a class cannot be manipulated. If we artificially control the top scorer, e.g., by reassigning the top scorer to a different class, or by providing contrived feedback about which student is the top scorer, then it is possible to observe the counter-stereotype role model effect. The aim here is to focus on the enhancement of the performance of already high-performing girls by reducing stereotype threat: by deemphasizing threatened social identities (Ambady et al., 2004; Shih et al., 1999), by encouraging self-affirmation (Croizet et al., 2001; Martens et al., 2006), or by emphasizing an incremental view of ability (Dar-Nimrod & Heine, 2006; Thoman et al., 2008). Finally, understanding how class climate interacts with the effects of top scorer gender over time may help address issues of gender disparity. Interventions aimed at increasing girls' engagement in STEM should favor a class context that does not shield students from stereotype threat but alters their defensive reactions to that threat. Focusing on math skills and using a self-determination motivational style could create an environment that supports mastery goals (e.g., Smeding et al., 2013) and psychological need satisfaction (Nadler & Komarraju, 2016), which in turn may negate the effect of stereotype threat.

Conclusion

In conclusion, our results add to the growing body of literature on the effect of a counter-stereotypical role model for girls who are negatively stereotyped in mathematics. These results add empirical support to the previous research findings concerning the importance of role models. Using a longitudinal design, we observed for the first time that exposure to a counter-stereotypical role model in class reinforced the buffering effect of being a girl, in the sense that it reduced the drop in math achievement for girls. When girls develop in a class where a girl is the math top scorer, they receive the message that math lies within the realm of possibilities for them. Girls will perform at the same level as their male classmates when they are encouraged to succeed, and a good way to provide such an opportunity is to have visible girl role models excelling in mathematics, such as the top math scorer. Our study also identified open questions about why role models are more beneficial for some students and in certain conditions. Future research could try to answer these questions which are particularly relevant to STEM gender diversification. However, we caution that by simply focusing on role models as a solution for better math performance risks sending the message that exposing students to counter-stereotype role models is an easy route to improving math achievement. This strategy can only succeed if it is part of a broader set of measures to make math more valuable and appealing for all students.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11199-022-01289-y.

Funding The authors did not receive support from any organization for the submitted work.

Availability of Data and Material Data and results of the analyses can be obtained on request from the first author.

Compliance with Ethical Standards

Ethics Approval The authors certify that the study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki. This study is non-interventional research and the approval of the French state authorities was not required.

Informed Consent Informed consent was obtained from all individual participants, school authorities and legal guardians included in the study.

Conflicts of Interest The authors have no relevant financial or non-financial interests to disclose.

References

Ahn, J. N., Hu, D., & Vega, M. (2020). Do as I do, not as I say: Using social learning theory to unpack the impact of role models on students' outcomes in education. *Social and Personality Psychology Compass*, *14*(2), e12517. https://doi.org/10.1111/spc3.12517

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Second international symposium on information theory* (pp. 267–281). Akademiai Kiado.

Ambady, N., Paik, S. K., Steele, J., Owen-Smith, A., & Mitchell, J. P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology, 40*(3), 401–408. https://doi.org/10.1016/j.jesp.2003.08.003

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271. https://doi.org/10.1037/0022-0663.84.3.261

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*(1), 29–46. https://doi.org/10.1006/jesp.1998.1371

Bagès, C., & Martinot, D. (2011). What is the best model for girls and boys faced with a standardized mathematics evaluation situation: A hardworking role model or a gifted role model? *British Journal of Social Psychology, 50*(3), 536–543. https://doi.org/10.1111/j.2044-8309.2010.02017.x

Bagès, C., Verniers, C., & Martinot, D. (2016). Virtues of a hardworking role model to improve girls' mathematics performance. *Psychology of Women Quarterly, 40*, 55–64. https://doi.org/10.1177/0361684315608842

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology, 1*(6), 589–595. https://doi.org/10.1037/h0022070

Bandura, A. (1969). Social learning of moral judgments. *Journal of Personality and Social Psychology, 11*(3), 275–279. https://doi.org/10.1037/h0026998

Bandura, A. (1972). Modeling theory: Some traditions, trends and disputes. In R. D. Park (Ed.), *Recent trends in social learning theory* (pp. 35–61). Academic Press. https://doi.org/10.1016/B978-0-12-545050-8.50008-0

Bandura, A. (1977). *Social learning theory*. Prentice-Hall.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*. https://doi.org/10.18637/jss.v067.i01

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., & Doms, M. E. (2011). Women in STEM: A gender gap to innovation. *Economics and Statistics Administration Issue Brief*, (04–11). https://doi.org/10.2139/ssrn.1964782

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spillover. *Journal of Experimental Psychology: General, 136*, 256–276. https://doi.org/10.1037/0096-3445.136.2.256

Bertrand, M., & Duflo, E. (2017). Field experiments on discrimination. *Handbook of Economic Field Experiments, 1*, 309–393. https://doi.org/10.1016/bs.hefe.2016.08.004

Betz, D. E., & Sekaquaptewa, D. (2012). My fair physicist? Feminine math and science role models demotivate young girls. *Social Psychological and Personality Science, 3*, 738–746. https://doi.org/10.1177/1948550612440735

Bigler, R. S., & Liben, L. S. (1990). The role of attitudes and interventions in gender- schematic processing. *Child Development, 61*, 1440–1452. https://doi.org/10.1111/j.1467-8624.1990.tb02873.x

Bouffard, T., Pansu, P., Boissicat, N., Vezeau, C., & Cottin, F. (2014). Questionnaire de la comparaison de soi scolaire pour enfants et adolescents [Questionnaire of the comparison of academic self for children and adolescents]. *Canadian Journal of Behavioural Science/revue Canadienne Des Sciences Du Comportement, 46*(2), 84–94. https://doi.org/10.1037/a0028429

Bryk, S. W., & Raudenbush, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Sage Publications.

Buunk, B. P., Collins, R. L., Taylor, S. E., VanYperen, N. W., & Dakof, G. A. (1990). The affective consequences of social comparison: Either direction has its ups and downs. *Journal of Personality and Social Psychology, 59*(6), 1238–1249. https://doi.org/10.1037/0022-3514.59.6.1238

Buunk, A. P., & Dijkstra, P. (2017). Social comparisons and well-being. In M. D. Robinson & M. Eid (Eds.), *The happy mind: Cognitive contributions to well-being* (pp. 311–330). Springer International Publishing/Springer Nature. https://doi.org/10.1007/978-3-319-58763-9_17

Cantin, S., & Boivin, M. (2004). Change and stability in children's social network and self-perceptions during transition from elementary to junior high school. *International Journal of Behavioral Development, 28*, 561–570. https://doi.org/10.1080/01650250444000289

Cantley, I., & McAllister, J. (2021). The gender similarities hypothesis: Insights from a multilevel analysis of high-stakes examination results in mathematics. *Sex Roles, 85*, 481–496. https://doi.org/10.1007/s11199-021-01234-5

Cheryan, S., & Plaut, V. C. (2010). Explaining underrepresentation: A theory of precluded interest. *Sex Roles, 63*(7–8), 475–488. https://doi.org/10.1007/s11199-010-9835-x

Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Social Psychological and Personality Science, 2*, 656–664. https://doi.org/10.1177/1948550611405218

Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin, 119*, 51–69. https://doi.org/10.1037/0033-2909.119.1.51

Collins, R. L. (2000). Among the better ones: Upward assimilation in social comparison. In J. Suls & L. Wheeler (Eds.), *The Plenum series in social/clinical psychology. Handbook of social comparison: Theory and research* (pp. 159–171). Dordrecht, Netherlands: Kluwer Academic Publishers. https://doi.org/10.1007/978-1-4615-4237-7_9

Correll, S. J. (2001). Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology, 106*, 1691–1730. https://doi.org/10.1086/321299

Croizet, J.-C., Désert, M., Dutrévis, M., & Leyens, J.-P. (2001). Stereotype threat, social class, gender, and academic under-achievement: When our reputation catches up to us and takes over. *Social Psychology of Education: an International Journal, 4*(3–4), 295–310. https://doi.org/10.1023/A:1011336821053

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science, 314*(5798), 435. https://doi.org/10.1126/science.1131100

Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counter stereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology, 40*, 642–658. https://doi.org/10.1016/j.jesp.2004.02.003

Dasgupta, N. (2011). Ingroup experts and peers as social vaccines who inoculate the self-concept: The stereotype inoculation model. *Psychological Inquiry, 22*(4), 231–246. https://doi.org/10.1080/1047840X.2011.607313

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media. https://doi.org/10.1007/978-1-4899-2271-7

Eagly, A. H., & Wood, W. (2011). Social role theory. In P. van Lange, A. Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories in Social Psychology*, (vol. 2, pp. 458–476). Thousand Oaks, CA: Sage Publications. https://doi.org/10.4135/9781446249222.n49

Easterbrook, M. J., & Hadden, I. R. (2021). Tackling educational inequalities with social psychology: Identities, contexts, and interventions. *Social Issues and Policy Review, 15*(1), 180–236. https://doi.org/10.1111/sipr.12070

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*, 103–127. https://doi.org/10.1037/a0018053

Frawley, T. J. (2008). Gender schema and prejudicial recall: How children misremember, fabricate, and distort gendered picture book information. *Journal of Research in Childhood Education, 22*, 291–303. https://doi.org/10.1080/02568540809594628

Fryer, R. G., Jr., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics, 2*, 210–240. https://doi.org/10.1257/app.2.2.210

Ghasemi, E., Burley, H., & Safadel, P. (2019). Gender differences in general achievement in Mathematics: An international study. *New Waves Educational Research & Development*, *22*(1), 27–54. https://files.eric.ed.gov/fulltext/EJ1229446.pdf

Goldstein, H. (2011). *Multilevel statistical models*. Wiley. https://doi.org/10.1002/9780470973394

González-Pérez, S., Mateos de Cabo, R., & Sáinz, M. (2020). Girls in STEM: Is it a female role-model thing? *Frontiers in Psychology*, *11*(2204). https://doi.org/10.3389/fpsyg.2020.02204

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology, 29*, 17–28. https://doi.org/10.1016/j.appdev.2007.10.004

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology, 102*(4), 700–717. https://doi.org/10.1037/a0026659

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. *Science, 320*, 1164–1165. https://doi.org/10.1126/science.1154094

Hadjar, A., Krolak-Schwerdt, S., Priem, K., & Glock, S. (2014). Gender and educational achievement. *Educational Research, 56*, 117–125. https://doi.org/10.1080/00131881.2014.898908

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*, 1–51. https://doi.org/10.1111/j.1529-1006.2007.00032.x

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology, 92*(2), 316–330. https://doi.org/10.1037/0022-0663.92.2.316

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal* *of Educational Psychology, 94*(3), 562–575. https://doi.org/10.1037/0022-0663.94.3.562

Hartley, B. L., & Sutton, R. M. (2013). A stereotype threat account of boys' academic underachievement. *Child Development, 84*(5), 1716–1733. https://doi.org/10.1111/cdev.12079

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*(5220), 41–45. https://doi.org/10.1126/science.7604277

Helwig, R., Anderson, L., & Tindal, G. (2001). Influence of elementary student gender on teachers' perceptions of mathematics achievement. *The Journal of Educational Research, 95*, 93–102. https://doi.org/10.1080/00220670109596577

Hoffmann, F., & Oreopoulos, P. (2009). Professor qualities and student achievement. *The Review of Economics and Statistics, 9*, 83–92. https://doi.org/10.1162/rest.91.1.83

Hoyt, C. L., & Simon, S. (2011). Female leaders: Injurious or inspiring role models for women*?*. *Psychology of Women Quarterly, 35*, 143–157. https://doi.org/10.1177/0361684310385216

Huguet, P., & Régner, I. (2009). Counter-stereotypic beliefs in math do not protect school girls from stereotype threat. *Journal of Experimental Social Psychology, 45*, 1024–1027. https://doi.org/10.1016/j.jesp.2009.04.029

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin, 136*(3), 422–449. https://doi.org/10.1037/a0018947

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990a). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139–155. https://doi.org/10.1037/0033-2909.107.2.139

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990b). Gender comparisons of mathematics attitudes and affect. *Psychology of Women Quarterly, 14*, 299–324. https://doi.org/10.1111/j.1471-6402.1990.tb00022.x

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A., & Williams, C. (2008). Gender similarities characterize math performance. *Science, 321*, 494–495. https://doi.org/10.1126/science.1160364

Hyde, J. S., & Linn, M. C. (2006). Gender similarities in mathematics and science. *Science, 314*, 599–600. https://doi.org/10.1126/science.1132154

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences, 108*(22), 8801–8807. https://doi.org/10.1073/pnas.0901265106

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*, 365–371. https://doi.org/10.1111/1467-9280.00272

Inzlicht, M., & Ben-Zeev, T. (2003). Do high-achieving female students underperform in private? The implications of threatening environments on intellectual processing. *Journal of Educational Psychology, 95*, 796–805. https://doi.org/10.1037/0022-0663.95.4.796

Jacobs, J. E., & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology, 63*, 932–944. https://doi.org/10.1037/0022-3514.63.6.932

Kit, K. A., Tuokko, H. A., & Mateer, C. A. (2008). A review of the stereotype threat literature and its application in a neurological population. *Neuropsychology Review, 18*, 132–148. https://doi.org/10.1007/s11065-008-9059-9

Leder, G. C., & Forgasz, H. J. (2018). Measuring who counts: Gender and mathematics assessment. *ZDM, 50*(4), 687–697. https://doi.org/10.1007/s11858-018-0939-z

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*, 1123–1135. https://doi.org/10.1037/a0021276

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*, 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Little, R. J. A., & Rubin, D. B. (2002). Statistical analysis with missing data. *John Wiley & Sons*. https://doi.org/10.1002/9781119013563

Lockwood, P. (2006). Someone like me can be successful: Do college students need same-gender role models? *Psychology of Women Quarterly, 30*, 36–46. https://doi.org/10.1111/j.1471-6402.2006.00260.x

Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology, 73*, 91–103. https://doi.org/10.1037/0022-3514.73.1.91

Lockwood, P., & Kunda, Z. (1999). Increasing the salience of one's best selves can undermine inspiration by outstanding role models. *Journal of Personality and Social Psychology, 76*, 214–228. https://doi.org/10.1037/0022-3514.76.2.214

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology, 42*(2), 236–243. https://doi.org/10.1016/j.jesp.2005.04.010

Marx, D. M., & Ko, S. J. (2012). Prejudice, discrimination, and stereotypes (Racial Bias). In V. S. Ramachandran (Ed.), *Encyclopedia of Human Behavior (Second Edition)* (Second Edition, p. 160-166). Academic Press. https://doi.org/10.1016/B978-0-12-375000-6.00388-8

Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin, 28*, 1183–1193. https://doi.org/10.1177/01461672022812004

McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology, 39*, 83–90. https://doi.org/10.1016/S0022-1031(02)00513-9

Morgenroth, T., Ryan, M. K., & Peters, K. (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. *Review of General Psychology, 19*, 1–19. https://doi.org/10.1037/gpr0000059

Nadler, D. R., & Komarraju, M. (2016). Negating stereotype threat: Autonomy support and academic identification boost performance of African American college students. *Journal of College Student Development, 57*(6), 667–679. https://doi.org/10.1353/csd.2016.0039

Nguyen, H.-H.D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314–1334. https://doi.org/10.1037/a0012702

Nhundu, T. J. (2007). Mitigating gender-typed occupational preferences of Zimbabwean primary school children: The use of biographical sketches and portrayals of female role models. *Sex Roles, 56*, 639–649. https://doi.org/10.1007/s11199-007-9204-6

Nollenberger, N., Rodríguez-Planas, N., & Sevilla, A. (2016). The math gender gap: The role of culture. *American Economic Review, 106*, 257–251. https://doi.org/10.1257/aer.p20161121

O'Brien, L. T., Hitti, A., Shaffer, E., Camp, A. R. V., Henry, D., & Gilbert, P. N. (2017). Improving girls' sense of fit in science: Increasing the impact of role models. *Social Psychological and Personality Science, 8*, 301–309. https://doi.org/10.1177/1948550616671997

OECD. (2015). *The ABC of gender equality in education: Aptitude*. PISA, OECD Publishing. https://doi.org/10.1787/9789264229945-en

OECD. (2016). *PISA 2015 results (Volume II): Policies and practices for successful schools*. PISA, OECD Publishing. https://doi.org/10.1787/9789264267510-en

OECD. (2018). *Résultats du PISA 2018 (Volume I): Savoirs et savoir-faire des élèves*. PISA, Éditions OCDE. https://doi.org/10.1787/ec30bc50-fr

Olsson, M., & Martiny, S. E. (2018). Does exposure to counter stereotypical role models influence girls' and women's gender stereotypes and career choices? A review of social psychological research. *Frontiers in Psychology*, *9*, 2264. https://doi.org/10.3389/fpsyg.2018.02264

Paredes, V. (2014). A teacher like me or a student like me? Role model versus teacher bias effect. *Economics of Education Review, 39*, 38–49. https://doi.org/10.1016/j.econedurev.2013.12.001

Plant, E. A., Baylor, A. L., Doerr, C. E., & Rosenberg-Kima, R. B. (2009). Changing middle-school students' attitudes and performance regarding engineering with computer-based social models. *Computers & Education, 53*, 209–215. https://doi.org/10.1016/j.compedu.2009.01.013

Rattan, A., Good, C., & Dweck, C. S. (2012). It's ok - Not everyone can be good at math: Instructors with an entity theory comfort (and demotivate) student. *Journal of Experimental Social Psychology, 48*, 731–737. https://doi.org/10.1016/j.jesp.2011.12.012

Ryan, K. E., & Ryan, A. M. (2005). Psychological processes underlying stereotype threat and standardized math test performance. *Educational Psychologist, 40*(1), 53–63. https://doi.org/10.1207/s15326985ep4001_4

Régner, I., Steele, J. R., Ambady, N., Thinus-Blanc, C., & Huguet, P. (2014). Our future scientists: a review of stereotype threat in girls from early elementary school to middle school. *International Review of Social Psychology*, *27*, 13–51. https://hal.archives-ouvertes.fr/hal-01432455

Roberts, T.-A. (1991). Gender and the influence of evaluations on self-assessments in achievement settings. *Psychological Bulletin, 109*(2), 297–308. https://doi.org/10.1037/0033-2909.109.2.297

Rodríguez-Planas, N., & Nollenberger, N. (2018). Let the girls learn! It is not only about math … it's about gender social norms. *Economics of Education Review, 62*, 230–253. https://doi.org/10.1016/j.econedurev.2017.11.006

Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A., & Doerr, C. E. (2008). Interface agents as social models for female students: The effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers in Human Behavior, 24*, 2741–2756. https://doi.org/10.1016/j.chb.2008.03.017

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*, 581–592. https://doi.org/10.1093/biomet/63.3.581

Rydell, R. J., & Boucher, K. L. (2017). Stereotype threat and learning. In J. M. Olson (Ed.), *Advances in experimental social psychology: Advances in experimental social psychology* (pp. 81–129). Elsevier Academic Press.

Rydell, R. J., Rydell, M. T., & Boucher, K. L. (2010). The effect of negative performance stereotypes on learning. *Journal of Personality and Social Psychology, 99*, 883–896. https://doi.org/10.1037/a0021139

Sáinz, M., & Eccles, J. (2012). Self-concept of computer and math ability: Gender implications across time and within ICT studies. *Journal of Vocational Behavior, 80*, 486–499. https://doi.org/10.1016/j.jvb.2011.08.005

Savenye, W. C. (1990). Role models and student attitudes toward nontraditional careers. *Educational Technology Research and Development, 38*, 5–13. https://doi.org/10.1007/BF0229817

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*, 440–452. https://doi.org/10.1037/0022-3514.85.3.440

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*, 336–356. https://doi.org/10.1111/1467-9280.00111

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464. https://doi.org/10.1214/aos/1176344136

Sherblom, S. A., Marshall, J. C., & Sherblom, J. C. (2006). The relationship between school climate and math and reading achievement. *Journal of Research in Character Education*, *4*(1–2), 19+. https://link.gale.com/apps/doc/A234791293/AONE?u=anon~9f53ecca&sid=googleScholar&xid=3edb67bf

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80–83. https://doi.org/10.1111/1467-9280.00111

Shin, J. E. L., Levy, S. R., & London, B. (2016). Effects of role model exposure on STEM and non-STEM student engagement. *Journal of Applied Social Psychology, 46*, 410–427. https://doi.org/10.1111/jasp.12371

Smeding, A., Darnon, C., Souchal, C., Toczek-Capelle, M. -C., & Butera, F. (2013). Reducing the socio-economic status achievement gap at university by promoting mastery-oriented assessment. *PLoS One*, *8*(8), e71678. https://doi.org/10.1371/journal.pone.0071678

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science? A critical review. *American Psychologist, 60*, 950–958. https://doi.org/10.1037/0003-066X.60.9.950

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4–28. https://doi.org/10.1006/jesp.1998.1373

Spinath, B., & Spinath, F. M. (2005). Longitudinal analysis of the link between learning motivation and competence beliefs among elementary school children. *Learning and Instruction, 15*, 87–102. https://doi.org/10.1016/j.learninstruc.2005.04.008

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629. https://doi.org/10.1037/0003-066x.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology, 100*, 255–270. https://doi.org/10.1037/a0021385

Taylor, V. J., & Walton, G. M. (2011). Stereotype threat undermines academic learning. *Personality and Social Psychology Bulletin, 37*, 1055–1067. https://doi.org/10.1177/0146167211406506

Thoman, D. B., White, P. H., Yamawaki, N., & Koishi, H. (2008). Variations of gender-math stereotype content affect women's vulnerability to stereotype threat. *Sex Roles, 58*(9–10), 702–712. https://doi.org/10.1007/s11199-008-9390-x

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology, 92*, 144–151. https://doi.org/10.1037/0022-0663.92.1.144

Tomasetto, C., & Appoloni, S. (2013). A lesson not to be learned? Understanding stereotype threat does not protect women from stereotype threat. *Social Psychology of Education: an International Journal, 16*, 199–213. https://doi.org/10.1007/s11218-012-9210-6

Valero, P. (2017). Mathematics for all, economic growth, and the making of the citizen-worker. In T. S. Popkewitz, J. Diaz, & C. Kirchgasler (Eds.), *Political sociology and transnational educational studies: The styles of reason governing teaching, curriculum and teacher education* (pp. 117–132). Routledge.

Van Camp, A., Gilbert, P., & O'Brien, L. T. (2019). Testing the effects of a role model intervention on women's STEM outcomes. *Social Psychology of Education, 22*, 649–671. https://doi.org/10.1007/s11218-019-09498-2

Weisgram, E. S., & Bigler, R. S. (2007). Effects of learning about gender discrimination on adolescent girls' attitudes toward and interest in science. *Psychology of Women Quarterly, 31*, 262–269. https://doi.org/10.1111/j.1471-6402.2007.00369.x

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81. https://doi.org/10.1006/ceps.1999.1015

Wijsman, L. A., Warrens, M. J., Saab, N., van Driel, J. H., & Westenberg, P. M. (2016). Declining trends in student performance in lower secondary education. *European Journal of Psychology of Education, 31*, 595–612. https://doi.org/10.1007/s10212-015-0277-2

Wiseman, A. W. (2008). A culture of (in)equality?: A cross-national study of gender parity and gender segregation in national school systems. *Research in Comparative and International Education, 3*(2), 179–201. https://doi.org/10.2304/rcie.2008.3.2.179