UNIVERSITÀ DEGLI STUDI DI MILANO

ECONOMIC SOCIOLOGY AND LABOUR STUDIES - 34th Cohort

Dipartimento di Scienze Sociali e Politiche

The stubbornness of gender stereotypes in education: A quantitative and experimental study on beliefs, attitudes and role models among high-school students in Italy

Elena De Gioannis

Supervisor: Prof. Flaminio Squazzoni

Director of Doctoral Program: Prof. Gabriele Ballarino

A.A. 2020/2021

Table of Contents

List of Tables

List of Figures

Acknowledgements

This thesis benefited from the encouragement, support and critical guidance that I received from many people during these last three years.

My supervisor, Prof. Flaminio Squazzoni, deserves a special thanks for allowing me a tremendous amount of freedom and support to make certain important decisions for my work. I have come to realize how unusual this sort of freedom is. I am grateful for the encouragement he showed to me over these years, even and especially at critical stages of my work. His suggestions, expertise and availability has been valuable on many occasions.

I would like to acknowledge the faculty of my department, especially Prof. Gabriele Ballarino and Dr. Niccolò Casnici, for their valuable suggestions and encouragement received on several occasions during the Ph.D. programme. I wish to extend my special thanks to Dr. Federico Bianchi for his helpful advices and support.

I thank my friends and colleagues for emotional support, stimulating discussions and happy distractions to rest my mind outside of my research. I would like to thank my boyfriend for being supportive, patiently listening when I was thinking aloud and helping me to clear my mind. Lastly, I am grateful to my parents, who encouraged me to enrol in a Ph.D. programme and whose support has never lacked. A special thanks goes to my mother, who helped me collecting information on schools, and to my sister, who created the videos for the experiment.

Chapter 1 ***Introduction***

The underrepresentation of women in STEM (Science, Technology, Engineering, and Mathematics) disciplines is a widely recognized problem. Even if women on average outnumber men in tertiary education (Vincent-Lancrin, 2008), in most countries still girls in STEM courses are a minority (UNESCO, 2017). Conversely, they are the majority in fields traditionally considered 'feminine', such as humanities. Women's educational choices at secondary and tertiary levels are subsequently reflected in career paths and determine the low engagement of women in STEM jobs.

Reducing this gender gap is both necessary and desirable for multiple reasons. From an ethical point of view, empowering women in STEM is pivotal if we want to achieve gender equality. Women working in STEM face the same challenges as working women, e.g., sexual harassment and lower salaries, however, compared to non-STEM working women, they are more likely to experience discrimination in the workplace (Funk & Parker, 2018). Reducing discrimination, especially in a 'masculine' field where women face further obstacles to their careers is imperative for a society that wants to ensure equality. On the other hand, increasing women's participation in STEM jobs is also a critical economic challenge that would have beneficial effects not only for women but for the whole society. The projected growth rate and the increasing number of opportunities in this sector are more than double that for other fields (U.S. Bureau of Labor Statistics 2021). Women are, thus, necessary to fill the surplus of STEM jobs that is likely to further increase in the future. The European Institute of Gender Equality (2017) estimated that closing the STEM gap could lead to an additional 1.2 million jobs and an improvement in GDP per capita by 2.2% to 3% in 2050.

The gender gap in STEM persists in almost all countries of the world, with differences among both countries and fields. Data on BA graduates in OECD countries (2021) suggest that some STEM fields are less unequal than others, i.e., on average in the European Union female graduates are 56% in natural sciences, mathematics and statistics, 26% in engineering, manufacturing and construction, and 19% in ICTs. To some extent, we also observe heterogeneity within each field among countries: the percentage of female graduate students ranges from 29% to 74% in mathematics, from 8% to 41% in engineering, and from 4% to 66% in ICTs. Compared to other countries, in Italy, the percentage of female graduates is slightly above the European average in natural sciences, mathematics/statistics and engineering but is below the European average in ICTs.

Because of the Italian organization of upper secondary schools, a diverging path of girls and boys in this country is already observable in the choice of high school. High schools are either academic or vocational and the former are further classified depending on the field in which they specialize, i.e., humanities (*liceo classico*), science (*liceo scientifico*), art (*liceo artistico*), social sciences (*liceo delle scienze umane*), music (*liceo musicale e coreutico*), modern languages (*liceo linguistico*). Girls represent 82% of students in *licei delle scienze umane*, 78% of students in *licei linguistici*, 70% of students in *licei classici*, and 42% of students in *licei scientifici* (MIUR, 2021). When moving to college, this gap persists and even widens depending on the field of study. As regards STEM, female undergraduates are 14% in IT, 26% in engineering, 31% in physics, while 47% in chemistry and 50% in mathematics. As regards female-dominated fields, male undergraduates are 7% in education, 15% in modern languages, 17% in psychology and 39 % in literary studies (AlmaLaurea, 2021).

Several studies have tried to identify the causes of the STEM gender gap, with explanations ranging from individual characteristics to environmental aspects. Some of these focused on the pervasive presence of gender stereotypes. According to the 'social role theory' (Eagly & Wood, 2012), gender stereotypes derive from the perception of women and men in different social roles and occupations. In the context of STEM, seeing that STEM professionals are more frequently male would induce to associate the STEM sector with a gender, which in turn would influence women's experience in this sector. When asked to perform in a domain culturally associated with men, indeed, women may underperform or decide not to compete at all to avoid failure and judgment (Spencer et al., 1999), a mechanism known as 'stereotype threat' (Spencer et al., 2016). As stated by Dasgupta et al. (2015, p. 4988):

> 'What seems like a free choice is constrained by subtle cues in achievement contexts, such as its sex composition, that signal who naturally belongs in STEM and is likely to succeed and who else is a dubious fit'.

Whilst there is a large consensus on the causes underlying the STEM gender gap, less is known about the reasons underlying the persistence of the gender gap in traditionally feminine sectors, an issue mostly disregarded in both academic research and public debates. However, while in the last years the gender gap in traditionally masculine sectors has narrowed (Martinez & Christnacht, 2021), female-dominated sectors still lag behind. Croft et al. (2015) discussed the asymmetry in gender roles and social status, i.e., the fact that communal (caregiving) roles associated with women and agentic (breadwinning) roles associated with men are traditionally attributed different levels of prestige, the former being considered subordinated and less important and prestigious than the latter. Given these premises they observed (Croft et al., 2015, p. 351):

> 'The asymmetry of changing gender roles can be understood as a manifestation of a more general process whereby lower status groups aspire to possess the traits and attributes associated with those of higher status, whereas higher status groups readily devalue the personal importance of traits and attributes associated with lower status groups.'

Even though a shift of men into female-dominated sectors seems undesirable, still men's choice of communal roles may be beneficial for society, including men themselves and their families (Croft et al., 2015).

Being assessed that gender stereotypes are pervasive and play a pivotal role in the (STEM) gender gap, several studies tested when and how these stereotypes influence women's outcomes in STEM. Results suggest that gender stereotypes influence females' performance in scientific tasks (Cvencek et al., 2015; Kiefer & Sekaquaptewa, 2007; Nosek et al., 2002; Ramsey & Sekaquaptewa, 2011; Smeding, 2012; Steffens et al., 2010), self-concept (Cvencek et al., 2015; Ertl et al., 2017; Nosek et al., 2002; Nosek & Smyth, 2011; Steffens et al., 2010), and attitudes toward math (Nosek et al., 2002; Nosek & Smyth, 2011). There is evidence that they can also affect choices and behaviours, e.g., career aspirations (Schuster & Martiny, 2017; Smyth et al., 2009; Steffens et al., 2010). As regards men, contributions to the theme are scarce and inconclusive. Kalokerinos et al. (2017) found that stereotype threat applies also to men in female-dominated fields, even if to a lower extent compared to what is usually reported for women. When engaged in an upward social comparison with a female worker, male child protection workers were more likely to express turnover intentions, while no effect was found on female workers. On the contrary, Chaffee et al. (2020) did not find evidence that stereotype threat directly impairs men's performance on language tests or suppresses men's belonging in language domains.

Even if stereotypes may be difficult to eradicate, Allport's intergroup contact theory (1954) affirms that under appropriate conditions interpersonal contact can effectively reduce prejudice. This relies on the fact that seeing and being in contact with a minority group in a stereotyped condition can expand people's horizons and reduce prejudices. Following this idea of the malleability of stereotypes, Eagly and Steffen (1984) theorize that gender stereotypes are learned and maintained by people's observations of the unequal distribution of women and men in various social roles. The two theories are the basis of the increasing research on the impact that (counterstereotypical) role models can exert. Being exposed to women who were able to succeed in scientific fields or that pursue a career in STEM disciplines, can transmit a positive message to other women, e.g., 'if someone else who is like me (a female) did it, I can do it too'.

One popular way to expose targets to positive exemplars is to involve them in meetings with women working in the STEM sector, usually referred to as role models or mentors (Townsend, 2002). The results of the studies on the effect of role models on women are mixed. Some have obtained a positive effect, others a negative one, still others no effect at all. The effectiveness of role models changes also depending on the context in which they are applied. Many of the existing studies tested the impact of role models on leadership, while others focused on the impact on girls' relation with science and math. In the latter case, the effects of this type of intervention on female students are on average positive, with some exceptions (Betz & Sekaquaptewa, 2012). For example, role models proved to increase girls' performance (Marx & Roman, 2002; McIntyre et al., 2003), attitudes, and self-efficacy (Cheryan et al., 2013) in STEM disciplines, to guide girls toward the choice of scientific courses (Breda et al., 2018) and also to decrease their stereotypes on women and science (Asgari et al., 2012; Cheryan et al., 2011).

This thesis aims to contribute to the literature studying gender stereotypes, the mechanism through which they affect choices and behaviours, and the efficacy of role models' interventions in solving this issue. It was designed as a collection of papers, each focusing on one aspect of the theme and heterogenous in the type of contribution, still related to the common thread of gender stereotypes in the STEM and humanities sectors. The thesis consists of five papers, of which two are reviews conducted using a systematic approach and three are empirical studies on self-collected data.

As regards the empirical studies, two data collections were conducted, as the first was unfortunately interrupted before completion due to the outbreak of the COVID-19 pandemic in February 2020. Data from the first collection were used in both Chapter 3 and Chapter 4. The original plan was to conduct a field experiment on high school students, randomized at the class level. However, the data collected before school closure in February 2020 was not adequate to be used for the experimental analysis. This is why I decided to conduct a second data collection. However, since students did not go back to school but attended lessons online from home, it was impossible to conduct a field experiment as planned. I adapted the design to an online experiment and randomized students assignment to treatments.

The thesis is organized as follows. The discussion starts from the primary element of the theme, i.e., gender stereotypes. Stereotypes do not have a unique definition, being mostly considered a generalized belief on the quality and characteristics of members of specific groups or social categories. One of the consequences of this is that various scales and measurements have been proposed to assess the endorsement of beliefs on the association between gender and science/liberal arts. **Chapter 2** summarizes, compares and discusses those measures, distinguishing between explicit, implicit and indirect measures. The review of the literature highlighted a huge but unrecognized heterogeneity in the constructs of gender stereotypes, especially for explicit measures. This can hamper findings comparability, reduce scales' validity, affect the correlation between implicit and explicit measurements, and bias their interpretations due to ambiguous terminologies.

Once assessed what gender stereotypes are, the question is whether and how they influence choices and behaviours. Chapter 3 and Chapter 4 present two studies showing, first, that bias on abilities and gender persists and, second, that gender stereotypes are directly associated with educational choices.

Informed by the status characteristics theory (Berger et al., 1972), according to which, when gender is salient, expectations on abilities might reflect gender constructs even when information on individual performance is available, **Chapter 3** tests this hypothesis in a network study on students from ten high school classes in Milan. I asked 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. I did not find any statistically significant results for the math competition. I also found that the female students were less likely to nominate themselves for any competition, regardless of the subject, even controlling for their own performance and self-concept. The study has been submitted to an international journal for peer-reviewed publication, co-authored by Federico Bianchi and Flaminio Squazzoni (University of Milan).

According to Eccles' model, gender stereotypes would indirectly influence major choices by shaping expectations of success and the values attached to the viable options. However, empirical findings on the link between implicit gender-science stereotypes and college major intentions are limited and controversial. To fill this gap, **Chapter 4** examines the association between implicit gender stereotypes and major choices in STEM and humanities, both for male and female high-school students. Logistic regression analysis revealed that implicit gender stereotypes were directly associated with females' intention of majoring in STEM, while the same association could not be confirmed for males' intention of majoring in humanities. Unlike previous findings, the mediation analysis could not confirm that other relevant factors, i.e., interest in the subject, performance at school, self-concept, and value attributed to the job's salary and social utility, moderated this association. The study has been submitted to an international journal for peer-reviewed publication.

The persistent gender imbalance among STEM workers has made women's empowerment a preeminent goal for most countries. Exposing women and girls to female role models is considered pivotal for breaking down gender-stereotypical beliefs on STEM interest and engagement. However, evidence is controversial regarding the efficacy of these interventions. **Chapter 5** provides a scoping review of empirical research on these interventions and assesses their efficacy and potential pitfalls. Results report the characteristics of sixty-eight studies, focusing on their research method, target, intervention, type of role models/mentors, variables of interest, and effects. This review indicates that research is considerably heterogeneous in terms of examples, interventions, variables of interest, and effects. Despite the substantial number of studies, certain topics require further investigation, i.e., interventions targeting preschoolers, a more systematic comparison of role models and mentors, and detection of mechanisms as to why role models sometimes have detrimental rather than beneficial effects. I suggest that more appealing interventions should be designed when targeting young students, e.g., online games and apps. The study has been submitted to an international journal for peer-reviewed publication, co-authored by Gian Luca Pasin and Flaminio Squazzoni (University of Milan).

Finally, **Chapter 6** contributes to the literature on role models by presenting an online experiment conducted on 325 high-school Italian students. Participants were asked to see a video collecting interviews with professionals coming from both the STEM and the humanities-related sectors. They were randomly assigned to one of the following conditions, (1) exposure to counterstereotypical workers (treatment A), (2) exposure to both stereotypical and counterstereotypical workers (treatment B), (3) no information about workers' gender (control). Results were mixed. Female students assigned to treatment A were more likely to perceive some humanistic studies as female-dominated and more likely to believe that engineering is a male-dominated sector. Female students assigned to treatment B were more likely to consider social pressure as more relevant in explaining the gender gap in STEM than biological characteristics. Male students assigned to treatment B were more likely to believe that there is an equal number of women and men in psychological studies and more likely to believe social pressure counts less in explaining the gender gap in STEM than biological characteristics. Those assigned to treatment A were, instead, more likely to believe that there are almost all men in physics-related studies. The study has been submitted to an international journal for peer-reviewed publication.

References

Allport, G. W. (1954). *The Nature Of Prejudice*. Cambridge, Mass: Addison-Wesley Pub. Co.

AlmaLaurea. (2021). *XXIII Indagine Profilo dei Laureati*. AlmaLaurea. https://www.almalaurea.it/universita/indagini/laureati/profilo

Asgari, S., Dasgupta, N., & Stout, J. G. (2012). When do counterstereotypic ingroup members inspire versus deflate? The effect of successful professional women on young women's leadership self-concept. *Personality & Social Psychology Bulletin*, *38*(3), 370–383. https://doi.org/10.1177/0146167211431968

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status Characteristics and Social Interaction. *American Sociological Review*, *37*(3), 241–255. JSTOR. https://doi.org/10.2307/2093465

Betz, D. E., & Sekaquaptewa, D. (2012). My fair physicist? Feminine math and science role models demotivate young girls. *Social Psychological and Personality Science*, *3*(6), 738–746. https://doi.org/10.1177/1948550612440735

Breda, T., Grenet, J., Monnet, M., & Effenterre, C. van. (2018). *Can female role models reduce the gender gap in science? Evidence from classroom interventions in French high schools* (halshs-01713068; PSE Working Papers). HAL. https://ideas.repec.org/p/hal/psewpa/halshs-01713068.html

Chaffee, K. E., Lou, N. M., & Noels, K. A. (2020). Does Stereotype Threat Affect Men in Language Domains? *Frontiers in Psychology*, *11*, 1302. https://doi.org/10.3389/fpsyg.2020.01302

Cheryan, S., Drury, B. J., & Vichayapai, M. (2013). *Enduring Influence of Stereotypical Computer Science Role Models on Women's Academic Aspirations*. https://doi.org/10.1177/0361684312459328

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*(6), 656–664. https://doi.org/10.1177/1948550611405218

Croft, A., Schmader, T., & Block, K. (2015). An Underexamined Inequality: Cultural and Psychological Barriers to Men's Engagement With Communal Roles. *Personality and Social Psychology Review*, *19*(4), 343–370. https://doi.org/10.1177/1088868314564789

Cvencek, D., Kapur, M., & Meltzoff, A. N. (2015). Math achievement, stereotypes, and math self-concepts among elementary-school students in Singapore. *Learning and Instruction*, *39*, 1–10. https://doi.org/10.1016/j.learninstruc.2015.04.002

Dasgupta, N., Scircle, M. M., & Hunsinger, M. (2015). Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. *Proceedings of the National Academy of Sciences*, *112*(16), 4988–4993. https://doi.org/10.1073/pnas.1422822112

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*(4), 735–754. https://doi.org/10.1037/0022-3514.46.4.735

Eagly, A. H., & Wood, W. (2012). Social Role Theory. In *Handbook of Theories of Social Psychology: Volume 2* (Vol. 1–2, pp. 458–476). SAGE Publications Ltd. https://doi.org/10.4135/9781446249222

Ertl, B., Luttenberger, S., & Paechter, M. (2017). The Impact of Gender Stereotypes on the Self-Concept of Female Students in STEM Subjects with an Under-Representation of Females. *Frontiers in Psychology*, *8*, 703. https://doi.org/10.3389/fpsyg.2017.00703

European Institute for Gender Equality. (2017). *Economic benefits of gender equality in the EU: How gender equality in STEM education leads to economic growth.* Publications Office. https://data.europa.eu/doi/10.2839/652355

Funk, C., & Parker, K. (2018). *Women and Men in STEM Often at Odds over Workplace Equity* (p. 158). Pew Research Center. https://vtechworks.lib.vt.edu/bitstream/handle/10919/92671/WomenSTEEMWorkplace.pdf?sequence=1&isAllowed=y

Kalokerinos, E. K., Kjelsaas, K., Bennetts, S., & Hippel, C. von. (2017). Men in pink collars: Stereotype threat and disengagement among male teachers and child protection workers. *European Journal of Social Psychology*, *47*(5), 553–565. https://doi.org/10.1002/ejsp.2246

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*(5), 825–832. https://doi.org/10.1016/j.jesp.2006.08.004

Martinez, A., & Christnacht, C. (2021). *Women Making Gains in STEM Occupations but Still Underrepresented.* U.S. Bureau of Labor Statistics. https://www.census.gov/library/stories/2021/01/women-making-gains-in-stem-occupations-but-still-underrepresented.html

Marx, D. M., & Roman, J. S. (2002). Female Role Models: Protecting Women's Math Test Performance. *Personality and Social Psychology Bulletin*, *28*(9), 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*(1), 83–90. https://doi.org/10.1016/S0022-1031(02)00513-9

MIUR. (2021). *Studenti per anno di corso, classe e genere. Scuola statale anno scolastico*. https://dati.istruzione.it/opendata/opendata/catalogo/elements1/leaf/?area=Studenti&datasetId=DS0030ALUCORSOINDCLASTA

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*(1), 44–59. https://doi.org/10.1037/0022-3514.83.1.44

Nosek, B., & Smyth, F. L. (2011). Implicit Social Cognitions Predict Sex Differences in Math Engagement and Achievement. *American Educational Research Journal*, *48*(5), 1125–1156. https://doi.org/10.3102/0002831211410683

OECD. (2021). *Education at a Glance 2021 [Share of graduates by gender in fields of education]*. OECD. https://stats.oecd.org/index.aspx?queryid=109881#

Ramsey, L. R., & Sekaquaptewa, D. (2011). Changing stereotypes, changing grades: A longitudinal study of stereotyping during a college math course. *Social Psychology of Education*, *14*(3), 377–387. https://doi.org/10.1007/s11218-010-9150-y

Schuster, C., & Martiny, S. E. (2017). Not Feeling Good in STEM: Effects of Stereotype Activation and Anticipated Affect on Women's Career Aspirations. *Sex Roles*, *76*(1–2), 40–55. https://doi.org/10.1007/s11199-016-0665-3

Smeding, A. (2012). Women in Science, Technology, Engineering, and Mathematics (STEM): An Investigation of Their Implicit Gender Stereotypes and Stereotypes' Connectedness to Math Performance. *Sex Roles*, *67*(11–12), 617–629. https://doi.org/10.1007/s11199-012-0209-4

Smyth, F. L., Greenwald, A. G., & Nosek, B. A. (2009). *Implicit gender-science stereotype outperforms math scholastic aptitude in identifying science majors* [Unpublished manuscript]. https://www.colorado.edu/odece/sites/default/files/attached-files/rba13-sb4converted_2.pdf

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype Threat. *Annual Review of Psychology*, *67*(1), 415–437. https://doi.org/10.1146/annurev-psych-073115-103235

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, *35*(1), 4–28. https://doi.org/10.1006/jesp.1998.1373

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology*, *102*(4), 947–963. https://doi.org/10.1037/a0019920

Townsend, G. C. (2002). People who make a difference: Mentors and role models. *ACM SIGCSE Bulletin*, *34*(2), 57–61. https://doi.org/10.1145/543812.543831

UNESCO. (2017). *Cracking the code girls' and women's education in science, technology, engineering and mathematics (STEM).* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000253479

Vincent-Lancrin, S. (2008). The Reversal of Gender Inequalities in Higher education: An On-going Trend. In *Higher Education to 2030*. Centre for Educational Research and Innovation, OECD.

Chapter 2. ***The conundrum of gender-science stereotypes: A review and discussion of measurements***

1. Introduction

In the last years, gender stereotypes have been acknowledged a crucial role in determining and contributing to the underrepresentation of women in STEM. Several studies have concurrently attempted to test their effect on girls' and women's aspirations, performance, interests and sense of belongingness in this field (Cundiff et al., 2013; Kiefer & Sekaquaptewa, 2007a, 2007b; Lane et al., 2012; Nosek & Smyth, 2011; Reuben et al., 2014). On the other hand, others have tested how these gender stereotypes can be effectively reduced. Interventions with this purpose include, for instance, exposing girls to counterstereotypical role models (Betz & Sekaquaptewa, 2012; Dasgupta & Asgari, 2004; McIntyre et al., 2003) and making them aware of the detrimental influence of stereotypes' endorsement (Farrell et al., 2020; Jackson et al., 2014; Johns et al., 2005).

Despite being similar in the addressed issue, these studies show great variability in the target population (children, adolescents, adults), the variables of interest, the setting (laboratory or field) and the research design. This heterogeneity may explain the lack of a unique and shared scale to measure gender-science stereotypes. Conversely, scales of this type exist for stereotypes on other gender-related issues, e.g., the Attitudes toward Women Scale (Spence & Helmreich, 1972) for gender roles and the Ambivalent Sexism Inventory (Glick & Fiske, 1996) for sexism.

Furthermore, stereotypes do not have a unique definition and the lack of a univocal meaning may also explain this heterogeneity in measurements. In their review of instruments for gender roles, McHugh and Frieze (1997) blamed the proliferation of scales that make any comparison difficult. While the existence of several different instruments is a problem, the multi-facet nature of gender-related characteristics and gender-belief systems would require avoiding single measurements. Indeed, gender stereotypes address several issues, i.e., traits, attitudes, interests, cognitive skills, family roles and occupations (Hentschel et al., 2019; Six & Eckes, 1991).

Given the existence of such a multitude of indicators, previous articles reviewed and synthesized scales and other types of measures used to assess gender stereotypes on roles (Beere, 1990; McHugh & Frieze, 1997), traits and attitudes (Kite et al., 2008; Smiler, 2004) but, to the best of my knowledge, only Zitelny et al. (2017) grouped studies on gender-science stereotypes and listed the instruments used. However, their aim was not to review these instruments, but rather to focus the attention on the correlation between implicit and explicit measures. There is a need for filling this gap and comparing measures of stereotypes on gender and STEM. This is the first aim of this paper.

The second part of the article will discuss potential consequences deriving from the heterogeneity of these instruments and some of their limitations. Expanding on the problem of findings comparison argued by McHugh and Frieze (1997), it is suggested here that the proliferation of scales affects also the correlation between implicit and explicit measures (Zitelny et al. 2017). As regards potential limitations of existing instruments, it is argued that while the focus of current studies is typically on math, also other scientific fields should be investigated. Moreover, since some questions behind measurements are often too generic so leaving space for interpretability, there is a need for understanding how this can compromise the interpretation of final scores.

Once acknowledged the increasing interest in gender stereotypes and the need to understand the manyfold role they play in STEM, this study could be beneficial for research on the theme in two ways. On the one hand, researchers have access to a general overview of the instruments available to test gender stereotypes. The summary of results facilitates the identification of instruments and their psychometric soundness while showing also their diffusion in previous studies. This could help reduce the tendency to create new *ad hoc* measures. On the other hand, the discussion about the instruments' limitations set the ground for a refinement in the measurement of gender stereotypes and suggests an unexplored field of research on the theme.

2. Literature review

2.1. Stereotypes and gender

The word ‘stereotype’ was first used by the journalist Lippman (1922) to indicate general cognitive structures that serve as mental pictures of social groups. However, since then, the meaning and definition of stereotypes have changed and evolved (for a review, see Schneider 2005).

Several definitions of stereotypes exist, differing in whether they describe stereotypes as inaccurate, consider stereotypes disagreeable in both the formation process and the consequences and represent stereotypes as shared among people or as individual beliefs (Schneider, 2005). To mention just a few, stereotypes were defined as ‘Beliefs and opinions about the characteristics, attributes and behaviours of members of various groups’ (Hilton & von Hippel, 1996, p. 240), as ‘Both positive and negative beliefs or overgeneralizations about the attributes of a group and its members’ (Marx & Ko, 2012, p. 160), and more recently as ‘General expectations about members of particular social groups […] that leads people to overemphasize differences between groups and underestimate variations within groups’ (Ellemers, 2018).

The social groups affected by stereotypes are various. Early research focused on stereotypes about race and ethnicity, while, starting from the 1970s and 1980s, the widespread interest in the discrimination against women led to an expansion in research on gender stereotypes (Schneider, 2005). There are several beliefs on gender differences, ranging from characteristics to roles, and, consequently, gender stereotypes consist of multiple components. This paper focuses on the belief that women and men would differ in their mathematical and scientific abilities, with men traditionally considered to outperform women in STEM.

2.2. Gender-science stereotypes

The belief that women would perform poorly in STEM and, conversely, that STEM would be the natural domain of men traditionally derived from the unfounded conviction that women's and men’s brains differ, the latter being more apt to logical thinking (Kersey et al., 2019). Furthermore, the observation of women’s underrepresentation in STEM (Eagly & Wood, 2012) and the association of success in these fields with being agentic – a characteristic traditionally attributed to men (Sczesny et al., 2018) – contributed to the reinforcement of beliefs on gender differences not only in abilities but also in interests and aptitudes (e.g., Plante et al., 2009).

It is still unclear when boys and girls start endorsing gender-science stereotypes. On the one hand, to a certain age children tend to consider their gender the smartest (Grow et al., 2016) – ingroup favouritism – on the other hand, in some studies there was evidence that children associated maths with boys by the age of six (Master et al., 2017; Tomasetto et al., 2012).

In the last years, research on gender stereotypes in this context has deeply increased and numerous studies found evidence of the detrimental effect of gender stereotypes on women in STEM. To mention some, Cundiff et al. (2013) found that among college students, women endorsing stronger gender–science stereotypes had weaker science identification and, in turn, weaker science career aspirations. Kiefer and Sekaquaptewa (2007a) found a negative association between stereotypes and performance. Female students with low implicit gender stereotypes performed better in a calculus course in college compared to those with stronger implicit stereotypes. Finally, Nosek and Smyth (2011) found that stronger implicit gender-science stereotypes predicted women's higher negativity toward maths, lower participation in STEM, and worse achievement in maths.

The relevance of addressing the issue of stereotypes is widely recognized, to the extent that the Committee on the Elimination of Discrimination against Women (CEDAW) – an international human rights treaty – regulates states' obligations to address stereotypes and stereotyping affecting women (Cusack, 2013). Given these premises, it is not difficult to recognize the importance of using valid instruments, especially, when testing strategies that can potentially be applied on a larger scale.

2.3. Instruments to measure gender-science stereotypes

In the context of stereotypes on gender and STEM, it is not possible to identify a widely adopted instrument assessing stereotypes endorsement. A subscale of the Fennema-Sherman Mathematics Attitudes Scales (1976), the Mathematics as a male domain scale, could have fulfilled this role. However, despite further refinement and validations of this scale (Leder & Forgasz, 2002) researchers have tended to create new instruments, with fewer items, shaped on the aim of their study, rather than adopting the existing scales. This resulted in a proliferation of heterogeneous measurements.

The lack of such scales has several negative consequences. As observed by McHugh and Frieze (1997, p. 3) for gender roles, 'When each researcher […] develops his/her own scale, it becomes increasingly difficult to make comparisons across studies, across samples, across cultures and over time. It is unlikely that each researcher has developed a valid and reliable measure, and even more unlikely that each is measuring a unique, enduring, and important construct'. The existence of multiple instruments measuring the same constructs requires and justifies reviews summarizing and reporting them. While several reviews were published on gender-role attitudes and sexism (Beere, 1990; Kite et al., 2008; McHugh & Frieze, 1997; Smiler, 2004), to my knowledge there is a lack of reviews on gender-science stereotypes.

A partial summary of these instruments is reported in the appendix of the study by Zitelny et al. (2017). They aimed to analyse the correlation between implicit, i.e., the Implicit Association Test (IAT), and explicit measures of gender-science stereotypes to suggest that the former should not be interpreted as a counterpart of the latter. In the appendix, they summarized twenty-four studies in which both explicit and implicit measures were used and reported the correlation between the two. The authors suggested that the observed heterogeneity in the correlation between the two measures may be due to the use of different self-reported instruments.

They distinguished between self-reported beliefs (about natural ability, natural interest, and prevalence), and self-reported association, i.e., the extent to which participants associated science with males versus females, and liberal arts with males versus females. They then discussed the relation between the two and the IAT scores. Results indicated that, among beliefs and self-reported association, the latter is the one that correlates with the IAT the most. This suggests that 'the IAT taps into different constructs than those tapped by the explicit measures used in research on the gender-science stereotype' (Zitelny et al., 2017, p. 6). Consequently, the authors suggest that a distinction among constructs of stereotypes and a more specific choice of one over the other may be relevant when both explicit and implicit measurements are used.

Expanding the summary table in Zitelny et al. (2017), the current study reports a comprehensive overview of papers including an instrument of gender-science stereotypes' endorsement. More specifically, I included explicit, implicit and indirect instruments (Whitley & Kite, 2016). The first refers to instruments based on participants' self-reports, the second to those measuring the mental association among concepts and the third to instruments that, as explicit measurements, asked participants about opinions or beliefs. However, unlike previous ones, in the latter, concepts are only indirectly linked to gender stereotypes.

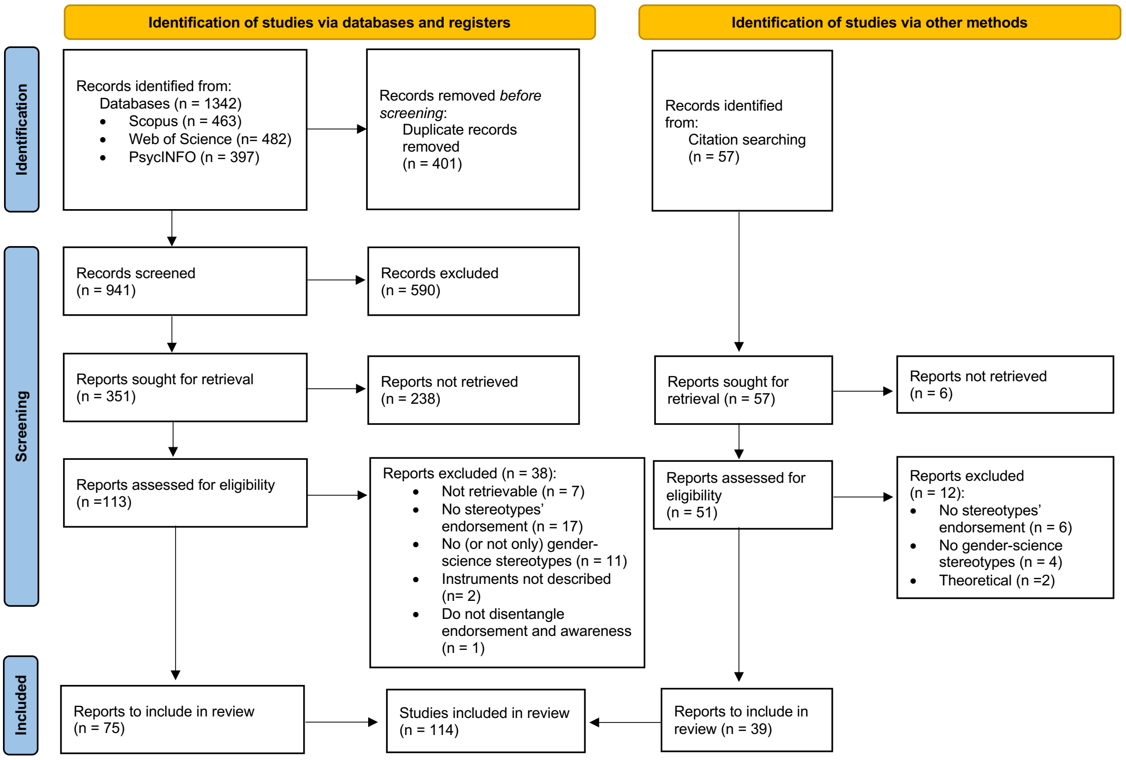
The distinction between implicit and explicit measures is not limited to whether they assess automatic or self-reported beliefs. Explicit instruments mostly ask participants' opinions about the single gender-STEM association. Conversely, implicit indicators, e.g., the IAT, mostly provide a final score computed as the difference between the gender-STEM and the gender-humanities automatic associations. As such, 'the IAT is limited to measuring the relative strengths of pairs of associations rather than absolute strengths of single associations. In practice, however, the IAT can nevertheless be effectively used because many socially significant categories form complementary pairs' (Greenwald & Farnham, 2000, p. 1023).

Being this distinction between 'complementary stereotypes' (Jost & Kay, 2005) potentially relevant when comparing and discussing instruments of stereotypes (Gilbert et al., 2015), here it is also reported when and how studies investigating gender-science stereotypes included a measure of the gender-humanities association. Note that here both 'humanities' and 'language' terms are used when referring to this complementary association. This is because studies investigating the association between gender and careers/majors usually refer to humanities (or liberal arts), while those investigating the association between gender and abilities refer to language-related, writing or reading skills (see Table 2.1 and Table 2.2).

3. Methodology

Figure 2.1 summarizes the procedure followed for the selection of studies. First, a combination of key terms – (gender AND stereotyp* AND (STEM OR math* OR scien* OR engineer* OR techn*) AND (instrument* OR measur*)) – was searched in relevant databases (Web of Science, PsycINFO, Scopus) in January 2021 and arranged according to the rules of the databases. After a screening of the reports resulting from this first stage, relevant references cited in those reports were then screened. This second stage of the selection process was pivotal to backtrack to the source who proposed the instrument in the first place.

Figure 2.1 *Prisma flow diagram*



To be eligible for inclusion, studies should have used an instrument to measure the endorsement of stereotypes on gender and STEM-related domains. The eligibility criteria were left wide on purpose, as one of the aims of this review is to highlight the heterogeneity in instruments and how gender stereotypes are described in the studies. Consequently, all studies stating to measure gender stereotypes in STEM were included in the review (e.g., gender-science stereotypes, gender-math stereotypes), while those investigating stereotypes' awareness rather than their endorsement were excluded. Furthermore, studies using the Draw-a-scientist test were excluded, as these have already been summarized in recent reviews (Ferguson & Lezotte, 2020; Miller et al., 2018). Finally, studies not in English or Italian were excluded.

One hundred fourteen studies, published from 1993 to the end of 2020 resulted eligible for being included in the review. The instruments used were classified into three macro-categories, i.e., explicit, implicit and indirect. Explicit instruments were, then, evaluated based on the construct measured. In particular, eight types of constructs were identified based, in part, on the classification proposed in previous studies (Nurlu, 2017; Zitelny et al., 2017).

- *Skills*: instruments asking participants to evaluate skills, abilities, or brain differences between men and women.
- *Conformance*: instruments asking participants to give their opinion about the need for women or men to conform to the opposite gender’s behaviour/attitudes in the domain of interest.
- *Gendered domain*: instruments generically referring to masculinity or femininity of domains, without specifying whether this refers to abilities, interests, or other characteristics.
- *Interest*: instruments investigating differences in interest.
- *Relevance*: instruments asking participants an opinion on the relevance the domains of interest have for people.
- *Gender imbalance*: instruments asking participants to evaluate the representativeness of women and men in occupations, and academic courses.
- *Suitability*: instruments asking participants explicitly an opinion on the suitability of people in certain domains.

Other information extracted from the selected studies were the domains of interest (e.g., science, maths) and whether and how they measured the association between gender and humanities or language-related skills. Implicit measures were evaluated based on (1) the type of test, and (2) the type of target and categories used (Whitley & Kite, 2016), while indirect measures were too few to be further classified.

4. Results

4.1. Explicit measurements

Most studies included in this review adopted an explicit measure to assess gender-science stereotypes’ endorsement. Statements were usually evaluated using Likert scales, the most popular being the 5-point Likert scale.

Table 2.1 summarizes the characteristics of the explicit measurements. The most frequently investigated construct is that relative to skills, followed by gendered domain, gender imbalance and suitability. However, there are differences even within constructs. These differences regard, in some cases, the content of the items, while in others, the question and the phrasing used for statements (see Table 2.3 in the Appendix).

Table 2.1 *Characteristics of explicit instruments*

Characteristic	Instruments (% of the total)
Construct	
Skills	67%
Gendered domain	19%
Interest	12%
Suitability	10%
Gender imbalance	10%
Attribution	10%
Conformance	8%
Relevance	8%
Domain of interest	
Maths/numbers/calculus	67%
Science	25%
Computing/Programming/ICT	12%
Engineering/mechanical	12%
Physics	9%
Technology/technical	8%
Geometry/mental rotation/spatial	7%
STEM	7%
Chemistry	4%

Analytic/reasoning/logic	4%
Nature/Geography	2%
Astronomy	1%
Both STEM and non-STEM domains	38%
Non-STEM domains	
Language	13%
Native language (e.g., English)	8%
Other	8%
Reading	7%
Arts	7%
Liberal Arts	4%
Humanities	3%
Writing	2%

Note. Total number of explicit instruments: 91

Total number of studies using an explicit instrument: 104

As regards *skills*, items ask participants to what extent they agree with the belief on the outperformance of men in the STEM field, either directly or indirectly. In other cases, it is asked to rate who is better between men and women, leaving thus the possibility to also detect cases in which the association is even reversed. Finally, in some instruments the cause underlying gender differences is explicitly mentioned, e.g., ‘Men are naturally better at advanced math (mechanical things) than women’ (Riegle-Crumb & Morton, 2017). It is noteworthy that, in one study, researchers showed an explicit interest in distinguishing between descriptive and prescriptive stereotypes (McGuire et al., 2020), the first referring to beliefs on what people do, e.g., ‘Who do you think is usually good at […]’, the second on what people should do, e.g., ‘Who do you think should be good at […]’.

As regards *gendered domain*, questions are similar one to the other and ask participants to rate how much they associate a list of domains with males or females (see, for example, Greenwald et al., 2003; White & White, 2006; Young et al., 2013), or whether they agree with statements such as ‘Math is rather a typical subject for girls (boys)’ (Steffens et al., 2010). Similarly, to assess the opinion on the *representativeness of women and men*, researchers usually ask participants either to estimate the percentage of male and female workers in certain occupations or to provide their agreement with statements on this representativeness, e.g., ‘There are more men in science-related jobs’ (Breda et al., 2018).

In some cases, participants are asked to compare the *suitability* of women and men to STEM or humanities fields or to give their opinion on the better suitability of men to STEM fields, e.g., ‘It is possible that men are better suited to studying at the technical university than women’ (Jasko et al., 2019). As for other constructs, questions on gender differences in *interest* ask participants their opinion on the higher interest in STEM of men, e.g., ‘Boys (girls) are more interested in careers which require mathematical ability than girls (boys) are’ (Nurlu, 2017). Similarly, questions on the *relevance* of STEM for men and women, asked, for example, ‘It is more important for boys to understand physical science than girls’ (Buck et al., 2002). On the contrary, instruments assessing *conformance* are quite different from one another. Ertl et al. (2017) generically asked whether ‘Females that are working in the field of STEM have to be like men’, Betz and Sekaquaptewa (2012) if ‘Do being good at math and being girly go together?’, while both Plante et al. (2009) and Nurlu (2017) asked about the association between popularity and abilities in STEM or reading.

In some studies, participants were directly asked their opinion on the potential explanations for gender differences. Contrary to other constructs, in the case of *attribution*, the existence of gender differences is taken for granted. The interest is in verifying whether participants are more likely to attribute the gender gap in STEM to biological rather than cultural and social factors, e.g., ‘Boys (girls) are encouraged more than girls (boys) to choose a career in a math-related area’ (Nurlu, 2017), ‘Males perform better than females in science because of greater natural ability’ (Nosek et al., 1998).

Most instruments asked participants for an opinion on ‘Maths’, 25% of them an opinion on ‘Science’, while in a minority of cases instruments mentioned other STEM-related fields (see Table 2.1). In some cases (38%) instruments asked an opinion on both STEM and non-STEM fields, the latter being specified in different ways, e.g., ‘Language’, ‘Liberal Arts’, ‘Humanities’. Similarly to implicit measurements’ scores, some authors computed the final score as the difference between the answers given on STEM domains and those given on non-STEM domains. For instance, del Rio et al. (2019, 2020) asked participants ‘Please rate how much you associate mathematics with males or females’ and ‘Please rate how much you associate language with males or females’. The math-gender stereotype scale was then computed as the difference between the two items.

Others followed a similar procedure but computed the final scale as the difference between the average male score and the average female score. For example, Steffens et al. (2010) asked participants to what extent they agreed with the statements ‘Boys are often talented for doing math’ and ‘Math is rather a typical subject for boys’. They then asked the same question referring to girls (e.g., ‘Girls are often...’). The same four questions were then asked using German as the subject. Two scales were computed, one measuring the math-boy stereotype, i.e., the difference between math items on girls and math items on boys, and the other measuring the language-girl stereotype, i.e., the difference between the language items on girls and the language items on boys.

4.2. Implicit measurements

While there was far less heterogeneity in the type of implicit measurements generally used to test stereotypes compared to the explicit ones, still there is some variability. The Implicit Association Test (Greenwald et al., 1998) is the most popular measurement when studying (gender) stereotypes. It was designed by Greenwald, McGhee and Schwartz (1998) to measure individual differences in implicit cognition. This is done by measuring the difference in the time needed to do an association between compatible constructs (e.g., women and humanities, men and STEM) and the time needed to do an association with incompatible constructs (e.g., women and STEM, men and humanities).

However, other tests similar to the IAT were used, i.e., the Implicit Relational Assessment Procedure, IRAP (Barnes-Holmes et al., 2006), the Affect Misattribution Procedure, AMP (Payne et al., 2005), the Go/No-Go Association Task, GNAT (Nosek & Banaji, 2001) and the Sorting Paired Feature Task, SPF (Bar-Anan et al., 2009). Contrary to the IAT, the AMP, IRAP and GNAT allow disentangling the two tested associations. Further details on implicit measures can be found in Gawronski and De Houwer (2013) and an application in the context of stereotypes in Whitley and Kite (2016).

Table 2.2 summarizes the types and characteristics of implicit instruments. More details can be found in the Appendix (Table 2.4).

Table 2.2 *Characteristics of explicit instruments*

Characteristic	Instruments (% of the total)
Type of test	
IAT	52%
GNAT	14%
Child-IAT	10%
IRAP	10%
AMP	5%
SA-IAT	5%
SPF	5%
Domain of interest	
Math	48%
Science	24%
STEM	14%
Engineering	10%
Space	5%
Spatial	5%
Both STEM and non-STEM domains	95%
Non-STEM domains	
Language	19%
Arts	14%
Reading	14%
Liberal Arts	14%
Humanities	10%
English	10%
Other	10%

Language Arts	5%

Note. Total number of implicit instruments: 21

Total number of studies using an implicit instrument: 55

Similarly to what was observed for explicit instruments but to a lower extent, most instruments used ‘Maths’ as a category while some used ‘Science’. However, in the most adopted version of the IAT, *stimuli* were STEM-related majors. As regards non-STEM fields, the choice of the category was more heterogeneous, as *stimuli* referred to either ‘Language’, ‘Arts’, ‘Reading’, or ‘Liberal Arts’. In one study (Guizzo et al., 2019) only one of the two associations, i.e., gender and space-related concepts, was tested.

4.3. Indirect measurements

Seven studies created and applied indirect instruments (see Table 2.5 in the Appendix). They differ from explicit and implicit measures because there is not an explicit reference to gender. Participants were usually shown two (or more) pictures, one showing a man/boy and the other showing a woman/girl, and asked which one possessed some characteristics, e.g., interest and giftedness in math (Nurnberger et al., 2016). However, two instruments stand out. In Tomasetto, Galdi and Cadinu (2012), children were told a story about an island where inhabitants would not consider boys and girls equally good in school subjects. At the end of the story, participants were asked whether, in their opinion, the inhabitants of the island considered boys or girls better in math. While, in Ambady et al. (2001), participants were asked to repeat a brief story about a student good in math and the experimenter noted whether they used the pronoun ‘he’ or ‘she’ when appointing the student.

5. Consequences of instrument heterogeneity

This review suggests that measurements of gender-science stereotypes show great heterogeneity on a variety of features, i.e., underlying constructs, domains of interest (e.g., science, math, reading, native language), types of scale and number of items, types of stereotypes (descriptive or prescriptive), age of the people on which the belief was asked (children, adolescents, adults), whether the opinion regarded school subjects, majors or occupations, number and type (stereotypical and/or counterstereotypical) of associations. This variability could have important implications. As mentioned before, some of these implications were already discussed for instruments on gender-roles stereotypes by McHugh and Frieze (1997).

Compared to gender roles, the problem with gender-science stereotypes is even more complicated. Indeed, in this case, considering proper scales is impossible because the set of items in the questionnaire was in most studies not chosen following a development process (Kyriazos & Stalikas, 2018). Rather, researchers have tended to create *ad hoc* statements, evaluating them using Likert scales. Assumptions and hypotheses behind items' selection were not always reported (more details in the Appendix). Hence, the problem of validity and reliability of measurements, discussed in McHugh and Frieze (1997), applies even more in this context. However, comparability and reliability of instruments are not the only cons of scale proliferation.

As mentioned before, Zitelny et al. (2017) suggested that the variability in the correlation between explicit measures and the IAT may be due to the use of different self-reported instruments. Therefore, a distinction among constructs of stereotypes and a more specific choice of one over the other may be relevant when both explicit and implicit measurements are used. However, other features of explicit instruments may affect the correlation with the implicit ones. As anticipated, implicit measures' scores, especially for the IAT, are the result of two stereotypical associations, one of men and science, the other of women and humanities. However, only 38% of explicit measurements tested both associations and only in a few cases the final score was computed as the difference between the two (e.g., Liu et al., 2010; Rentas, 2015). For instance, Plante and Theoret (2009) created two scales, i.e., the male domain scale and the female domain scale, each including 16 items on abilities, usefulness, attitudes, typicality, effort and support in both math and language. The final score of gender stereotypes was calculated by subtracting the mean scale score of the female domain from the male domain's mean scale score. This distinction is relevant as the reversal of the typical stereotype emerged from the analysis. As suggested by the authors, 'it appears that stereotypes favouring girls in mathematics can emerge when the instrument that is used allows this possibility' (Plante et al., 2009, p. 398). Consequently, implicit and explicit scores reflect two different things.

6. Limitations of existing instruments

This review found certain weaknesses caused by the way items were constructed. The first limitation regards the domain associated with gender. As mentioned above, in most cases questions asked an opinion about mathematics. However, women's underrepresentation does not characterize only math-related areas, but in general the entire scientific field. Indeed, when accounting for the discontinued education and career paths of females in the scientific field, researchers refer to a STEM leaky pipeline, not a math leaky pipeline (Grogan, 2019).

Among the reviewed studies, while those interested in the association between gender and occupations usually referred to STEM or science-related careers, those interested in the association between gender and school subjects referred more frequently to maths. If the aim is to assess medium- and long-term effects rather than attesting to the mere existence of stereotypes' endorsement, extending the domain to include also science could be more appropriate. Note that the gender gap is narrower in math majors compared to other scientific fields of study. In 2020, the percentage of females among Italian undergraduate students was equal to 50% in mathematics. Conversely, female undergraduates were 14% in IT, 26% in engineering, and 31% in physics (AlmaLaurea, 2021).

Furthermore, research tended also to disregard the association between gender and reading skills or the corresponding humanities field. As mentioned above, ignoring such an association may be problematic when the implicit association test is combined with explicit measurements. Moreover, the gender gap in humanities-related majors is at least as wide as in STEM majors. In 2020 in Italy, the percentage of male undergraduate students was 7% in education, 15% in modern languages, 17% in psychology and 39 % in literary studies. As suggested by Plante and Theoret (2009) to justify the reversal of beliefs about female math abilities in their sample, initiatives aimed at reducing the underrepresentation of women may contribute to a change in stereotypical beliefs on mathematics, but not on other domains.

Some instruments, especially those on the gendered domain, are based on quite generic statements. In particular, those asking to rate to what extent a certain domain would be feminine or masculine, leave the respondent the freedom to choose which aspects of gender differences the questionnaire is referring to. Femininity may derive from representativeness, which would imply that a domain is feminine when more women than men are working in that area, the opposite for masculine. On the other hand, when considering ability, the domain would be labelled as feminine when the respondent believes that women are better than men in that specific domain. Leaving too much space for interpretations can bias results, as assessing whether the final score would measure the same thing for all individuals in the sample would be impossible.

An issue that is not taken into account by studies on the theme is the gender stereotypes' attribution. Few items proposed in Nurlu (2017) distinguished between differences due to natural abilities and differences due to discrimination, e.g., 'Compared to boys (girls), girls (boys) mostly increase their mathematical achievement, because of the support of their teachers'. However, the final scale does not distinguish these from the other items. Project Implicit, a website on which users can perform several IATs, including the gender-science IAT, asked users their agreement on potential explanations for the STEM gender gap, e.g., 'On average, men and women differ in their willingness to devote the time required by such high-powered positions', 'Directly or indirectly, boys and girls tend to receive different levels of encouragement for developing scientific interest'. However, it seems from the review that the implications of making explicit the different causes for the gender gap have not been investigated.

This refinement could be quite relevant in the context of gender-science stereotypes, as initiatives aimed at reducing them are frequently based on the exposure to role models (Betz & Sekaquaptewa, 2012; Gilbert, 2015; Van Camp et al., 2019; Young et al., 2013). Further research would be necessary to understand whether attributions to gender differences in science may affect the efficacy of the exposure to role models. If we believe the gender gap is due to biological, innate differences, we will be less likely to modify our opinion even when evidence of equality of performance is provided. On the contrary, if other justifications are given for the gender imbalance in the sector, information on inaccuracy is likely to change previous beliefs.

Since the data collected on the Project Implicit website includes questions on both gender stereotypes' attribution, rating of masculinity and femininity of the STEM sector and an implicit indicator of gender stereotypes, a comparison of the three types of instruments could give some insights on the above-mentioned issue. To avoid overloading the chapter, I reported an analysis of those data in the Appendix. Results from a structural equation model analysis revealed that attributions to the gender gap in science are related to two distinct factors, which I called 'personal' and 'social', the first including causes pertaining to individual characteristics (e.g., biological differences), the latter including causes referring to others' discrimination and behaviours. Furthermore, when looking at the association between the two factors and explicit and implicit indicators of gender stereotypes, I found that the attribution of masculinity to the STEM domain was associated with the social component of attribution for women, while for men it was associated with the personal component. This would suggest that when questions on the association between gender and science are ambiguous, the deriving scales can measure different constructs of gender stereotypes, at least in part related to the respondent's gender.

Partially related to stereotypes' attribution, is the distinction between descriptive and prescriptive stereotypes. Descriptive stereotypes can be defined as the description of what group members are typically like – e.g., STEM is a masculine sector – while prescriptive stereotypes are the description of the behaviour group members should uphold to avoid derision – e.g., STEM should be a masculine sector (Gill, 2004). Solid reasons would suggest that making distinctions and refining what is generically called a stereotype is beneficial as descriptive and prescriptive beliefs might have different effects on outcomes. For instance, in a laboratory experiment on sex discrimination in hiring, Gill (2004) found that descriptive stereotypes did predict gender bias in neither the choice of job applicants nor the evaluation of candidates. On the other hand, results indicated that prescriptive stereotypes fostered a bias among male participants against females enacting a masculine role. Similarly, in their review of the literature on descriptive and prescriptive stereotypes in sex discrimination and sexual harassment, Burgess and Borgida (1999) called for clear-cut distinctions between these two components as they resulted in different types of sex discrimination. Descriptive stereotypes would lead to an unintentional form of discrimination, which may be modified when information on the inaccuracy of the gender bias is provided. Prescriptive stereotypes lead to a stronger form of discrimination and prejudice, which is not dented by any information. Unfortunately, little is known about the effect of these two components in the specific area of gender bias in STEM and humanities. A remarkable exception is McGuire et al. (2020) who collected information on three distinct types of stereotypes, i.e., awareness (descriptive component, e.g., 'who do you think is usually good at…'), endorsement (prescriptive component, e.g., 'who do you think should be good at…') and flexibility (e.g., 'who do you think can be good at…'). However, the authors did not distinguish among the three components when reporting the effect of gender stereotypes on aspirations, performance, and other outcomes of interest.

7. Discussion and conclusions

Greater attention has been recently devoted to stereotypes and their influence on gender issues. Moreover, these are on top of the political agenda in several countries (Cusack, 2013). Academic researchers have, concurrently, conducted empirical studies testing the effect of gender stereotypes on women's engagement in STEM. This has stimulated the creation of scales and other instruments to measure gender stereotypes. However, compared to other fields of gender bias, there is a lack of properly developed scales to assess associations between gender and science/humanities. This led to a proliferation of instruments, which in turn can explain the variability of findings and certain terminological ambiguities.

The current review extended that of Zitelny et al. (2017) on instruments for gender stereotypes in science and summarized implicit, explicit and indirect measures adopted by researchers in 114 articles. Explicit measures were classified based on the underlying construct of stereotypes as follows: attribution to gender differences; conformity to behaviours and attitudes of the prevalent sex in the field; masculinity/femininity of the domain; interest in the subject; representativeness of men and women in the sector; suitability in the domain; performance in the subject; and relevance attributed to the subject by men and women. Research prevalently identified stereotypes with differences in abilities in maths/science and reading. However, the instruments differed in several features, such as type of ability, domain investigated, and type of scale.

The summary of implicit indicators detected a certain degree of heterogeneity, though less than expected. The most popular test is the Implicit Association Test, yet different versions are used. These versions can be distinguished on the type of categories and stimuli adopted to design the test, i.e., words related to majors, occupations, or features of STEM and humanities. As regards indirect measurements, they are mostly adopted when testing stereotypes on children and ask them to associate boys and girls seen in pictures with characteristics related to maths/science and language.

I then discussed certain pitfalls due to the heterogeneity and proliferation of scales. First, the adoption of indicators varying in multiple aspects can eventually invalidate findings' comparison and scale reliability (McHugh & Frieze, 1997). This heterogeneity may also explain the variability in the correlation between explicit and implicit measures (Zitelny et al., 2017). While the IAT score is the difference between two associations, i.e., male and science vs female and humanities, explicit measures usually address either one of the two associations or do not construct the final score as a difference. Furthermore, some of the revised instruments may suffer certain potential limitations. In particular, most instruments focus on math, which leaves aside other scientific fields (e.g., science, technology and engineering) and language-related domains, which are affected by stereotypical beliefs as well.

Another limitation regards more specifically the way questions on the masculinity and femininity of science and humanities are posed. Questions of this type are generic and do not specify what is meant by masculinity or femininity, thereby leaving the interpretation to the respondent. This impairs the final scores of whatever instruments, as they may assume different meanings. Furthermore, further refinement of instruments may shed light on the difference between the attribution of gender-science stereotypes to either biological or social/discriminatory reasons and the relevance of distinguishing between descriptive and prescriptive stereotypes.

Finally, it is worth noting that the definitions of stereotypes have changed over time assuming now simpler and less restrictive forms compared to the past. As suggested by Nelson, the latter included also ‘inaccuracy, negativity, and overgeneralization. It is unfortunate that we have let those original requirements go – after all, they really are the heart of why we care about the topic at all.’ (Nelson, 2009, p. 2). Nowadays, we tend to use more ‘neutral’ definitions, which depict stereotypes as beliefs on groups’ characteristics, attributes and behaviours. The existence of multiple definitions and their neutrality are a double-edged sword. On one hand, they allow us to catch the multifaceted nature of beliefs on gender differences. On the other hand, they led to a proliferation of instruments.

A more accurate choice of instruments in empirical research and refinement of the type of constructs measured by the proposed indicators is required to advance our understanding of these important puzzles. Researchers studying gender stereotypes should prefer using already existing and shared instruments when the aim of their work allows it. Being aware of this, further specification on the subtypes and constructs of gender stereotypes should be given when presenting studies’ results. Furthermore, factor analysis should be performed when the chosen items refer to different constructs. Stereotypical beliefs on ability may have different implications than beliefs on the representativeness of women and men in careers. on choices, behaviours and attitudes and so biasing our measurements. Eventually, this would contribute to the comprehension of the issue of women in STEM, by facilitating the comparison of similar studies.

References

AlmaLaurea. (2021). *XXIII Indagine Profilo dei Laureati*. AlmaLaurea. https://www.almalaurea.it/universita/indagini/laureati/profilo

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype Susceptibility in Children: Effects of Identity Activation on Quantitative Performance. *Psychological Science*, *12*(5), 385–390. https://doi.org/10.1111/1467-9280.00371

Bar-Anan, Y., Nosek, B., & Vianello, M. (2009). The Sorting Paired Features Task: A Measure of Association Strengths. *Experimental Psychology*, *56*(5), 329–343. https://doi.org/10.1027/1618-3169.56.5.329

Barnes-Holmes, D., Barnes-Holmes, Y., Hayden, E., Milne, R., & Stewart, I. B. (2006). Do You Really Know What You Believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a Direct Measure of Implicit Beliefs. *The Irish Psychologist*, *32*(7), 169–177.

Beere, C. A. (1990). *Gender Roles: A Handbook of Tests and Measures*. Greenwood Press. http://journals.sagepub.com/doi/10.1177/027046769301300329

Betz, D. E., & Sekaquaptewa, D. (2012). My fair physicist? Feminine math and science role models demotivate young girls. *Social Psychological and Personality Science*, *3*(6), 738–746. https://doi.org/10.1177/1948550612440735

Breda, T., Grenet, J., Monnet, M., & Effenterre, C. van. (2018). *Can female role models reduce the gender gap in science? Evidence from classroom interventions in French high schools* (halshs-01713068; PSE Working Papers). HAL. https://ideas.repec.org/p/hal/psewpa/halshs-01713068.html

Buck, G. A., Leslie-Pelecky, D., & Kirby, S. K. (2002). Bringing Female Scientists into the Elementary Classroom: Confronting the Strength of Elementary Students' Stereotypical Images of Scientists. *Journal of Elementary Science Education*, *14*(2), 1–9.

Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy, and Law*, *5*(3), 665–692. https://doi.org/10.1037/1076-8971.5.3.665

Cundiff, J. L., Vescio, T. K., Loken, E., & Lo, L. (2013). Do gender–science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Social Psychology of Education*, *16*(4), 541–554. https://doi.org/10.1007/s11218-013-9232-8

Cusack, S. (2013). *Gender stereotyping as a human rights violation*. Office of the High Commissioner for Human Rights. https://www.esem.org.mk/pdf/Najznachajni%20vesti/2014/3/Cusack.pdf

Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology*, *40*(5), 642–658. https://doi.org/10.1016/j.jesp.2004.02.003

del Rio, M. F., Strasser, K., Cvencek, D., Susperreguy, M. I., & Meltzoff, A. N. (2019). Chilean kindergarten children's beliefs about mathematics: Family matters. *Developmental Psychology*, *55*(4), 687–702. https://doi.org/10.1037/dev0000658

del Rio, M. F., Susperreguy, M. I., Strasser, K., Cvencek, D., Iturra, C., Gallardo, I., & Meltzoff, A. N. (2020). Early sources of children's math achievement in chile: The role of parental beliefs and feelings about math. *Early Education and Development*. https://doi.org/10.1080/10409289.2020.1799617

Eagly, A. H., & Wood, W. (2012). Social Role Theory. In *Handbook of Theories of Social Psychology: Volume 2* (Vol. 1–2, pp. 458–476). SAGE Publications Ltd. https://doi.org/10.4135/9781446249222

Ellemers, N. (2018). Gender Stereotypes. *Annual Review of Psychology*, *69*(1), 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

Ertl, B., Luttenberger, S., & Paechter, M. (2017). The Impact of Gender Stereotypes on the Self-Concept of Female Students in STEM Subjects with an Under-Representation of Females. *Frontiers in Psychology*, *8*, 703. https://doi.org/10.3389/fpsyg.2017.00703

Farrell, L., Nearchou, F., & McHugh, L. (2020). Examining the effectiveness of brief interventions to strengthen a positive implicit relation between women and STEM across two timepoints. *Social Psychology of Education*, *23*(5), 1203–1231. https://doi.org/10.1007/s11218-020-09576-w

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman Mathematics Attitudes Scales: Instruments Designed to Measure Attitudes Toward the Learning of Mathematics by females and Males. *Journal for Research in Mathematics Education*, *7*(5), 324–326.

Ferguson, S., & Lezotte, S. (2020). Exploring the state of science stereotypes: Systematic review and meta-analysis of the Draw-A-Scientist Checklist. *School Science And Mathematics*, *120*(1), 55–65. https://doi.org/10.1111/ssm.12382

Gawronski, B., & De Houwer, J. (2013). Implicit Measures in Social and Personality Psychology. In H. T. Reis & C. M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology* (2nd ed., pp. 283–310). Cambridge University Press. https://doi.org/10.1017/CBO9780511996481.016

Gilbert, P. N. (2015). *The role of role models: How does identification with STEM role models impact women's implicit STEM stereotypes and STEM outcomes?* [Ph.D. Dissertation, Tulane University]. https://digitallibrary.tulane.edu/islandora/object/tulane%3A27945

Gilbert, P. N., O'Brien, L. T., Garcia, D. M., & Marx, D. M. (2015). Not the Sum of Its Parts: Decomposing Implicit Academic Stereotypes to Understand Sense of Fit in Math and English. *Sex Roles*, *72*(1–2), 25–39. https://doi.org/10.1007/s11199-014-0428-y

Gill, M. J. (2004). When information does not deter stereotyping: Prescriptive stereotyping can foster bias under conditions that deter descriptive stereotyping. *Journal of Experimental Social Psychology*, *40*(5), 619–632. https://doi.org/10.1016/j.jesp.2003.12.001

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491–512. https://doi.org/10.1037/0022-3514.70.3.491

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*(6), 1022–1038. https://doi.org/10.1037/0022-3514.79.6.1022

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Grogan, K. E. (2019). How the entire scientific community can confront gender bias in the workplace. *Nature Ecology & Evolution*, *3*(1), 3–6. https://doi.org/10.1038/s41559-018-0747-4

Grow, A., Takács, K., & Pál, J. (2016). Status Characteristics and Ability Attributions in Hungarian School Classes: An Exponential Random Graph Approach. *Social Psychology Quarterly*, *79*(2), 156–167. https://doi.org/10.1177/0190272516643052

Guizzo, F., Moe, A., Cadinu, M., & Bertolli, C. (2019). The role of implicit gender spatial stereotyping in mental rotation performance. *ACTA PSYCHOLOGICA*, *194*, 63–68. https://doi.org/10.1016/j.actpsy.2019.01.013

Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The Multiple Dimensions of Gender Stereotypes: A Current Look at Men’s and Women’s Characterizations of Others and Themselves. *Frontiers in Psychology*, *10*, 11. https://doi.org/10.3389/fpsyg.2019.00011

Hilton, J. L., & von Hippel, W. (1996). STEREOTYPES. *Annual Review of Psychology*, *47*(1), 237–271. https://doi.org/10.1146/annurev.psych.47.1.237

Jackson, S. M., Hillard, A. L., & Schneider, T. R. (2014). Using implicit bias training to improve attitudes toward women in STEM. *Social Psychology of Education*, *17*(3), 419–438. https://doi.org/10.1007/s11218-014-9259-5

Jasko, K., Dukala, K., & Szastok, M. (2019). Focusing on gender similarities increases female students’ motivation to participate in STEM. *Journal of Applied Social Psychology*, *49*(8), 473–487. https://doi.org/10.1111/jasp.12598

Johns, M., Schmader, T., & Martens, A. (2005). Knowing Is Half the Battle: Teaching Stereotype Threat as a Means of Improving Women’s Math Performance. *Psychological Science*, *16*(3), 175–179. https://doi.org/10.1111/j.0956-7976.2005.00799.x

Jost, J. T., & Kay, A. C. (2005). Exposure to Benevolent Sexism and Complementary Gender Stereotypes: Consequences for Specific and Diffuse Forms of System Justification.

Journal of Personality and Social Psychology, *88*(3), 498–509. https://doi.org/10.1037/0022-3514.88.3.498

Kersey, A. J., Csumitta, K. D., & Cantlon, J. F. (2019). Gender similarities in the brain during mathematics development. *Npj Science of Learning*, *4*(1), 19. https://doi.org/10.1038/s41539-019-0057-x

Kiefer, A. K., & Sekaquaptewa, D. (2007a). Implicit Stereotypes, Gender Identification, and Math-Related Outcomes: A Prospective Study of Female College Students. *Psychological Science*, *18*(1), 13–18. https://doi.org/10.1111/j.1467-9280.2007.01841.x

Kiefer, A. K., & Sekaquaptewa, D. (2007b). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*(5), 825–832. https://doi.org/10.1016/j.jesp.2006.08.004

Kite, M. E., Deaux, K., & Haines, E. L. (2008). Gender stereotypes. In *Psychology of women: A handbook of issues and theories, 2nd ed.* (pp. 205–236). Praeger Publishers/Greenwood Publishing Group.

Kyriazos, T. A., & Stalikas, A. (2018). Applied Psychometrics: The Steps of Scale Development and Standardization Process. *Psychology*, *09*(11), 2531–2560. https://doi.org/10.4236/psych.2018.911145

Lane, K. A., Goh, J. X., & Driver-Linn, E. (2012). Implicit Science Stereotypes Mediate the Relationship between Gender and Academic Participation. *Sex Roles*, *66*(3–4), 220–234. https://doi.org/10.1007/s11199-011-0036-z

Leder, G., & Forgasz, H. (2002). *Two new instruments to probe attitudes about gender and mathematics* (p. 29). La Trobe University. https://www.academia.edu/20451640/Two_New_Instruments_To_Probe_Attitudes_about_Gender_and_Mathematics

Lippmann, W. (1922). *Public Opinion*. Harcourt, Brace & Co.

Liu, M., Hu, W., Jiannong, S., & Adey, P. (2010). Gender Stereotyping and Affective Attitudes Towards Science in Chinese Secondary School Students. *International Journal of Science Education*, *32*(3), 379–395. https://doi.org/10.1080/09500690802595847

Marx, D. M., & Ko, S. J. (2012). Prejudice, Discrimination, and Stereotypes (Racial Bias). In *Encyclopedia of Human Behavior* (pp. 160–166). Elsevier. https://doi.org/10.1016/B978-0-12-375000-6.00388-8

Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. N. (2017). Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology*, *160*, 92–106. https://doi.org/10.1016/j.jecp.2017.03.013

McGuire, L., Mulvey, K. L., Goff, E., Irvin, M. J., Winterbottom, M., Fields, G. E., Hartstone-Rose, A., & Rutland, A. (2020). STEM gender stereotypes from early childhood through adolescence at informal science centers. *Journal of Applied Developmental Psychology*, *67*, 101109. https://doi.org/10.1016/j.appdev.2020.101109

McHugh, M. C., & Frieze, I. H. (1997). The Measurement of Gender-Role Attitudes: A Review and Commentary. *Psychology of Women Quarterly*, *21*(1), 1–16. https://doi.org/10.1111/j.1471-6402.1997.tb00097.x

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*(1), 83–90. https://doi.org/10.1016/S0022-1031(02)00513-9

Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The Development of Children's Gender-Science Stereotypes: A Meta-analysis of 5 Decades of U.S. Draw-A-Scientist Studies. *Child Development*, *89*(6), 1943–1955. https://doi.org/10.1111/cdev.13039

Nelson, T. D. (Ed.). (2009). *Handbook of prejudice, stereotyping, and discrimination*. Psychology Press.

Nosek, B., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, *19*(6), 625–666. https://doi.org/10.1521/soco.19.6.625.20886

Nosek, B., Banaji, M. R., & Greenwald, A. G. (1998). *Project Implicit*. https://implicit.harvard.edu/implicit/index.jsp

Nosek, B., & Smyth, F. L. (2011). Implicit Social Cognitions Predict Sex Differences in Math Engagement and Achievement. *American Educational Research Journal*, *48*(5), 1125–1156. https://doi.org/10.3102/0002831211410683

Nurlu, Ö. (2017). Developing a Teachers Gender Stereotype Scale toward Mathematics. *International Electronic Journal of Elementary Education*, *10*(2), 287–299. https://doi.org/10.26822/iejee.2017236124

Nurnberger, M., Nerb, J., Schmitz, F., Keller, J., & Sutterlin, S. (2016). Implicit gender stereotypes and essentialist beliefs predict preservice teachers’ tracking recommendations. *Journal of Experimental Education*, *84*(1), 152–174. https://doi.org/10.1080/00220973.2015.1027807

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*(3), 277–293. https://doi.org/10.1037/0022-3514.89.3.277

Plante, I., Théorêt, M., & Favreau, O. E. (2009). Student gender stereotypes: Contrasting the perceived maleness and femaleness of mathematics and language. *Educational Psychology*, *29*(4), 385–405. https://doi.org/10.1080/01443410902971500

Rentas, C. A. (2015). The effects of a perceived causal relationship on the strength of stereotypes. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *76*(6-A(E)).

Reuben, E., Sapienza, P., & Zingales, L. (2014). How stereotypes impair women’s careers in science. *Proceedings of the National Academy of Sciences*, *111*(12), 4403–4408. https://doi.org/10.1073/pnas.1314788111

Riegle-Crumb, C., & Morton, K. (2017). Gendered Expectations: Examining How Peers Shape Female Students’ Intent to Pursue STEM Fields. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00329

Schneider, D. J. (2005). *The psychology of stereotyping* (Paperback ed). Guilford Press.

Sczesny, S., Nater, C., & Eagly, A. H. (2018). Agency and communion: Their implications for gender stereotypes and gender identities. *Agency and Communion in Social Psychology*, 103–116. https://doi.org/10.4324/9780203703663

Six, B., & Eckes, T. (1991). A closer look at the complex structure of gender stereotypes. *Sex Roles*, *24*(1), 57–71. https://doi.org/10.1007/BF00288703

Smiler, A. P. (2004). Thirty Years After the Discovery of Gender: Psychological Concepts and Measures of Masculinity. *Sex Roles*, *50*(1/2), 15–26. https://doi.org/10.1023/B:SERS.0000011069.02279.4c

Spence, J. T., & Helmreich, R. (1972). The attitudes toward women scale: An objective instrument to measure attitudes toward the rights and roles of women in contemporary society. *Catalog of Selected Documents in Psychology*, *2*(66).

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology*, *102*(4), 947–963. https://doi.org/10.1037/a0019920

Tomasetto, C., Galdi, S., & Cadinu, M. (2012). Quando l'implicito precede l'esplicito: Gli stereotipi di genere sulla matematica in bambine e bambini di 6 anni. *Psicologia sociale*, *2*, 169–186. https://doi.org/10.1482/37693

Van Camp, A. R., Gilbert, P. N., & O'Brien, L. T. (2019). Testing the effects of a role model intervention on women's STEM outcomes. *Social Psychology of Education*, *22*(3), 649–671. https://doi.org/10.1007/s11218-019-09498-2

White, M. J., & White, G. B. (2006). Implicit and Explicit Occupational Gender Stereotypes. *Sex Roles*, *55*(3–4), 259–266. https://doi.org/10.1007/s11199-006-9078-z

Whitley, B. E., & Kite, M. E. (2016). *Psychology of prejudice and discrimination* (Third Edition). Routledge/Taylor & Francis Group.

Young, D. M., Rudman, L. A., Buettner, H. M., & McLean, M. C. (2013). The Influence of Female Role Models on Women's Implicit Science Cognitions. *Psychology of Women Quarterly*, *37*(3), 283–292. https://doi.org/10.1177/0361684313482109

Zitelny, H., Shalom, M., & Bar-Anan, Y. (2017). What Is the Implicit Gender-Science Stereotype? Exploring Correlations Between the Gender-Science IAT and Self-Report Measures. *Social Psychological and Personality Science*, *8*(7), 719–735. https://doi.org/10.1177/1948550616683017

Chapter 3. ***Gender bias in the classroom: A network study on self and peers' ability attribution among high-school students in Italy***

1. Introduction

In Italy, women make up only one-third of the workforce in STEM sectors. While the underrepresentation of women in STEM is a worldwide issue, the percentage of female scientists and engineers in Italy is below the European average, 41% (Eurostat, 2019). Although organizations have introduced diversity and inclusion policies targeting women in various sectors, including education, 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 & 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 shape (re-shape) “social relational contexts”, which in turn would reproduce these gender patterns (Ridgeway & 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 maths and reading abilities was prominent among children and their parents (Galdi et al., 2017; Passolunghi et al., 2014; Tomasetto et al., 2012, 2015), as well as among teachers (Carlana, 2019).

The endorsement of these gender stereotypes often lead to faulty assessments, which in turn influence expectations on the performance of individuals, thus biasing opportunities and work outcomes for both men and women (Hentschel et al., 2019). As theorized by the 'status characteristic theory' (Berger et al., 1972), when gender is paramount to the situation, gender-based stereotypes 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.

Previous research has examined the effect of gender stereotypes in STEM on a heterogeneous set of individual characteristics, e.g., performance (Kiefer & Sekaquaptewa, 2007; Ramsey & Sekaquaptewa, 2011), self-concept (Cvencek et al., 2015; Ertl et al., 2017; Steffens et al., 2010), and attitudes toward math (Nosek et al., 2002; Nosek & Smyth, 2011). However, less is known about *expected* performance, especially in 'outside the lab', real-world settings (Grow et al., 2016; Kisfalusi et al., 2019; Mann & DiPrete, 2016), and how this could trigger gender-biased behaviour. Furthermore, previous studies have mostly focused on individuals' general endorsement of gender stereotypes, while neglecting the dynamics of social influence, i.e., the opinion of other individuals embedded in the same specific social context. Indeed, previous studies on stereotype dynamics in social groups found that peers influence choices and intentions, especially for female students in STEM (Dasgupta et al., 2015; Riegle-Crumb & Morton, 2017; Robnett & Leaper, 2013; van der Vleuten et al., 2018).

If we consider that information, expectations, and social group pressures as part of a context-specific social construction process are key to reconstructing the link between stereotypes and behaviour, it is still hard to find suitable data to capture such complexity. On the one hand, it would be necessary to identify a context where a small social group engages in potential gender-biased behaviour. On the other hand, data should collect information on both personal and relational characteristics of all a social group's members.

To fill this gap, this study reports the outcomes of a network study ran 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', it is tested whether gender could affect the formation of expectations on classmates' abilities, even in contexts where information on all students' actual performance was publicly 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' abilities.

The objective here was twofold. On the one hand, the study aims to understand whether gender influences students' perception of their classmates' ability in disciplines typically considered either feminine or masculine. By modelling peers' perceived abilities as social networks (Grow et al., 2016; Kisfalusi et al., 2019; Paluck, 2011; Shepherd & Paluck, 2015), we can disentangle the effect of peers' gender on perceived abilities from observed performance and other relationships between classmates (Grow et al., 2016). On the other hand, the study aims to test the effect of gender in the self-evaluation of these skills.

The study's contribution to research on gender bias in STEM and humanities is threefold. First, it provides 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-the-lab' setting. Being the network study performed in a school context, the population target is similar to a lab-experiment design, i.e., young students, with the advantage of taking observations from a 'natural' social environment. Second, it proposes a new and indirect method to test gender bias on reading/mathematical skills, which allows to avoid social desirability bias (details in the methodology section). Finally, it expands this recent line of research on student bias to the case of stereotypes about boys and reading, often neglected in previous research.

2. Literature review

2.1. Status characteristic theory

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 & 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', being 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. 1972:242; Wagner and Berger 1997). Status characteristics can be either ‘diffuse’ (e.g., gender) or ‘specific’ (e.g., math ability) with important implications in terms of assumptions. Expectations based on diffuse status characteristics can be both general (e.g., men would be generally more intelligent than women) and specific, while those based on specific status characteristics can only be specific (e.g., people good at math would perform better than others in math tests). For a certain task, individuals in the low state “receive less attention, they are given lower evaluations and they exert less influence” (Foddy & Smithson, 1999, p. 308), while “those with a status advantage will adopt a repertoire of attitudes and behaviours that is associated with their status superiority” (Wagner et al., 1986, p. 48). In contexts in which gender is a salient characteristic, individuals may form expectations on others’ and personal abilities based on gender.

2.2. Multiple status characteristics

Berger at al. (1972) argued that in a situation in which multiple salient status characteristics can be activated, subjects tend to combine status information rather than selecting it. Wagner and Berger (1997) also argued that whenever information on multiple status characteristics is available, e.g., gender and ability, the combined outcome on expectation formation depends upon information consistency. This implies that ability would magnify gender-based expectations if the information were consistent. Otherwise, these expectations would change slightly or even drastically, according to new information.

Simpson and Walker (2002, p. 28) proposed a reformulation of the fourth assumption of SCT, which states that “similarly signed states of diffuse and specific characteristics have exactly the same impact on performance expectations”. They rather claimed that the impact of diffuse characteristics is greater than that of specific characteristics. Consequently, whenever a certain specific characteristic is relevant to a task (e.g., ability in maths), the impact of a diffuse characteristic (e.g., gender) could be reduced - although not eliminated.

This reformulated assumption was confirmed by Pugh and Wahrman (1983). They 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 behaviour compared to the control group (no information).

2.3. Ability attribution in a school setting

Schools are an ideal setting to examine gender bias and especially, ability attribution in 'outside the lab', real-world social settings. Unfortunately, only a few studies have studied 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 gender and ethnicity as 'clever students'.

However, these studies focused on estimating gender effects on general academic ability rather than on specific skills. While children tend to consider their gender as the cleverest, (young) adults are less sensitive toward in-group favouritism, i.e., the tendency to favour and magnify in-group members' achievements (e.g., same gender) to indirectly enhance self-esteem (Tajfel & Turner, 2004). However, the latter are more sensitive to societal stereotypes (Passolunghi et al., 2014). 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 maths with women having higher reading abilities (Steffens & 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.

2.4. Competitive context and expectations

In a competitive task setting, other individual characteristics - which others are unaware of - could influence expectations on performance and consequently self-nominations. Previous studies found a general tendency for women to be less competitive than men (Buser et al., 2012), suffer more from anxiety triggered by performance and competition (Baraskar & Shinde, 2018), and be more risk-averse (Fisk, 2018). This tendency is further exacerbated when women compete in a masculine field.

More specifically concerning science, research has indicated that women would be more reluctant to be involved in mathematics and science competitions (for a review and discussion of these studies, 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. This would suggest that whenever the task is a competition, women would more probably combine information on gender and ability while considering their attitude towards competition.

3. Hypotheses

Assuming that higher status should be attributed to female students for the reading competition and male students for the maths/science competition, the following hypotheses are formulated:

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 this setting, we can expect 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.

4. Methodology

4.1. Data

Data were collected from a sample of high school classes in Milan (Italy) between January-February 2020, using a computer-assisted survey. I 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.

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, I was able to collect data in only five out of ten schools who agreed to participate, with a total amount 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 performance. 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 total of 5 years, on average. It is, therefore, plausible to assume 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,* differing in the specialized courses' orientation*,* all have certain core subjects in common, including reading, maths, and science. Due to the different specializations, usually, we do not observe gender balance among students and schools in this sample make no exception. Most students in humanities-oriented classes were females (83%), while the opposite occurred in science-oriented schools (32%). More details on the percentage of male and female students in each class is reported in the Appendix.

4.2. Instruments

Following Grow et al. (2016) and Kisfalusi et al. (2019), the data collected were both relational, i.e., subjects' perceived academic ability of their classmates and friendship ties, and individual, i.e., relating to the subjects' socio-demographic characteristics and academic performance. Data were collected through a computer-assisted questionnaire administered to each class during regular school time.

4.3. 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 maths 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 these hypothetical school competitions in reading. By eliciting nominations for this fictional competition and integrating this information with data on the nominees' gender, we obtained an indirect measurement of gender bias.

Compared to traditional approaches, this instrument allows 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 that reflects what they believe is socially acceptable rather than what they 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, 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' evaluation of their peers' abilities, it was necessary to design an instrument that stimulated students to report their opinion rather than that of their teachers. If, for instance, we asked students to nominate the best classmates in a subject, they would probably nominate those with the highest grades, without thinking much about it. We would, thus, end with an instrument that measures teachers' evaluation rather than classmates' evaluation of students.

Nevertheless, school grades here play an important role. On the one hand, they are used as a proxy for performance information available to students. On the other hand, in evaluating gender bias in nominations, we assume that school grades measure the student's real ability in the subject. This instrument has, thus, two potential weaknesses if (1) students do not know their classmates' grades and thus do not have information about abilities, and (2) school grades do not necessarily measure actual ability in a subject. As regards the first issue, Italian students have constant access to information on all class members' school performance. Indeed, students are continuously evaluated through multiple tests during school terms, which are frequently discussed in class. Furthermore, final term grades are publicly displayed at the end of the year and are often a topic of conversation among students. The assumption that students in the study were aware of their classmates' performance in school is, thus, realistic.

As regards teachers' evaluations, indeed, grades can 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, Jonsson, and van Tubergen 2017), e.g., skipping class and effort in school, as well as by teachers' bias, a problem known as the 'gender grading gap' (Protivínský & Münich, 2018). Evidence shows that teachers generally tend to favour female students and give them higher grades (Voyer & Voyer, 2014), except for maths (Hofer, 2015; Spear, 1984a, 1984b). Being aware of these limitations, in a school context, especially in Italy where standard tests are rare, teachers' evaluations are the cornerstone of the school system. Grades can be partially biased, but on average they are a good measure of students' actual performance in school.

I decided to add science as a third context because, while related to the same sector, the gender gap in maths and science do differ. In Italy in 2019, the percentage of female undergraduates was 75% in biology, 32% in physics, and 49% in mathematics (AlmaLaurea, 2019). Furthermore, while most studies on the theme focus on the math domain, other scientific sectors are more rarely investigated and, thus, need more attention.

4.3.1. Relational data

Relational data were collected using sociometric questions formatted according to the conventional repeated roster method for name-generators (Kilduff & Krackhardt, 2008).

Nominations for the competition: Each student was asked to nominate the four best candidates from their classmates for three hypothetical inter-class competitions in reading, maths, and science. Three tie-variables were then built, 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 ($i = j$).

Friendship: Students were asked to nominate those classmates whom they considered to be 'friends', so that $x_{ij} = 1$ if i considered j to be a 'friend', while $x_{ij} = 0$ otherwise, with $i \neq j$.

4.3.2. Individual data

Gender: Students were asked to report their gender by selecting one of the following categories: ‘male’, ‘female’ and ‘other’. Given that only one student selected the latter option, that record was excluded from the analysis.

Grades: Students were asked to report their grades obtained in Italian, maths, and science, as published as the final evaluation of the previous school year. In Italy, final grades range between one and ten, however, students with a final grade below six are not allowed to the following academic year unless they pass a make-up exam with a grade at least equal to six before the beginning of the new academic year. This implies that, formally, all students in the last year had a grade at least equal to six, either obtained as the final evaluation of the previous year or as the evaluation of the make-up exam. In practice, few students reported a grade equal to 5. The variable, thus, ranged from 5 to 10.

Identification with the subject: Following Gilbert (2015), students were asked to rate on a 5-point Likert scale the extent to which they identified with reading, math, and science (more details in the Appendix) on a 4-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 maths, 0.89 in science).

4.4. Analytical methods

The two hypotheses are tested using two different analytical approaches. 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 allow to model the statistical dependencies of student nominations. Hypothesis 2 was tested using a logistic regression model with fixed effects and cluster-robust standard errors.

4.4.1. Hypothesis 1

Gender difference in others’ nominations were estimated using a set of ERGMs for each subject (reading, maths, and science). Three classes out of ten were excluded from the sample because of the high percentage of missing data on ties. The use of ERGMs allows 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., school performance) or relational interdependent processes (i.e., friendship).

Following Snijders and Baerveldt (2003), I performed a meta-analysis of the estimated ERGM parameters for each subject. This allows to estimate macro-level parameters across classes as weighted least square (WLS) means which, unlike a simple mean, model heterogeneity in standard error. I then used *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 Appendix. Missing data on students' characteristics were estimated via multivariate imputation by chained equations (see the Appendix), 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 & Groothuis-Oudshoorn, 2011).

All ERG models were specified for the same parameter set. To test Hypothesis 1, the model included a *gender receiver* parameter, whose coefficient indicated the likelihood that a female rather than a male student was nominated. The *performance receiver* parameter controlled for students' performance. A positive coefficient indicates the tendency of students with higher grades to receive more nominations.

I 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, the *friendship* parameter, specified as a dyadic covariate, controls for the likelihood that students simply nominated friends. Note that gender homophily was not included in the final model because of multicollinearity issues with friendship. Finally, given that students were asked to nominate four students, the maximal outdegree was set to four in the model specification. More details are provided in the Appendix.

4.4.2. Hypothesis 2

Being self-nominations of students modelled as a binary variable, Hypothesis 2 was tested using a fixed-effects logistic regression model with clustered standard errors, to account for the clustered nature of data while minimizing bias due to the low number of clusters (McNeish & Kelley, 2019; McNeish & Stapleton, 2016). I 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 *grades* and *identification* with the specified subjects as control variables.

5. Results

Table 3.1 shows the results of the meta-analysis of parameter estimates of ERG models. To test Hypothesis 1, we look at the mean coefficient estimates of the *receiver gender* parameters. As regards the competition in reading, the estimated positive and statistically significant coefficient shows that female students were more likely to be nominated than male students, even accounting for the intervening effect of performance and the confounding effect of other model parameters, thus confirming hypothesis H1a. Conversely, regarding maths and science, 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 is only partially confirmed, as data could not confirm the difference between female and male students was different from zero in the case of maths.

Unsurprisingly, results showed that performance 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 are reported in the Appendix. They showed a typical situation found in other studies on status-related networks, i.e., students tended not to reciprocate nominations, whereas they did tend to nominate students who had already been nominated by their own nominations (e.g., Krackhardt 1994).

Table 3.1 *Meta-analysis results of ERG models*

	Estimate	SE	t-ratio	p-value
Reading competition				
Receiver: Gender [female]	1.49	.10	14.56	.000
Receiver: School grade	1.26	.19	6.72	.000
Friendship	.99	.09	11.99	.000
Math competition				
Receiver: Gender [female]	-.43	.53	-.81	.209
Receiver: School grade	1.31	.21	6.34	.000
Friendship	1.41	.13	10.62	.000
Science competition				
Receiver: Gender [female]	-.33	.17	-1.95	.026
Receiver: School grade	.43	.06	6.76	.000
Friendship	.61	.15	4.18	.000

Note. SE = Standard errors

Concerning Hypothesis 2, Table 3.2 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 performance and identification with the subject. 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 stronger identification with 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 performance and identification with the subject.

I also tested whether this gender gap varied according to the grade and identification levels, by adding two interaction terms in the model, i.e., gender and grade, and gender and identification. However, estimated coefficients of these interactions did not provide clear evidence of any effect in either case.

Table 3.2 *Regression results on self-nominations*

	Estimate	SE	Percentile bootstrap 95% CI		p-value
			LB	UB	
Italian					
Gender [Female]	-1.33	.21	-1.77	-1.25	.006
School grade in Italian	1.07	.12	.84	1.32	.002
Identification with Italian	1.02	.20	.63	1.43	.007
Math					
Gender [Female]	-1.49	.85	-2.51	-.79	.005
School grade in math	.92	.19	.96	1.34	.000
Identification with math	.36	.11	.46	.65	.038
Science					
Gender [Female]	-2.17	.47	-3.62	-1.67	.000
School grade in science	.64	.43	-.03	1.81	.116
Identification with science	.77	.24	.44	1.27	.005

Note. N = 195

SE = bootstrap standard errors; CI = confidence interval; LB = Lower bound; UB = Upper bound

[a] Reference category 'Above high school'

[b] Reference category 'Humanities-oriented'

6. Discussion and conclusions

As stated, I 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 about students' own and others' abilities in a stereotypical context (i.e., a competition requiring reading and mathematical/scientific skills), when performance information 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.

Note that two characteristics of the school system of the sample were pivotal when testing the hypotheses. On the one hand, students attended the same class with the same peers for four years, thereby establishing long-term relationships with countless occasions to observe each other's abilities and attitudes. On the other hand, grades were publicly available to students, who thereby had complete information on their classmates' performance.

Consistent with the ‘status characteristic theory’, I hypothesized that students would have formed expectations based on gender stereotypes even in cases where the performance of classmates contradicted the stereotypical attribution of different abilities in STEM and reading. Using Exponential Random Graph Models, I tested whether female students were nominated less frequently for competitions in maths and science, and more frequently for the reading competition. I also performed a fixed effects logistic regression analysis to check whether female students were less likely to nominate themselves for these competitions.

The findings confirmed some of the initial hypotheses. Regarding expectations about peers, even controlling for grades in school and friendship relationships, which were both influential, gender still played a role in the case of reading and science competitions. Female students were more likely to be nominated for the reading competition and less likely for the science competition. Even if, as expected, female students were less likely to be nominated also for the maths competition, in that case, data could not reject the hypothesis of non-significance.

This would suggest that students in the sample may have still endorsed gender stereotypes on reading skills, and this affected their expectations on others’ abilities, even when they were aware of their performance. The difference in findings between math and science, if not related to the sample, could confirm previous research showing that the math 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, 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 (Plante et al., 2009). These include, for instance, awareness of role models (Olsson & Martiny, 2018), public debate and discussion of the gaps within schools (UN International day of women and girl in science), and laboratories and initiatives targeted at women to increase their knowledge and interest in science (e.g., “Shecodes”, “sciencegirlslab” in the U.S.).

Even given that the small size of the sample prevents us from making any generalizations, the findings on the difference between maths 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, Luttenberger, and Paechter (2017) and Blažev et al. (2017) included STEM and not only math in the instruments used to measure stereotypes. Improving our understanding of the association between gender bias and science is also key to reducing the mismatch between education and jobs in future professions.

While the network analysis suggests a tendency to favour female students in reading and male students in science, I did not find a similar difference in the formation of self-expectations. Indeed, as expected, 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 identification. The gap was greater for nominations in science and math but was also present for those in reading.

Results corroborate previous findings on the general tendency of women to avoid competition (Gneezy et al., 2003) and suffer more from anxiety than men (Baraskar & Shinde, 2018). As emphasized by Niederle and Vesterlund (2010), there is 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 & 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 been equal, 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, this study provides an example of the application of the SCT's assumptions on the information-combining process for ability's expectations in an 'out-of-the-lab' setting. Assuming that, in a school setting, teachers' evaluations are a good proxy of classmates' information on others' abilities, even when performance information is available, gender may still influence individual expectations about personal and others' abilities. Findings would, thus, suggest that young adults may be affected by gender stereotypes in judging skills and abilities even in a familiar environment, such as school, where everyone has full information about other classmates.

Having noted this, this research does have certain limitations. First, the sample size was not sufficiently large to support any generalization. Furthermore, data were cross-sectional and did not allow us to examine the evolution of relationships among students in the class, as well as how these relationships shaped and influenced performance and beliefs. A longitudinal approach could help us to explain whether students are affected by stereotype threats in the class due to prejudices endorsed by their classmates or whether these prejudices are changed or consolidated because of the performance of students in academic exams.

Another limitation is related to the unequal representation of female and male students within classes. As previously mentioned, students in the sample came from different types of high schools, some male-dominated, others more female-dominated. While these differences did not affect the measurement of gender bias, given that all students studied reading, math, and science, they could have affected student nominations. However, I did not find any pattern among ERGMs estimated on networks from the same type of school. Note also that information on grades was self-reported thus potentially subject to inaccuracy, especially for low-achieving students (Cassady, 2001; Mayer et al., 2007) and gender differences (Caskie et al., 2014). Furthermore, in addition to the issues of teachers' subjectivity in students' evaluation, discussed before, grades of low-achieving students may also partially reflect teachers' adjustments in the last period of the year. Since students with a grade lower than six cannot pass on to the next year, either teachers adjust grades before the decision on final evaluations or students are required to pass a re-sit exam.

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. The 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 of information, abilities and prospective performance in competitive settings could help to disentangle possible causal links between these factors (Baldassarri & Abascal, 2017).

In conclusion, findings suggest that there is still a need to intervene in the STEM area at least in science, but also in the female-dominated sector of humanities. We should contemplate initiatives targeted at reducing all types of stereotypes associating gender and abilities, instead of focusing only on certain areas, e.g., maths. 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. Findings suggest that this association still seems to be strong. However, further studies are necessary to understand how these stereotypical associations affect the choices and behaviours 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 labour markets and traditional work-family division of labour. The gendered expectations on their peers could even exacerbate this demotivation, and negatively affect their STEM career intentions (Riegle-Crumb & Morton, 2017; Robnett & Leaper, 2013; van der Vleuten et al., 2018).

References

AlmaLaurea. (2019). *XXII Indagine Profilo dei Laureati*. Almalaurea. https://www.almalaurea.it/sites/almalaurea.it/files/docs/universita/profilo/profilo2020/almalaurea_profilo_rapporto2020.pdf

Baldassarri, D., & Abascal, M. (2017). Field Experiments Across the Social Sciences. *Annual Review of Sociology*, *43*(1), 41–73. https://doi.org/10.1146/annurev-soc-073014-112445

Baraskar, V. J., & Shinde, P. P. (2018). To study the gender difference in personality factor and anxiety in adolescents – sociability self-confidence ambition and anxiety. *Journal of Evidence Based Medicine and Healthcare*, *5*(18), 1458–1461. https://doi.org/10.18410/jebmh/2018/305

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status Characteristics and Social Interaction. *American Sociological Review*, *37*(3), 241–255. JSTOR. https://doi.org/10.2307/2093465

Blažev, M., Karabegović, M., Burušić, J., & Selimbegović, L. (2017). Predicting gender-STEM stereotyped beliefs among boys and girls from prior school achievement and interest in STEM school subjects. *Social Psychology of Education*, *20*(4), 831–847. https://doi.org/10.1007/s11218-017-9397-7

Buser, T., Niederle, M., & Oosterbeek, H. (2012). *Gender, Competitiveness and Career Choices* (Working Paper No. 18576). National Bureau of Economic Research. https://doi.org/10.3386/w18576

Carlana, M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias*. *The Quarterly Journal of Economics*, *134*(3), 1163–1224. https://doi.org/10.1093/qje/qjz008

Caskie, G. I. L., Sutton, M. C., & Eckhardt, A. G. (2014). Accuracy of Self-Reported College GPA: Gender-Moderated Differences by Achievement Level and Academic Self-Efficacy. *Journal of College Student Development*, *55*(4), 385–390. https://doi.org/10.1353/csd.2014.0038

Cassady, J. C. (2001). Self-Reported GPA and SAT: A Methodological Note. *Practical Assessment, Research & Evaluation*, *7*(12). https://doi.org/10.7275/5HYM-Y754

Cvencek, D., Kapur, M., & Meltzoff, A. N. (2015). Math achievement, stereotypes, and math self-concepts among elementary-school students in Singapore. *Learning and Instruction*, *39*, 1–10. https://doi.org/10.1016/j.learninstruc.2015.04.002

Dasgupta, N., Scircle, M. M., & Hunsinger, M. (2015). Female peers in small work groups enhance women’s motivation, verbal participation, and career aspirations in engineering. *Proceedings of the National Academy of Sciences*, *112*(16), 4988–4993. https://doi.org/10.1073/pnas.1422822112

Deaner, R. O., McClellan, A., Zeigler-Hill, V., & Benenson, J. F. (2020). Sex differences in exclusion and aggression on single-sex sports teams. *Evolutionary Behavioral Sciences*. https://doi.org/10.1037/ebs0000247

Diekman, A. B., Weisgram, E. S., & Belanger, A. L. (2015). New Routes to Recruiting and Retaining Women in STEM: Policy Implications of a Communal Goal Congruity Perspective: Communal Processes in STEM. *Social Issues and Policy Review*, *9*(1), 52–88. https://doi.org/10.1111/sipr.12010

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*(4), 735–754. https://doi.org/10.1037/0022-3514.46.4.735

Ellemers, N. (2018). Gender Stereotypes. *Annual Review of Psychology*, *69*(1), 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

Ertl, B., Luttenberger, S., & Paechter, M. (2017). The Impact of Gender Stereotypes on the Self-Concept of Female Students in STEM Subjects with an Under-Representation of Females. *Frontiers in Psychology*, *8*, 703. https://doi.org/10.3389/fpsyg.2017.00703

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

Fisk, S. R. (2018). Who’s on Top? Gender Differences in Risk-Taking Produce Unequal Outcomes for High-Ability Women and Men. *Social Psychology Quarterly*, *81*(3), 185–206. https://doi.org/10.1177/0190272518796512

Foddy, M., & Smithson, M. (1999). Can gender inequalities be eliminated? *Social Psychology Quarterly*, *62*(4), 307–324. https://doi.org/10.2307/2695831

Galdi, S., Mirisola, A., & Tomasetto, C. (2017). On the relations between parents' and children's implicit and explicit academic gender stereotypes. *Psicologia sociale*, *2*, 215–238. https://doi.org/10.1482/87248

Geven, S., O. Jonsson, J., & van Tubergen, F. (2017). Gender Differences in Resistance to Schooling: The Role of Dynamic Peer-Influence and Selection Processes. *Journal of Youth and Adolescence*, *46*(12), 2421–2445. https://doi.org/10.1007/s10964-017-0696-2

Gilbert, P. N. (2015). *The role of role models: How does identification with STEM role models impact women's implicit STEM stereotypes and STEM outcomes?* [Ph.D. Dissertation, Tulane University]. https://digitallibrary.tulane.edu/islandora/object/tulane%3A27945

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, *118*(3), 1049–1074. https://doi.org/10.1162/00335530360698496

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Grow, A., Takács, K., & Pál, J. (2016). Status Characteristics and Ability Attributions in Hungarian School Classes: An Exponential Random Graph Approach. *Social Psychology Quarterly*, *79*(2), 156–167. https://doi.org/10.1177/0190272516643052

Günther, C., Ekinci, N. A., Schwieren, C., & Strobel, M. (2010). Women can't jump?—An experiment on competitive attitudes and stereotype threat. *Journal of Economic Behavior & Organization*, *75*(3), 395–401. https://doi.org/10.1016/j.jebo.2010.05.003

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), 1–51. https://doi.org/10.1111/j.1529-1006.2007.00032.x

Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The Multiple Dimensions of Gender Stereotypes: A Current Look at Men's and Women's Characterizations of Others and Themselves. *Frontiers in Psychology*, *10*, 11. https://doi.org/10.3389/fpsyg.2019.00011

Hofer, S. I. (2015). Studying Gender Bias in Physics Grading: The role of teaching experience and country. *International Journal of Science Education*, *37*(17), 2879–2905. https://doi.org/10.1080/09500693.2015.1114190

Kahn, S., & Ginther, D. (2017). *Women and STEM* (No. w23525; p. w23525). National Bureau of Economic Research. https://doi.org/10.3386/w23525

Kersey, A. J., Csumitta, K. D., & Cantlon, J. F. (2019). Gender similarities in the brain during mathematics development. *Npj Science of Learning*, *4*(1), 19. https://doi.org/10.1038/s41539-019-0057-x

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*(5), 825–832. https://doi.org/10.1016/j.jesp.2006.08.004

Kilduff, M., & Krackhardt, D. (2008). *Interpersonal networks in organizations: Cognition, personality, dynamics, and culture*. Cambridge University Press.

Kisfalusi, D., Janky, B., & Takács, K. (2019). Double Standards or Social Identity? The Role of Gender and Ethnicity in Ability Perceptions in the Classroom. *The Journal of Early Adolescence*, *39*(5), 745–780. https://doi.org/10.1177/0272431618791278

Krackhardt, D. (1994). Graph Theoretical Dimensions of Informal Organization. *Computational Organization Theory*, *89*.

Krause, R. W., Huisman, M., Steglich, C., & Snijders, T. A. B. (2020). Missing data in cross-sectional networks—An extensive comparison of missing data treatment methods. *Social Networks*, *62*, 99–112.

Lusher, D., Koskinen, J., & Robins, G. (Eds.). (2012). *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press. https://doi.org/10.1017/CBO9780511894701

Mann, A., & DiPrete, T. (2016). The Consequences of the National Math and Science Performance Environment for Gender Differences in STEM Aspiration. *Sociological Science*, *3*, 568–603. https://doi.org/10.15195/v3.a25

Mayer, R. E., Stull, A. T., Campbell, J., Almeroth, K., Bimber, B., Chun, D., & Knight, A. (2007). Overestimation Bias in Self-reported SAT Scores. *Educational Psychology Review*, *19*(4), 443–454. https://doi.org/10.1007/s10648-006-9034-z

McNeish, D., & Kelley, K. (2019). Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychological Methods*, *24*(1), 20–35. https://doi.org/10.1037/met0000182

McNeish, D., & Stapleton, L. M. (2016). Modeling Clustered Data with Very Few Clusters. *Multivariate Behavioral Research*, *51*(4), 495–518. https://doi.org/10.1080/00273171.2016.1167008

Niederle, M., & Vesterlund, L. (2010). Explaining the Gender Gap in Math Test Scores: The Role of Competition. *Journal of Economic Perspectives*, *24*(2), 129–144. https://ideas.repec.org/a/aea/jecper/v24y2010i2p129-44.html

Niederle, M., & Vesterlund, L. (2011). Gender and Competition. *Annual Review of Economics*, *3*(1), 601–630. https://doi.org/10.1146/annurev-economics-111809-125122

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*(1), 44–59. https://doi.org/10.1037/0022-3514.83.1.44

Nosek, B., & Smyth, F. L. (2011). Implicit Social Cognitions Predict Sex Differences in Math Engagement and Achievement. *American Educational Research Journal*, *48*(5), 1125–1156. https://doi.org/10.3102/0002831211410683

OECD. (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. OECD. https://doi.org/10.1787/9789264229945-en

Olsson, M., & Martiny, S. E. (2018). Does Exposure to Counterstereotypical Role Models Influence Girls' and Women's Gender Stereotypes and Career Choices? A Review of Social Psychological Research. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02264

Palmerio, L. (2019). *I risultati italiani dell'indagine internazionale OCSE PISA 2018*. OECD. https://www.invalsi.it/invalsi/ri/pisa2018/docris/2019/Presentazione_Risultati_PISA2018.pdf

Paluck, E. L. (2011). Peer pressure against prejudice: A high school field experiment examining social network change. *Journal of Experimental Social Psychology*, *47*(2), 350–358. https://doi.org/10.1016/j.jesp.2010.11.017

Passolunghi, M. C., Rueda Ferreira, T. I., & Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. *Learning and Individual Differences*, *34*, 70–76. https://doi.org/10.1016/j.lindif.2014.05.005

Plante, I., Théorêt, M., & Favreau, O. E. (2009). Student gender stereotypes: Contrasting the perceived maleness and femaleness of mathematics and language. *Educational Psychology*, *29*(4), 385–405. https://doi.org/10.1080/01443410902971500

Protivínský, T., & Münich, D. (2018). Gender Bias in teachers' grading: What is in the grade. *Studies in Educational Evaluation*, *59*, 141–149. https://doi.org/10.1016/j.stueduc.2018.07.006

Pugh, M. D., & Wahrman, R. (1983). Neutralizing Sexism in Mixed-Sex Groups: Do Women Have to Be Better Than Men? *American Journal of Sociology*, *88*(4), 746–762. JSTOR. http://www.jstor.org/stable/2779483

Ramsey, L. R., & Sekaquaptewa, D. (2011). Changing stereotypes, changing grades: A longitudinal study of stereotyping during a college math course. *Social Psychology of Education*, *14*(3), 377–387. https://doi.org/10.1007/s11218-010-9150-y

Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the Gender System: A Theoretical Perspective on Gender Beliefs and Social Relations. *Gender & Society*, *18*(4), 510–531. https://doi.org/10.1177/0891243204265269

Riegle-Crumb, C., & Morton, K. (2017). Gendered Expectations: Examining How Peers Shape Female Students' Intent to Pursue STEM Fields. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00329

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 173–191. https://doi.org/10.1016/j.socnet.2006.08.002

Robins, G., Pattison, P., & Wang, P. (2009). Closure, connectivity and degree distributions: Exponential random graph (p*) models for directed social networks. *Social Networks*, *31*(2), 105–117. https://doi.org/10.1016/j.socnet.2008.10.006

Robins, G., Pattison, P., & Woolcock, J. (2004). Missing data in networks: Exponential random graph (p∗) models for networks with non-respondents. *Social Networks*, *26*(3), 257–283. https://doi.org/10.1016/j.socnet.2004.05.001

Robnett, R. D., & Leaper, C. (2013). Friendship Groups, Personal Motivation, and Gender in Relation to High School Students' STEM Career Interest. *Journal of Research on Adolescence*, *23*(4), 652–664. https://doi.org/10.1111/jora.12013

Shepherd, H., & Paluck, E. L. (2015). Stopping the Drama: Gendered Influence in a Network Field Experiment. *Social Psychology Quarterly*, *78*(2), 173–193. https://doi.org/10.1177/0190272515581212

Simpson, B., & Walker, H. A. (2002). Status Characteristics and Performance Expectations: A Reformulation. *Sociological Theory*, *20*(1), 24–40. https://doi.org/10.1111/1467-9558.00149

Snijders, T. A. B., & Baerveldt, C. (2003). A multilevel network study of the effects of delinquent behavior on friendship evolution. *The Journal of Mathematical Sociology*, *27*(2–3), 123–151. https://doi.org/10.1080/00222500305892

Spear, M. G. (1984a). The Biasing Influence of Pupil Sex in a Science Marking Exercise. *Research in Science & Technological Education*, *2*(1), 55–60. https://doi.org/10.1080/0263514840020107

Spear, M. G. (1984b). Sex bias in science teachers' ratings of work and pupil characteristics‡. *European Journal of Science Education*, *6*(4), 369–377. https://doi.org/10.1080/0140528840060407

Steegh, A. M., Höffler, T. N., Keller, M. M., & Parchmann, I. (2019). Gender differences in mathematics and science competitions: A systematic review. *Journal of Research in Science Teaching*, *56*(10), 1431–1460. https://doi.org/10.1002/tea.21580

Steffens, M. C., & Jelenec, P. (2011). Separating Implicit Gender Stereotypes regarding Math and Language: Implicit Ability Stereotypes are Self-serving for Boys and Men, but not for Girls and Women. *Sex Roles*, *64*(5–6), 324–335. https://doi.org/10.1007/s11199-010-9924-x

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology*, *102*(4), 947–963. https://doi.org/10.1037/a0019920

Tajfel, H., & Turner, J. C. (2004). *The Social Identity Theory of Intergroup Behavior*. Psychology Press. https://doi.org/10.4324/9780203505984-16

Tomasetto, C., Galdi, S., & Cadinu, M. (2012). Quando l’implicito precede l’esplicito: Gli stereotipi di genere sulla matematica in bambine e bambini di 6 anni. *Psicologia sociale*, *2*, 169–186. https://doi.org/10.1482/37693

Tomasetto, C., Mirisola, A., Galdi, S., & Cadinu, M. (2015). Parents’ math–gender stereotypes, children’s self-perception of ability, and children’s appraisal of parents’ evaluations in 6-year-olds. *Contemporary Educational Psychology*, *42*, 186–198. https://doi.org/10.1016/j.cedpsych.2015.06.007

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1–67.

van der Vleuten, M., Steinmetz, S., & van de Werfhorst, H. (2018). Gender norms and STEM: The importance of friends for stopping leakage from the STEM pipeline. *Educational Research and Evaluation*, *24*(6–7), 417–436. https://doi.org/10.1080/13803611.2019.1589525

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174–1204. https://doi.org/10.1037/a0036620

Vuletich, H. A., Kurtz-Costes, B., Cooley, E., & Payne, B. K. (2020). Math and language gender stereotypes: Age and gender differences in implicit biases and explicit beliefs. *PLOS ONE*, *15*(9), e0238230. https://doi.org/10.1371/journal.pone.0238230

Wagner, D. G., & Berger, J. (1997). Gender and Interpersonal Task Behaviors: Status Expectation Accounts. *Sociological Perspectives*, *40*(1), 1–32. https://doi.org/10.2307/1389491

Wagner, D. G., Ford, R., & Ford, T. (1986). Can gender inequalities be reduced? *American Sociological Association*, *51*(1), 47–61. https://doi.org/10.2307/2095477

Chapter 4. *Implicit gender-science stereotypes and college major intentions of Italian adolescents*

1. Introduction

The issue of the gender gap in STEM is notorious but still puzzling. In the last years, many studies have investigated the causes and proposed various solutions to reduce the underrepresentation of women in STEM, while the counterpart underrepresentation of men in female-dominated fields has only recently gained attention (Chaffee, Lou, & Noels, 2020; Dunlap & Barth, 2019; Heyder et al., 2017; Kalokerinos et al., 2017).

Various and often interrelated factors explaining this gender gap have been proposed. While early scholars argued that differences in performance, interests, and preferences were due to biological characteristics (Lueptow et al., 1995), now research has focused on socio-cultural factors, especially on gender stereotypes that frame societal expectations on attitudes and behaviours of both women and men (Guiso et al., 2008; Kersey et al., 2019; Reinking & Martin, 2018; Spelke, 2005; Wang & Degol, 2017).

Gender stereotypes, defined as beliefs that people have about the characteristics of males and females (Martin & Dinella, 2001), can be endorsed both at the conscious (explicit) and unconscious (implicit) level. Focusing on the latter, implicit gender-science stereotypes are automatic beliefs about the association between gender and STEM. Previous studies found that implicit gender-science stereotypes influence females' performance in scientific tasks and attitudes toward math (Cvencek et al., 2015; Kiefer & Sekaquaptewa, 2007; Nosek & Smyth, 2011; Smeding, 2012). All these factors eventually contribute to women's progressive abandonment of the scientific field, both in educational and professional choices.

As theorized by Eccles' model of achievement-related choices (Eccles, 1983), stereotypes would indirectly influence career-related and educational choices through expectations of success and the values attached to the viable options, i.e., perceived utility, interest, identification, and the cost related to the choice. While studies on implicit gender-science stereotypes report an indirect association of these stereotypes with major choice (e.g., Cundiff et al., 2013), some studies on explicit gender-science stereotypes also report a direct association between stereotypes and major choice (e.g., Plante et al., 2013).

Furthermore, while the focus of academic research and policies has been on women's aspirations in the STEM sector, the gender gap in humanities-related sectors requires careful attention. As a matter of fact, instruments measuring implicit gender-science stereotypes report the difference in the time needed to associate men with STEM and women with humanities, and the time needed to associate men with humanities and women with STEM. However, evidence of the link between (implicit) gender-science stereotypes and men's intentions to choose an educational or professional path in humanities is almost absent (Chaffee, Lou, & Noels, 2020; Kalokerinos et al., 2017).

This study aims to examine the direct and indirect association between implicit gender-science stereotypes and major choice intentions of both female and male high-school students. It contributes to the existing literature in two ways: first, it includes both male students and female-dominated majors. It then provides evidence of the direct association between implicit gender-science stereotypes and major intentions.

2. Literature review

2.1. The gender gap in STEM majors

At the end of high school, students usually face the first important choice of their life, i.e., whether they want to enrol at university or start a professional path. In both cases, their choice depends on personal characteristics, extrinsic factors, e.g., opportunities and salaries, and interpersonal relationships (Akosah-Twumasi et al., 2018). In this context, women and men differ both in their choices and in the type and the extent to which various factors are involved in the decision.

In almost all countries, men tend to choose more prestigious academic tracks, which are also those more math- and science-intensive (Buser et al., 2012). Conversely, women are more inclined to choose majors with a strong social and communal component and/or majors in which reading and language skills are required (AMACAD, 2015; Okahana & Zhou, 2017). If we look at the distribution of men and women in STEM-related majors, the underrepresentation of women appears prominent. In Italy, female undergraduates are 14% in IT, 26% in engineering, and 31% in physics. The opposite pattern is observed in female-dominated fields. Male undergraduates are 7% in education, 15% in modern languages, 17% in psychology, and 39 % in literary studies (AlmaLaurea, 2021).

2.2. Implicit gender-science stereotypes

One of the most relevant factors explaining the gender gap in STEM is the endorsement of gender-science stereotypes (Wang & Degol, 2017). Gender stereotypes can be defined as beliefs that people have about the characteristics of males and females (Martin & Dinella, 2001). They can be endorsed both at the explicit and the implicit level. In the first case, the endorsement is conscious while in the second case it is unconscious and automatic (Whitley & Kite, 2016).

In the context of STEM, gender stereotypes are beliefs that associate the STEM domain with men. Explicit gender-science stereotypes usually assess beliefs about the different abilities of women and men, the former being naturally talented for reading or language-related tasks, the latter being naturally talented for maths-related tasks (Schmader et al., 2004). Conversely, implicit gender-science stereotypes usually measure the automatic association of men with the STEM domain, and women with the humanities domain (Nosek et al., 2009). This study focuses on implicit gender-science stereotypes.

Previous studies tested the influence of these stereotypes on several factors. Nosek and Smyth (2011) found that stronger implicit stereotypes were associated with women's greater negativity toward math, weaker self-ascribed ability, and worse performance in math. Similarly, Ramsey and Sekaquaptewa (2011) found a negative relationship between an increase in implicit stereotypes and performance in a maths course for female college students, while Cvencek et al. (2015) found a negative association between implicit stereotypes and maths performance among elementary-school children. Finally, Steffens et al. (2010) found that implicit math-gender stereotypes predicted girls' implicit maths self-concepts, i.e., the automatic association of the concepts 'I' vs 'other' with STEM vs humanities fields, academic achievement, and enrolment preferences.

2.3. Implicit gender stereotypes and educational choices

Empirical evidence has confirmed the existence of an association between (implicit) gender-science stereotypes and educational choices in STEM. Two studies (Dunlap & Barth, 2019; Smeding, 2012) found that female students majoring in STEM held weaker implicit gender-science stereotypes than women majoring in more feminine fields. Smyth et al. (2009) found that female college students with stronger implicit gender-science stereotypes were less likely to major in science, while the opposite occurred for male students, who were conversely more likely to attend science majors. Interestingly, implicit gender stereotypes were a stronger correlate of science major than was maths SAT.

The mechanism through which gender stereotypes would affect educational (and career-related) choices is described by the model of achievement-related choices proposed by Eccles (1987, 1994). According to this model, educational and vocational choices are determined by both self-efficacy, i.e., the expectation of success in the task and confidence in personal abilities, and what Eccles calls subjective task values (STV), defined as ‘the value individuals attach to various achievement-related options they believe are available to them’ (Eccles et al., 1999, p. 163). STVs are further dived into four components, i.e., (1) the utility value of the task in facilitating goals’ achievement, (2) interest and enjoyment of the task, (3) attainment value, i.e., ‘the value an activity has in manifesting one’s social or personal identities and core values’ (Eccles, 2011, p. 197), and (4) the cost deriving from engaging in the task.

Both expectations of success and STVs are the outcome of personal experiences and perceptions that are, in turn, determined by stereotypes and other relevant people’s beliefs and behaviours (e.g., parents, teachers, role models, and peers). Therefore, according to this model, gender stereotypes would indirectly influence educational choices by influencing self-efficacy, perceived utility of the course, interest in the subject, identification with the course’s field, and the cost of choosing one course rather than another.

Previous studies found evidence for the application of Eccles’ model in the context of the STEM gender gap. In a study on college students, Lane et al. (2012) found that the gender gap in students’ intentions to pursue science was completely accounted for by implicit gender stereotypes. However, for women, this effect was mediated by implicit identification with science. Similarly, Young et al. (2013) found an indirect and negative influence of implicit science stereotypes on women’s career aspirations. The effect was mediated by both implicit and explicit attitudes toward and identification with science, the former referring to positive or negative feelings towards the scientific field (e.g., ‘I very much like doing science’) and the latter to the association between the self and the scientific field (e.g., ‘In general, being a science student is an important part of my current self-image’). Cundiff et al. (2013) found that among women enrolled in an introductory science course, stronger implicit gender-science stereotypes were associated with weaker science identification and, in turn, weaker science career aspirations.

Despite the assumptions of Eccles’ model, in some studies both direct and indirect effects of gender-science stereotypes were found on career intentions (e.g., Plante et al., 2013). However, stereotypes endorsement in those studies was measured using self-reported rather than implicit instruments. It is thus unclear whether implicit gender stereotypes are also directly associated with college-major choices. Furthermore, in all studies testing the association between implicit gender-science stereotypes and major choices the sample consisted of undergraduate or graduate students, in some cases already majoring in STEM. It would be interesting to test whether the association holds also for younger students who are required to choose what they want to major in. In this context, Eccles’ model would help predict students’ final choices.

2.4. Men’s gender stereotypes and educational choices

As mentioned before, implicit gender-science stereotypes are usually derived from the comparison of the propensity to automatically associate men with STEM and women with humanities and the propensity to associate men with humanities and women with STEM. What is measured is then the relative strengths of pairs of associations rather than the absolute strength of a single association (Greenwald & Farnham, 2000). We talk, in this case, about ‘complementary stereotypes’ (Jost & Kay, 2005), one attributing a strength to males (and a weakness to females), i.e., having a feel for maths/science, the other attributing a strength to females (and a weakness to males), i.e., having a feel for reading/verbal tasks.

Interestingly, Plante et al. (2009) found that when asked about the maleness and femaleness of math and language domains, elementary and high school Canadian students did not believe that mathematics was a male domain, while they clearly viewed language as a female domain. Conversely, in a study testing implicit gender stereotypes separately for math domain and language domain on a sample of adolescents and college students, Steffens and Jelenec found that males but not females endorsed implicit math-male stereotypes. On the contrary, females revealed strong language-female implicit stereotypes, whereas males showed language-male counterstereotypes, i.e., the strong association between men and language.

The influence gender-science stereotypes exert on men was rarely of interest. Some studies on the STEM domain on a mixed-gender sample reported also results for male participants. For instance, the study conducted by Cundiff et al. (2013), mentioned before, reported an opposite result for men compared to women. Stronger implicit gender-science stereotypes were associated with higher science identification and, in turn, with stronger science career aspirations. However, the gender gap in humanities, favouring women, has only recently gained attention. Therefore, only a few studies tested whether gender stereotypes are keeping men away from female-dominated majors and careers (Chaffee, Lou, & Noels, 2020; Chaffee, Lou, Noels, et al., 2020).

Furthermore, the existing studies were more interested in traditional gender ideologies on occupational choice, rather than specifically on gender-science stereotypes. This study contributes to the growing literature on gender stereotypes and men's choices by giving information on the link between implicit gender-science stereotypes and major intentions of men in the humanities field.

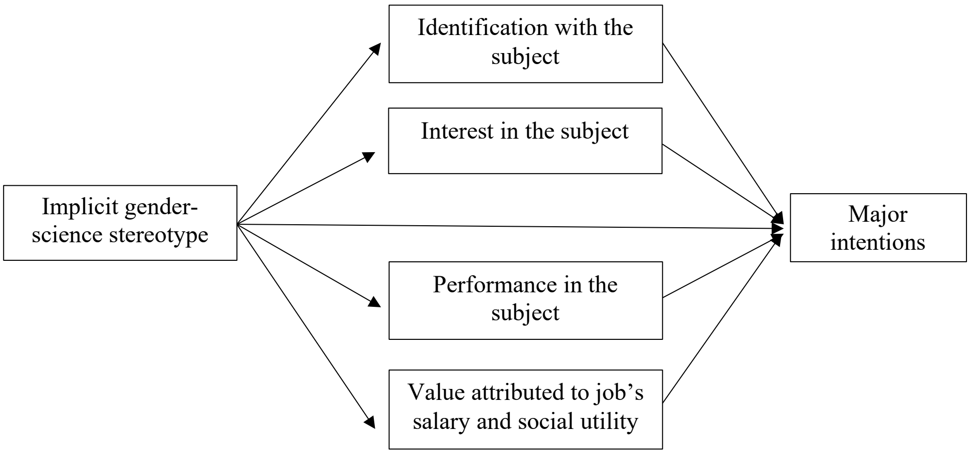
3. Research questions and hypotheses

This study aims to investigate the association between implicit gender-science stereotypes and major choice intentions of a sample of Italian high-school students. In particular, the three following research questions are considered:

1. Are implicit gender-science stereotypes associated with students' intentions of majoring in STEM and humanities?
2. Are there gender differences in the way and the extent to which implicit gender-science stereotypes are associated with major intentions?
3. Is the association mediated by identification with the field (STEM/humanities), interest in the field, performance in the related school subject, and value attributed to job salary (social utility)?

Figure 4.1 depicts the estimated model, with the direct and indirect links taken into account.

Figure 4.1 *The estimated model with direct and indirect links*



As regards questions 1 and 2, it was expected that implicit gender-science stereotypes were associated with major intentions in both male- and female-dominated fields, with different, opposite results for men and women. In particular, it was hypothesized a positive association in humanities for female students, negative for male students, the opposite association in STEM.

As regards question 3, based on Eccles' model (1987, 1994), I expected that gender stereotypes were indirectly associated with major intentions and that their effect was moderated by identification with the field, interest, performance, and value attributed to job salary (social utility in case of humanities).

4. Methodology

4.1. Participants

The sample consists of 302 Italian students (61% females), aged 18 years old, attending the last year of high school (Level 3, ISCEED 2011) and coming from five schools located in Milan and the surrounding area. The sample is the same of that in Chapter 3, however, it also includes students that participated to the data collection individually rather than by class and were thus not asked about the hypothetical competition, presented in Chapter 3.

As specified in Chapter 3, students in the sample came from either a *liceo classico*, *liceo scientifico,* or a *liceo linguistico*. The belonging school type does not preclude college-major choice, but it is likely to reflect students' inclinations and interests. Consequently, most students in humanities-oriented schools were females (83%), while the opposite occured in science-oriented schools (32%). Furthermore, students tend to choose a college major coherent with what they studied in high school. This implies that it is indispensable to control for the type of school when analyzing the association between gender and major choice (more details in the Appendix).

4.2. Instruments

Data were collected through a computer-assisted questionnaire administered in the participants' schools during regular school time. Students were asked to sign an informed consent before answering the questionnaire.

Major intentions. Students were asked whether they intended to enrol in university after high school and, if so, to indicate, among a list of all possible majors, those they were taking into account. The list contained 51 majors taken from the Italian official national list which were then classified into 7 macro-areas, i.e., health, vet or agrarian, STEM, law, economic/statistic, sociopolitical, arts and humanities. These macro-areas reflect the Italian disciplinary groups in which degree programs are organized at the national level. Those related to STEM and humanities were further grouped to be coherent with the international definition of these domains. The two groups of STEM and humanities were used, separately, as dependent variables, with a value equal to 1 if the student was interested in at least one of the majors in that field, 0 otherwise.

Independent variable. The main variable of interest was the implicit association between gender and majors. Students were asked to perform the Implicit Association Test (Greenwald et al., 1998), which is commonly used to assess automatic gender stereotypes (Greenwald et al., 2009). I adopted the version of the IAT used in the Project Implicit website for gender-science stereotypes (Nosek et al., 2009), where the target – gender – is represented by male and female names, while the attribute – majors – by STEM-related and humanities-related majors. The variable used in the analysis was the test score, i.e., the so-called IAT D measure (Greenwald et al., 2003), ranging from -2 to 2. More specifically, a negative value in the scale means that the student associated math with females and/or humanities with males more easily (counterstereotypical association), while a positive value means that the student associated math with males and/or humanities with females more easily (stereotypical association). Finally, values around zero indicate that the student did not show any associations between gender and majors (more details on the procedure and stimuli in the Appendix).

Mediators. Four mediating variables were included to test the indirect effect of stereotypes on major choice, i.e., identification with the subject, interest in the subject, performance, and value attributed to job salary or social utility.

Identification with STEM/humanities was measured using a battery of items suggested by (Brown & Josephs, 1999). This included: relevance of the subject, relevance attributed to others' opinion about personal 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 (see Chapter 3 Appendix for the full list of items) All five questions were asked for Italian, maths, and science subjects, separately, on a 5-point Likert scale (Cronbach alpha 0.9 in Italian, 0.8 in maths, 0.9 in science). Due to multicollinearity issues, math and science were combined to create a unique variable of identification with STEM (Cronbach alpha 0.9).

Interest in the subject was assessed by asking students to indicate, among a list of all possible subjects, their favourite ones. Two binary variables were then created grouping subjects in the STEM and humanities areas, with 1 if the student indicated at least one belonging to that area as a favourite subject, 0 otherwise.

Performance was derived from the final grade obtained in the previous academic year in Italian, maths, and science, as reported by students in the questionnaire. The variable on performance was continuous and ranged from 5 to 10 (a more refined discussion on school grades can be find in Chapter 3).

Finally, to obtain an indicator of the value attributed to jobs' salary and social utility, students were asked to indicate the factors that played a relevant role in their decision process for major choice among a list of potential. The list included factors on a future professional path (i.e., variety of career opportunities, job' social utility, salary, prestige), personal characteristics (i.e., interest in the subject, realize a dream, performance in the subject, relevant others' expectations and suggestions), and factors pertaining the program of study (i.e., challenging courses, number of years for graduating, course competitiveness, entry test). Two binary variables were created for salary and social utility, with 1 if the student selected that factor, 0 otherwise.

Control variables. The final model includes mother's level of education and type of school as control variables. The first is a categorical variable with three categories, i.e, below high school (value 1), completed high school or equivalent vocational school (value 2), degree, and post-degree (value 3).The second is a binary variable taking value equal to 1 for humanities-oriented type of schools (i.e., *liceo classico* and *liceo linguistico*) and value equal to 2 for STEM-oriented schools (i.e., *liceo scientifico*).

4.3. Analytical methods

I examined the association between implicit gender-science stereotypes and major intentions using two logistic regression analyses, i.e., one on STEM majors and the other on humanities majors. Since students in the sample were nested in school classes, clustered standard errors were used, computed using the percentile cluster bootstrap technique, suggested in the case of few clusters (Cameron et al., 2008; Cameron & Miller, 2015). More details are reported in the Appendix.

The analysis was performed in three steps separately for both dependent variables. First, to detect gender differences in the effect of regressors on major choice, a regression model including control and moderating variables was performed and each variable interacted with gender. Subsequently, a model including the implicit stereotype score and its interaction with gender was performed. Finally, a mediation analysis was used to compute the direct and indirect effects of implicit gender stereotypes via the four mediating variables. In this case, percentile bootstrap 95% confidence intervals were used for significance testing (Preacher & Hayes, 2008).

Data analysis was performed using Stata (StataCorp, 2021). Missing data were between 4% (IAT D score) and 10% (identification with STEM/humanities) of the total and were estimated using multiple imputations by chained equations (Azur et al., 2011). The number of imputations was set to 20 and the variables included in the imputation model are the same as those included in the regression model. Note that due to incompatibility between the commands available to compute indirect effects in the case of logistic regression models and the commands used to impute missing values, the mediation analysis is based on the non-imputed dataset. However, results from the regression models estimated using the imputed and non-imputed datasets are similar (see Table 4.9 in the Appendix).

5. Results

5.1. Logistic regression analysis

Regression analysis including an interaction term of gender with all the variables in the model indicated that there was no statistically significant gender difference in the association between the dependent variable and these factors. Therefore, only the interaction of gender with implicit stereotypes was included in the final model. Table 4.1 shows the results from the analysis of STEM majors' intentions.

Table 4.1 *Results from logistic regression analysis on STEM majors' intentions*

	Estimate	SE	Percentile bootstrap 95% CI		p-value
			LB	UB	
STEM					
Gender [Female]	.24	.49	-.47	1.15	.625
IAT D score	.35	.49	-.74	1.58	.472
Gender*IAT D	-1.44	.65	-3.39	-.22	.026
School grade in science	.29	.13	-.07	.74	.025
Identification with STEM	.91	.35	.39	1.78	.010
Interest in STEM subjects	1.06	.42	.24	1.86	.011
Relevance of job's salary	.73	.36	-.14	1.56	.042
Mother's education [a] [below high school]	-.01	.52	-1.31	1.05	.981
Mother's education [a] [high school]	-.52	.41	-1.33	.28	.207
Type of school [b] [STEM-oriented]	1.04	.36	.24	2.06	.004

Note. N = 302

CI = Confidence interval; LB = Lower bound; UB = Upper bound

[a] Reference category 'Above high school'

[b] Reference category 'Humanities-oriented'

Results suggest that identification with STEM, interest in STEM subjects at school and school grade in science were positively associated with intentions to choose STEM majors.

However, only the first two are statistically significant. The association between major intentions and the influence of the job's salary on the choice was positive and statistically significant. Not surprisingly, students attending a STEM-oriented high school were more likely to express the intention of majoring in STEM compared to those attending a humanities-oriented high school. The association was statistically significant.

Table 4.2 *Results from logistic regression analysis on humanities majors' intentions*

	Estimate	SE	Percentile bootstrap 95% CI		p-value
			LB	UB	
Humanities					
Gender [Female]	.45	.34	-.43	1.15	.192
IAT D score	-.89	.53	-2.02	.32	.087
Gender*IAT D	1.30	.49	-.02	2.95	.007
School grade in Italian	-.16	.10	-.54	.12	.111
Identification with humanities	.53	.22	.14	.98	.017
Interest in humanities-related subjects	.37	.31	-.26	1.15	.240
Relevance of job's utility	.35	.27	-.21	.99	.198
Mother's education [a] [below high school]	-.05	.46	-1.22	1.36	.919
Mother's education [a] [high school]	-.60	.27	-1.32	.13	.025
Type of school [b] [STEM-oriented]	-.78	.33	-1.56	-.19	.017

Note. Total N = 302

SE = Cluster-robust standard error; CI= Confidence interval; LB = Lower bound; UB = Upper bound

[a] Reference category 'Above high school'

[b] Reference category 'Humanities-oriented'

Results in Table 4.2 on humanities depicts a different picture. Among the included factors, only identification with the field and the type of school were statistically significant, while for the other factors we cannot reject the null hypothesis of no association. As expected, a higher identification with the humanistic field is associated with a higher probability of expressing the intention of majoring in humanities. Furthermore, students attending a humanities-oriented high school were more likely to be interested in majoring in humanities compared to those attending a STEM-oriented high school.

As regards gender and implicit gender stereotypes they were both not statistically significant in either model, while the interaction term was statistically significant in both models. The absence of a statistically significant result for the main effects suggests that implicit stereotypical beliefs have the opposite association for the two genders. Marginal effects were computed to understand how males and females differ and whether the association was statistically significant. Table 4.3 shows the average marginal effect (AME) of implicit gender stereotypes on major choice intentions for males and females. Confidence intervals (CI) can be used to determine the significance of the marginal effect, i.e., when the CI did not include zero, the effect was statistically significant.

Table 4.3 *AME of implicit stereotypes on major intentions by gender*

	AME	SE	Percentile bootstrap 95% CI	
			LB	UB
STEM				
Male	.052	.073	-.127	.233
Female	**-.151**	.047	-.346	-.056
Humanities				
Male	-.107	.067	-.260	.032
Female	.106	.074	-.096	.312

Note. AME = Average marginal effect; SE = Bootstrap standard error; CI = Confidence interval; LB = Lower bound; UB = Upper bound

As expected, implicit gender stereotypes were negatively associated with major choice intentions for male students, positively for female students in the case of STEM majors. The opposite occurred in the case of humanities, i.e., the stronger the stereotypical association, the less likely the intention was for male students to choose a major in humanities, whereas the more likely the intention was for female students to choose a major in that field. However, confidence intervals suggest that the association between implicit gender stereotypes and major choice intentions was statistically significant only for females in STEM.

5.2. Mediation analysis

Finally, a mediation analysis was performed to check whether the association between implicit gender stereotypes and major choice intentions was mediated by other factors. Given the gender differences in the association between implicit gender stereotypes and the variable of interest, the analysis was conducted separately on males and females. The impossibility of assessing an association between implicit gender stereotypes and intentions of majoring in humanities was confirmed, as for both women and men, there was neither a direct nor an indirect link between the two. The same held for men's intentions of majoring in STEM. As regards women, percentile bootstrap 95% confidence intervals suggest that the data could not reject the null hypothesis of no indirect effect. However, the direct significant association between implicit gender stereotypes and intentions of majoring in STEM was confirmed. Table 4.4 reports the results of the mediation analysis on intentions of majoring in STEM for women.

Table 4.4 *Mediation analysis on women's intentions of majoring in STEM*

	Estimate	SE	Percentile bootstrap 95% CI	
			LB	UB
Indirect effects				
School grade in science	.013	.018	-.006	.058
Identification with STEM	-.009	.038	-.073	.073
Interest in STEM subjects	-.027	.033	-.091	.046
Relevance of job's salary	.059	.042	-.018	.144
Total indirect	.036	.066	-.089	.179
Direct effect	-.237	.086	-.406	-.073
Total effect	-.201	.086	-.357	-.020

Note. SE = Bootstrap standard errors; CI = Confidence interval; LB = Lower bound; UB = Upper bound

Proportion of total effect mediated = -.180

Ratio of indirect to direct effect = -.153

Ratio of total to direct effect = .847

6. Discussion

The gender gap in STEM has progressively gained so much attention to stimulate several initiatives aimed to reduce the underrepresentation of women in this field (UNESCO, 2017).

Gender stereotypes are usually considered one of the root causes of this gender gap so various initiatives aiming to change stereotypical beliefs on gendered abilities in math and science have been undertaken. However, empirical findings on the association between implicit gender stereotypes and majors are still few, with mixed results, and restricted to the case of women and STEM, leaving aside stereotypes of men in female-dominated fields. This limits our understanding of whether and how stereotypes influence women's and men's professional paths.

The aim of the study was threefold. First, to verify the association between implicit gender-science stereotypes and major choice intentions. Then to detect possible gender differences in this association. Finally, to understand whether the association was direct, indirect, or both direct and indirect. The analysis was conducted on a sample of Italian high-school students using logistic regression models and mediation analysis.

Results revealed a statistically significant association of gender stereotypes with major choice intentions only in the STEM field for female students. Unlike previous studies (Dunlap & Barth, 2019; Plante et al., 2013) and Eccles' model (1987, 1994) assumptions, in this study, results from the mediation analysis could not reject the hypothesis of no indirect effects for any of the included mediating variables. As regards gender differences, results suggest that the association between implicit gender stereotypes and major choice intentions is positive for male students in the STEM field and female students in the humanities field. It is negative for female students in STEM and male students in humanities. However, the association was statistically significant only for female students in STEM.

The study contributes to the scarce literature on the effect of gender stereotypes on men in female-dominated fields. In particular, results suggest that men endorsing strong stereotypical beliefs are more likely to choose a STEM major and less likely to choose a major related to humanities. However, data could not reject the null hypothesis of no effect of implicit gender stereotypes on men. Future research should examine the mechanism through which stereotypes may influence men's choices and whether exposure to role models could be effective in changing these choices.

Having said this, this study also has limitations. First, the small size of the sample and the cross-sectional nature of data, in particular, limit any generalization. On the one hand, the small sample implies low statistical power, biased effects size estimation, and low reproducibility

(Button et al., 2013; Colquhoun, 2014). The cross-sectional nature of data hampers the possibility to make any inference on causal relationships. As noted by Smyth et al. (2009), longitudinal data has advantages when – as here – the purpose is to understand how stereotypical beliefs evolve and therefore when and how they can be reduced and eventually eliminated. While female students in this sample were more stereotyped than their male peers, these differences disappeared when we take into account the type of high school attended. With a cross-sectional dataset, it is impossible to determine whether female students in humanities-oriented schools were already more stereotyped or strengthened their endorsement of gender stereotypes whenever attending a female-dominated school. Since most studies on stereotypes rely on this type of data, future research should focus on longitudinal data to shed light on the causal inference in the identified paths.

7. Conclusions

The study confirmed the association between implicit gender-science stereotypes and college majors found in previous studies for female students (Jugović, 2017; Smyth et al., 2009). However, unlike other studies (Dunlap & Barth, 2019; Plante et al., 2013; Schuster & Martiny, 2017; Vleuten et al., 2016), only the direct effect resulted statistically significant.

The direct association of gender stereotypes with major intentions could have relevant implications for initiatives and policies aimed at reducing the gender gap in STEM. If the effect of stereotypes is not mediated by other factors, reducing the stereotypical association of STEM with men should in theory be effective in changing women's participation in STEM courses and careers. On the contrary, in the case of an indirect effect mediated by identification with the subject or self-efficacy, changing the representation of women and men in STEM would not be sufficient to have an impact on women's choices. In this case, interventions should reinforce the individual relationship of women with STEM, whereas showing how other women are successfully involved in STEM activities could be ineffective (Olsson & Martiny, 2018).

Finally, while the data allowed to confirm the association of gender-science stereotypes with major intentions only for female students, further research is necessary to verify whether men could be influenced by gender stereotypes in their attitudes toward female-dominated sectors.

References

Akosah-Twumasi, P., Emeto, T. I., Lindsay, D., Tsey, K., & Malau-Aduli, B. S. (2018). A Systematic Review of Factors That Influence Youths Career Choices—The Role of Culture. *Frontiers in Education*, *3*, 58. https://doi.org/10.3389/feduc.2018.00058

AlmaLaurea. (2021). *XXIII Indagine Profilo dei Laureati*. AlmaLaurea. https://www.almalaurea.it/universita/indagini/laureati/profilo

AMACAD. (2015). *Gender Distribution of Bachelor's Degrees in the Humanities* [Humanities Indicators]. American Academy of Arts and Sciences. https://www.amacad.org/humanities-indicators/higher-education/gender-distribution-bachelors-degrees-humanities#31609

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: What is it and how does it work?: Multiple imputation by chained equations. *International Journal of Methods in Psychiatric Research*, *20*(1), 40–49. https://doi.org/10.1002/mpr.329

Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of Personality and Social Psychology*, *76*(2), 246–257. https://doi.org/10.1037/0022-3514.76.2.246

Buser, T., Niederle, M., & Oosterbeek, H. (2012). *Gender, Competitiveness and Career Choices* (Working Paper No. 18576). National Bureau of Economic Research. https://doi.org/10.3386/w18576

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, *90*(3), 414–427. https://doi.org/10.1162/rest.90.3.414

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, *50*(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

Chaffee, K. E., Lou, N. M., & Noels, K. A. (2020). Does Stereotype Threat Affect Men in Language Domains? *Frontiers in Psychology*, *11*, 1302. https://doi.org/10.3389/fpsyg.2020.01302

Chaffee, K. E., Lou, N. M., Noels, K. A., & Katz, J. W. (2020). Why don't "real men" learn languages? Masculinity threat and gender ideology suppress men's language learning motivation. *Group Processes & Intergroup Relations*, *23*(2), 301–318. https://doi.org/10.1177/1368430219835025

Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of *p* -values. *Royal Society Open Science*, *1*(3), 140216. https://doi.org/10.1098/rsos.140216

Cundiff, J. L., Vescio, T. K., Loken, E., & Lo, L. (2013). Do gender–science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Social Psychology of Education*, *16*(4), 541–554. https://doi.org/10.1007/s11218-013-9232-8

Cvencek, D., Kapur, M., & Meltzoff, A. N. (2015). Math achievement, stereotypes, and math self-concepts among elementary-school students in Singapore. *Learning and Instruction*, *39*, 1–10. https://doi.org/10.1016/j.learninstruc.2015.04.002

Dunlap, S. T., & Barth, J. M. (2019). Career Stereotypes and Identities: Implicit Beliefs and Major Choice for College Women and Men in STEM and Female-Dominated Fields. *Sex Roles*, *81*(9–10), 548–560. https://doi.org/10.1007/s11199-019-1013-1

Eccles, J. (1983). Expectancies, Values, and Academic Behaviors. In *Achievement and Achievement Motives: Psychological and sociological approaches* (J.T. Spence (Ed.), pp. 75–146). San Francisco, CA: Free man.

Eccles, J. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development*, *35*(3), 195–201. https://doi.org/10.1177/0165025411398185

Eccles, J. S. (1987). Gender Roles and Women's Achievement-Related Decisions. *Psychology of Women Quarterly*, *11*(2), 135–172. https://doi.org/10.1111/j.1471-6402.1987.tb00781.x

Eccles, J. S. (1994). Understanding Women's Educational And Occupational Choices: Applying the Eccles et al. Model of Achievement-Related Choices. *Psychology of Women Quarterly*, *18*(4), 585–609. https://doi.org/10.1111/j.1471-6402.1994.tb01049.x

Eccles, J. S., Barber, B., & Jozefowicz, D. (1999). Linking gender to educational, occupational, and recreational choices: Applying the Eccles et al. model of achievement-related choices. In W. B. Swann, J. H. Langlois, & L. A. Gilbert (Eds.), *Sexism and stereotypes in modern society: The gender science of Janet Taylor Spence.* (pp. 153–192). American Psychological Association. https://doi.org/10.1037/10277-007

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*(6), 1022–1038. https://doi.org/10.1037/0022-3514.79.6.1022

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41. https://doi.org/10.1037/a0015575

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). DIVERSITY: Culture, Gender, and Math. *Science*, *320*(5880), 1164–1165. https://doi.org/10.1126/science.1154094

Heyder, A., Kessels, U., & Steinmayr, R. (2017). Explaining academic-track boys' underachievement in language grades: Not a lack of aptitude but students' motivational beliefs and parents' perceptions? *British Journal of Educational Psychology*, *87*(2), 205–223. https://doi.org/10.1111/bjep.12145

Jost, J. T., & Kay, A. C. (2005). Exposure to Benevolent Sexism and Complementary Gender Stereotypes: Consequences for Specific and Diffuse Forms of System Justification. *Journal of Personality and Social Psychology*, *88*(3), 498–509. https://doi.org/10.1037/0022-3514.88.3.498

Jugović, I. (2017). Students' Gender-Related Choices and Achievement in Physics. *Center for Educational Policy Studies Journal*, *7*, 71–95. https://doi.org/10.26529/cepsj.170

Kalokerinos, E. K., Kjelsaas, K., Bennetts, S., & Hippel, C. von. (2017). Men in pink collars: Stereotype threat and disengagement among male teachers and child protection workers. *European Journal of Social Psychology*, *47*(5), 553–565. https://doi.org/10.1002/ejsp.2246

Kersey, A. J., Csumitta, K. D., & Cantlon, J. F. (2019). Gender similarities in the brain during mathematics development. *Npj Science of Learning*, *4*(1), 19. https://doi.org/10.1038/s41539-019-0057-x

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*(5), 825–832. https://doi.org/10.1016/j.jesp.2006.08.004

Lane, K. A., Goh, J. X., & Driver-Linn, E. (2012). Implicit Science Stereotypes Mediate the Relationship between Gender and Academic Participation. *Sex Roles*, *66*(3–4), 220–234. https://doi.org/10.1007/s11199-011-0036-z

Lueptow, L. B., Garovich, L., & Lueptow, M. B. (1995). The persistence of gender stereotypes in the face of changing sex roles: Evidence contrary to the sociocultural model. *Ethology and Sociobiology*, *16*(6), 509–530. https://doi.org/10.1016/0162-3095(95)00072-0

Martin, C. L., & Dinella, L. (2001). Gender-related Development. In *International Encyclopedia of the Social & Behavioral Sciences* (pp. 6020–6027). Elsevier. https://doi.org/10.1016/B0-08-043076-7/01684-3

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., … Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593–10597. https://doi.org/10.1073/pnas.0809921106

Nosek, B., & Smyth, F. L. (2011). Implicit Social Cognitions Predict Sex Differences in Math Engagement and Achievement. *American Educational Research Journal*, *48*(5), 1125–1156. https://doi.org/10.3102/0002831211410683

Okahana, H., & Zhou, E. (2017). *Graduate Enrollment and Degrees: 2006 to 2016*. Washington, DC: Council of graduate schools. https://cgsnet.org/publication-pdf/4642/CGS_GED16_Report_Final.pdf

Olsson, M., & Martiny, S. E. (2018). Does Exposure to Counterstereotypical Role Models Influence Girls' and Women's Gender Stereotypes and Career Choices? A Review of Social Psychological Research. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02264

Plante, I., de la Sablonnière, R., Aronson, J. M., & Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. *Contemporary Educational Psychology*, *38*(3), 225–235. https://doi.org/10.1016/j.cedpsych.2013.03.004

Plante, I., Théorêt, M., & Favreau, O. E. (2009). Student gender stereotypes: Contrasting the perceived maleness and femaleness of mathematics and language. *Educational Psychology*, *29*(4), 385–405. https://doi.org/10.1080/01443410902971500

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Ramsey, L. R., & Sekaquaptewa, D. (2011). Changing stereotypes, changing grades: A longitudinal study of stereotyping during a college math course. *Social Psychology of Education*, *14*(3), 377–387. https://doi.org/10.1007/s11218-010-9150-y

Reinking, A., & Martin, B. (2018). The Gender Gap in STEM Fields: Theories, Movements, and Ideas to Engage Girls in STEM. *Journal of New Approaches in Educational Research*, *7*(2), 148–153. https://doi.org/10.7821/naer.2018.7.271

Schmader, T., Johns, M., & Barquissau, M. (2004). The Costs of Accepting Gender Differences: The Role of Stereotype Endorsement in Women's Experience in the Math Domain. *Sex Roles*, *50*(11), 835–850. https://doi.org/10.1023/B:SERS.0000029101.74557.a0

Schuster, C., & Martiny, S. E. (2017). Not Feeling Good in STEM: Effects of Stereotype Activation and Anticipated Affect on Women's Career Aspirations. *Sex Roles*, *76*(1–2), 40–55. https://doi.org/10.1007/s11199-016-0665-3

Smeding, A. (2012). Women in Science, Technology, Engineering, and Mathematics (STEM): An Investigation of Their Implicit Gender Stereotypes and Stereotypes' Connectedness to Math Performance. *Sex Roles*, *67*(11–12), 617–629. https://doi.org/10.1007/s11199-012-0209-4

Smyth, F. L., Greenwald, A. G., & Nosek, B. A. (2009). *Implicit gender-science stereotype outperforms math scholastic aptitude in identifying science majors* [Unpublished manuscript]. https://www.colorado.edu/odece/sites/default/files/attached-files/rba13-sb4converted_2.pdf

Spelke, E. S. (2005). Sex Differences in Intrinsic Aptitude for Mathematics and Science?: A Critical Review. *American Psychologist*, *60*(9), 950–958. https://doi.org/10.1037/0003-066X.60.9.950

StataCorp. (2021). *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC.

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *Journal of Educational Psychology*, *102*(4), 947–963. https://doi.org/10.1037/a0019920

UNESCO. (2017). *Cracking the code: Girls' education in science, technology, engineering and mathematics (STEM); report of the UNESCO International Symposium and Policy Forum*. UNESCO International Symposium and Policy Forum. https://unesdoc.unesco.org/ark:/48223/pf0000260079

Vleuten, M. van der, Jaspers, E., Maas, I., & Lippe, T. van der. (2016). Boys' and girls' educational choices in secondary education. The role of gender ideology. *Educational Studies*, *42*(2), 181–200. https://doi.org/10.1080/03055698.2016.1160821

Wang, M.-T., & Degol, J. L. (2017). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. *Educational Psychology Review*, *29*(1), 119–140. https://doi.org/10.1007/s10648-015-9355-x

Whitley, B. E., & Kite, M. E. (2016). *Psychology of prejudice and discrimination* (Third Edition). Routledge/Taylor & Francis Group.

Young, D. M., Rudman, L. A., Buettner, H. M., & McLean, M. C. (2013). The Influence of Female Role Models on Women's Implicit Science Cognitions. *Psychology of Women Quarterly*, *37*(3), 283–292. https://doi.org/10.1177/0361684313482109

Chapter 5. ***Empowering women in STEM: A scoping review of interventions with role models and mentors***

1. Introduction

In 2019, 9.9 million American workers were employed in Science, Technology, Engineering, and Mathematics (STEM), accounting for 6% of the total U.S. workforce. This sector is constantly expanding and jobs are expected to grow by 8% between 2019 and 2029 (U.S. Bureau of Labor Statistics, 2021).

The projected growth rate, increasing number of opportunities, and median salary are more than double that for other fields (U.S. Bureau of Labor Statistics, 2021), making STEM one of the most appealing sectors for college students. However, STEM education is attracting mostly men, while women are still underrepresented. Worldwide, fewer girls than boys aspire to STEM careers, even among top performers, which in turn determines the observed gender gap in tertiary education (Alam & Sanchez Tapia, 2020).

As highlighted in a report on female participation in STEM drawn up by UNESCO (2017), female participation can be influenced by many factors on four distinct levels, i.e., individual, family and peer, school, and society. Therefore, interventions can be quite heterogeneous, as they operate at these different levels and tackle different issues. However, one intervention in particular – exposure to role models – seems to be especially effective in reducing gender inequality. This consists of providing girls and women with female positive examples, usually successfully employed in STEM sectors, who challenge the stereotypical association of men and STEM (Dasgupta & Greenwald, 2001).

Indeed, given that gender stereotypes reflect the unequal representation of women in STEM professions (Eagly & Steffen, 1984), seeing female STEM workers would counteract entrenched stereotypes. Furthermore, these exemplars can also expand women's horizons by showing them pathways they had not previously perceived as being viable (Morgenroth et al., 2015).

However, the efficacy of these interventions is still controversial and few reviews summarized the existing contributions to studies on the effect of role models on female engagement in STEM (Olsson & Martiny, 2018; Prieto-Rodriguez et al., 2020). To fill this gap, this review follows a systematic methodology to summarize studies measuring the effect of role models and mentors addressing women in STEM areas.

2. Background

Although the words 'mentor' and 'role model' are often used interchangeably, the two terms refer to well established yet separately distinct figures. While both act as sources of social influence, a role model has a one-way and passive relationship ((Shapiro, Haseltine, and Rowe 1978), whereas mentors involve the concept of social exchange (Haggard et al. 2011).

Here we follow Fagenson's definition (1989: 312) in describing a mentor as 'someone in a position of power who looks out for you, or gives you advice, or brings your accomplishments to the attention of other people who have power in the company'. To be defined as such, mentors should show the core attributes identified by Haggard et al. (2011), i.e., reciprocity, developmental benefits and regular interactions. Unlike mentors, role models could be defined as individuals who are, or can be, admired and emulated by others. As suggested by Morgenroth et al. (2015), role models can be seen as people who show others (1) how to perform a skill and achieve a goal – behaviour, (2) that a goal is attainable – representation of the possible, (3) that a goal is desirable – inspiration.

Different initiatives have been undertaken to favour female empowerment in STEM via role models and mentors in various countries and communities. However, the results of research on the effect of role models on women are mixed. Some studies found a positive effect, others a negative one and others no effect at all. Furthermore, the effectiveness of role models also varies depending on the context in which the exposure is tested. When assigned to a female experimenter presented as highly competent in maths, female undergraduates performed better in a maths test compared to those assigned to a female experimenter described as not competent (Marx and Roman 2002). After a one-hour visit of a female scientist, female high-school students showed weaker stereotypes on gender and STEM and were more likely to apply and be admitted to a selective science major in college (Breda et al. 2020).

However, the efficacy of role models' intervention seems to depend on the characteristics of both the role models and the target. For example, Cheryan et al. (2013) engaged female undergraduates in interacting with senior students majoring in computer science, who shared information on their major and personal life. They manipulated role models' perception as 'nerds' through hobbies and clothes. Students who interacted with a woman embodying computer science stereotypes were less interested in majoring in computer science and, moreover, showed a weaker sense of belongingness in the field, compared to those who interacted with a non-stereotypical woman. Surprisingly, interacting with a non-stereotypical man was more effective in increasing students' interest than interacting with a stereotypical woman.

Given this heterogeneity in findings and types of interventions, the current debate on the efficacy of role models on the STEM gender gap would benefit from an exhaustive summary of studies on the theme. However, few reviews exist that unfortunately do not fulfil this purpose. Lenton et al. (2009) conducted a meta-analysis of interventions aimed at reducing automatic gender stereotypes, including – but not limited to – the exposure to female role models. Similarly, Lawner et al. (2019) conducted a meta-analysis on the effect of in-group role models on performance and interest in STEM of students belonging to an under-represented social group. However, they included only studies that reported an effect size of the interventions and, more broadly, concentrated on students belonging to a social group underrepresented in STEM.

Olsson and Martiny (2018) published an overview of research-based interventions, in which girls and women observed and interacted with counterstereotypical role models. Unfortunately, as acknowledged by the authors themselves, this review lacked a systematic approach. Finally, Prieto-Rodriguez et al. (2020) presented a systematic review of studies on STEM interventions targeted at girls. However, their summary included also interventions not involving role models and was restricted to studies on female students in secondary school. To fill this gap, this review summarizes studies measuring the effect of role models and mentors addressing women in STEM areas.

3. The aim of the study

This review aims to summarize studies measuring the effect of role models and mentors addressing women in STEM areas. In particular, the following six characteristics were the focus of the summary:

1. Research type
2. Targeted population
3. Type of examples which participants were exposed to
4. Type and content of the intervention
5. Variables of interest (dependent variables)
6. Effects

4. Methodology

The review was conducted following the guidelines of the JBI Manual for Evidence Synthesis (Peters et al., 2020). I, first, developed an *a priori* research protocol, by defining research questions, inclusion and exclusion criteria, search strategy and data sources. Studies' selection was conducted from March to April 2021 on four online databases (i.e., PsycInfo, Web of Science, Proquest, and Scopus) and through citation searching. Only studies published in English were included.

Starting from the search strategy in Lawner et al. (2019), I established the combination of search terms after three tests on PsycInfo and Proquest. The final combination included three strings: role models, gender, and STEM (see Table 5.4 in the Appendix). Results from the first screening into all four databases were then merged and duplicates removed. Title, abstract, and full-text screening were conducted by two researchers independently and based on the criteria established *a priori* in the research protocol. Any disagreements were resolved by consensus.

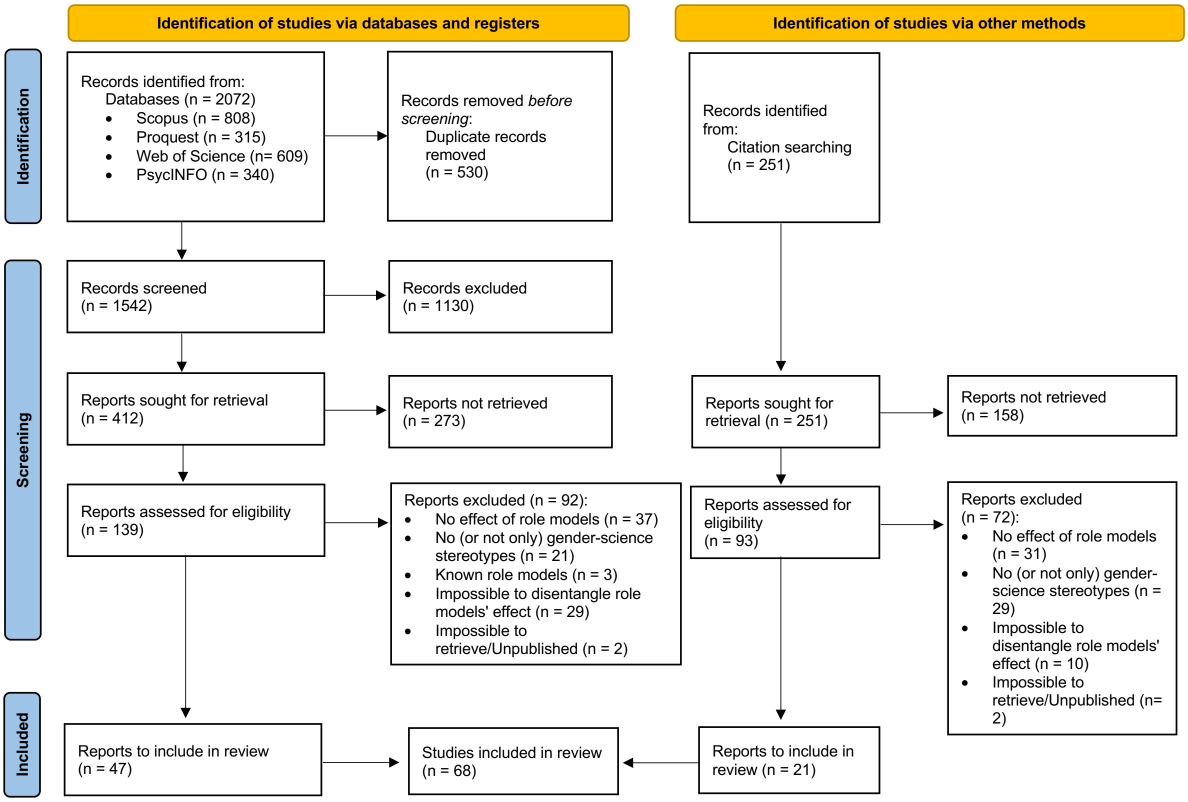
I established non-restrictive inclusion criteria to explore the literature widely. However, two characteristics were particularly relevant, i.e., unfamiliar or unacquainted role models and interventions addressing stereotypes in STEM. This review aims to inform future interventions targeting the female underrepresentation in STEM via exposure to role models. This involved gaining information on interventions that could potentially be replicated elsewhere and with various samples. Parents, peers, as well as teachers cannot be exogenously assigned to people.

This led us to exclude research that examined the effect of these types of role models (see Table 5.8 in the Appendix).

5. *Results*

The database search returned 2,072 references, of which 68 articles, published between 1981 and April 2021, were included in the review.

Figure 5.1 *PRISMA 2020 flow diagram*



5.1. *Research type*

Regarding the research type, most studies adopted a quantitative approach, mainly experimental (72%), while a minority used quasi-experimental (7%) or non-experimental (1%) designs. The remaining studies adopted either a qualitative (8%) or a mixed (10%) approach.

5.2. *Target*

Regarding the study target, the sample mostly consisted of students of various ages. Half of the studies targeted students aged 19 or below, while the other half were addressed to undergraduate and graduate students. Only in six studies, was the sample taken online from Amazon Turk and so varied in age, from 18 to 60 years old.

Regarding nationality, 77% of studies were conducted on individuals from the United States, 6% from France, 4% from Germany, 4% from Australia, and 3% from Canada, with small samples from other countries. 54% of studies targeted only females, while 46% both males and females.

5.3. Examples

In most studies, stereotypical or counterstereotypical exemplars were role models (85%), while in ten studies, participants were supported by mentors. Role models were usually STEM professionals, either from the academic sector or the private sector. In 41% of cases, they were peers at an advanced stage of their studies. For instance, for undergraduate students, the role model could be a senior student in college (Howard, 2015) or one attending an upper-division major (Nickerson et al., 2017). In cases of students in high school or below, role models were college students (Breda et al., 2020; Evans et al., 1995; Merritt et al., 2021).

In 54% of studies, participants were exposed only to women, while in 44% role models were both women and men. Only in one study (Pietri, Drawbaugh, et al., 2021), was the role model a man. There was heterogeneity also in terms of the number of models presented. In 44% of studies, participants were exposed only to one exemplar, whereas in the remaining 56% of studies, they were exposed to more than one.

5.4. Interventions

It was rare to find two or more studies investigating the effects of the same intervention, although, we can identify certain commonalities. I classified interventions based on the type of activities proposed, the location in which they took place and their duration. Table 5.1 includes interventions characteristics and the number of studies in which they were applied.

Table 5.1 *Interventions' characteristics*

Intervention		Number of studies	% of the total
Type			
	Reading biographies, essays, or articles	24	35%
	Watching a video	15	22%
	Listening to and interacting with keynote speakers	12	17%
	Mentoring program	10	15%
	Attending a workshop	6	9%
	Visit the job site guided by the worker	2	3%
	Seeing and listening to the experimenter	2	3%
	Playing a game with a virtual mentor	1	1%
	Seeing the image of a character in the exercise sheet	1	1%
Location			
	School	26	38%
	Laboratory	20	29%
	Online	13	19%
	Conference/workshop	8	12%
	Summer camp	1	1%
Duration			
	Less than 1 hour	12	17%
	1 to 3 hours	26	38%
	1 to 2 days	11	16%
	1 month or less	4	6%
	1-6 months	6	9%
	More than 6 months	9	13%

This classification identifies nine types of interventions, of which four were adopted in only one or two studies. Asking subjects to read something, e.g., a biography or an essay, was the most frequent option (35%). For instance, Betz and Sekaquaptewa (2012) asked participants to read magazine-type interviews of female university students. Hoffman and Kurtz-Costes (2018) provided participants with a biography of a scientist, describing how he/she first became interested in STEM, his/her training, and his/her civic service promoting STEM interest in youth. Finally, Bagès and Martinot (2011) asked participants to read a short text about an older student's success in maths.

Asking subjects to watch a video, whose content varied from study to study, was the second most frequent intervention (22%). More specifically, studies used interviews (e.g., Pietri, Johnson, et al., 2021), presentations (e.g., Baylor et al., 2006; Wessels, 1987), lessons on STEM contents showing scientists at work (e.g., Good et al., 2010), movies with a female leading character (Ziegler & Stoeger, 2008), footage of women and men interacting (LaCosse et al., 2016; Lewis et al., 2019; Van Loo & Rydell, 2014), and commercials (Davies et al., 2002; Lamers & Mason, 2018).

In 17% of studies, keynote speakers were invited to talk about their job and, usually, participants could ask questions and interact with them. For instance, in Buck et al. (2002), role models taught a class on science activities with magnetism and electricity and then talked about their job. Similarly, the effect of a workshop was investigated in 9% of studies, which differed from the former for the number of speakers and the less intimate type of interaction. Examples included Girls in STEM at Tulane (Merritt et al., 2021; O'Brien et al., 2017), and the Grad Cohort workshop (Stout et al., 2017).

Mentoring programmes (15%) exposed students to mentors rather than role models. Participants were typically required to engage for a long period (usually from six months to one year) in a one-to-one relationship with a chosen mentor. For instance, in the program 'MentorNet' (Single et al., 2005), college students were paired with professionals based on certain criteria, e.g., their interests. Mentors and protégées were advised to communicate via email at least twice a month for one academic year. Examples of discussion topics included the mentor's job, protégé's plans, college life, balancing career and family/life, managing time, and stress.

In two cases, role models guided participants around a site visit at their workplace while illustrating their job daily routine (Bamberger, 2014; Jethwani et al., 2017). In two other studies, the exposure to role models consisted of seeing and listening to an experimenter presenting a task in the laboratory session (Marx & Roman, 2002; Stout et al., 2011). Similarly, Morin-Messabel et al. (2017) exposed participants only to images of the role models, printed on the top of the task sheet. Finally, Cherchiglia (2019) proposed a virtual agent as a role model who then accompanied students while playing an online STEM game.

Other parameters varying according to the study were the location and the time spent with the exemplar. Activities took place either in school (38%), in a laboratory (29%), online (19%), at a conference or workshop site (12%), or at a summer camp (1%). The time spent with role models and mentors ranged from less than one hour to a year. Studies conducted in a laboratory generally lasted less than one hour, while talks with speakers from one to three hours could be repeated for a month or so. Workshops and conferences could last one or two days while mentoring programs had usually a longer length – from six to nine months.

5.5. Variables

The dependent variable on which the role models' effect was tested (here, called 'variable of interest') showed the greatest variation among studies. I tried to classify these variables by grouping together those measuring the same concept. This reduced the variation, but I still ended up with a long list of 41 variables, quite heterogeneous (see Table 5.2). I grouped them into four macro classes, i.e., target characteristics, target future, stereotypes and intervention.

The first collects all factors measuring participants' characteristics, i.e., opinions, feelings, attitudes, and behaviour. I decided to exclude from this macro class any attitude or behaviour referring to plans for the future, or stereotypes, which were grouped respectively, in the second and third macro classes. This distinction was informed by the 'expectancy-value theory' (Eccles, 1987), according to which, stereotypes indirectly affect career and educational-related choices by shaping certain individual characteristics, e.g., self-efficacy beliefs, interests, expectancies of success, and subjective values. Therefore, the first three macro classes distinguish the three components of this mediation model, whereas the last class collects opinions and attitudes toward the intervention itself.

Table 5.2 *Types of variables of interest*

Variable of interest	Number of studies	% of the total
Target's characteristic		
Self-concept in STEM	20	29%
Performance in STEM	19	28%
Self-efficacy in STEM	19	28%
Interest in STEM	18	26%
Opinion on the experience	10	15%

STEM enjoyment	10	15%
Self-esteem in STEM	5	7%
Self-assessed performance in STEM	5	7%
STEM utility	5	7%
Friends in STEM	2	3%
STEM activities	2	3%
Anxiety	1	1%
Evaluative certainty[1]	1	1%
Gender identification	1	1%
Impression-related concerns[2]	1	1%
STEM knowledge	1	1%
STEM task perceived difficulty	1	1%
Growth mindset	1	1%
Target's future		
STEM career intentions and aspirations	15	22%
STEM academic intentions and aspirations	12	18%
Academic choice	1	1%
Major retention	1	1%
Stereotypes		
Gender-science stereotypes endorsement	11	16%
Awareness of gender bias in STEM	5	7%
Attitudes toward women in science	5	7%
Stereotype threat	3	4%
Belief that computing skills are innate	1	1%
Masculinity bias awareness	1	1%
Messages about women in STEM read	1	1%
Role model/mentor		
Perceived similarity to role model	10	15%
Perceived image of scientists	6	9%
Perceived role model's likeability	4	6%
Perceived communal goal affordance	2	3%
Opinion on the role model/mentor	2	3%
Perceived role model's competence	2	3%
Empathy with the role model	1	1%
Having role models	1	1%

Perceived attainability of role model's competence	1	1%
Perceived role model's success	1	1%
Perceived role model's warmth	1	1%
Interest in having the role model as a professor/boss	1	1%

Note. ^{1} Belief of having performed poorly in a test

^{2} Concern about being successful and what people think of the individual

Regarding *target characteristics*, the role models and mentors' effect was tested on self-concept (29%), performance (28%), self-efficacy (28%), interest in STEM (26%), STEM enjoyment (15%), self-esteem (7%), self-assessed performance (7%), and utility attributed to STEM (7%). In a few studies, researchers measured the effect of role models on friendship networks with people in STEM (Ramsey et al., 2013; Walton et al., 2015), the engagement in STEM activities (Stoeger et al., 2013, 2016), and the knowledge of STEM concepts (Stoeger et al., 2013).

While related, these factors point to different mechanisms. Self-concept is defined as the collection of qualities (i.e., attributes and roles) that an individual attributes to himself/herself (Kinch, 1963). Self-efficacy refers to a 'person's belief in their ability to succeed in a particular situation' (Bandura, 1977). Finally, self-esteem is defined as 'one's positive or negative attitude toward oneself and one's evaluation of one's own thoughts and feelings overall about oneself' (Rosenberg, 1965). While these concepts point to various characteristics, they also share certain commonalities, including the fact that they reflect general 'attitudes' toward STEM. Contrary to the others summarized in the table, variables in this category were frequently assessed together, especially performance, interest in STEM, self-concept, STEM enjoyment, and self-efficacy.

Finally, some dependent variables answered specific research questions not directly linked to STEM or role models, yet related to gender or other task-specific factors. These included feelings of anxiety (Good et al., 2010), beliefs about the proposed task (Morin-Messabel et al., 2017; Thiem, 2016), gender identification (Walton et al., 2015), the concern about being successful and people's opinions (Marx et al., 2005), and growth mindset (Cherchiglia, 2019).

Regarding the *target future*, studies measured the interventions' effect on aspirations, intentions and possibly, future choices. In most cases, participants were asked to report their intentions and aspirations for a future professional career (22%) and a prospective academic path (18%). For instance, Breda et al. (2020) were able to obtain information on the concrete choice of higher education curricula. Using administrative data, they tested the role models' effect on the probability of being observed in a science-related track one year after the intervention. Similarly, Dennehy and Dasgupta (2017) collected data one year after the intervention, to examine college students' tendency to abandon the STEM major. They measured both intention to and actual choice of switching majors for female college students enrolled in an engineering major.

Regarding *stereotypes and gender bias*, studies tested role models' effect on the endorsement of gender-science stereotypes (16%), awareness of the existence of a gender bias in STEM (7%), and attitudes toward women in science (7%). 4% of studies tested the effect on stereotype threat, i.e., 'a concern that one might inadvertently confirm an unwanted belief about one's group' (Inzlicht & Schmader, 2011). Van Loo and Ridell (2014) provided an interesting example in the context of gender and STEM by asking their female participants whether they agreed on statements relating to their mathematical ability and gender after watching a brief video.

Stout et al. (2017) asked participants their opinion on the belief that computer skills are innate, from which gender-stereotypical beliefs typically arise. In a study on male-only role models, Pietri et al. (2021) measured participants' awareness that men would experience negative reactions when showing typically feminine features. Finally, Ramsey et al. (2013) asked participants the number of occasions in which they read messages/articles about women employed in a STEM sector.

In terms of *interventions*, studies examined especially participants' perceived similarity to the model (15%) and in general, their perceptions of role models' characteristics. This was informed by previous research on the mediating effect of these factors in the role modelling process (Asgari et al., 2012). As stated in the 'motivational role model theory' (Morgenroth et al., 2015), to be effective, role models should be perceived as being similar and desirable and their achievements as attainable.

Besides similarity, other role model characteristics were generic likeability (Betz & Sekaquaptewa, 2012; Lawner, 2014; Marx et al., 2013; Pietri, Johnson, et al., 2021), the affordance of communal goals in their profession (Clark et al., 2016; Lawner, 2014), inspired empathy (Pietri, Drawbaugh, et al., 2021), competence (Marx & Ko, 2012; Pietri, Johnson, et al., 2021) and its attainability (Marx & Ko, 2012), success (Lawner, 2014), and warmth (Pietri, Johnson, et al., 2021).

In more qualitative studies, participants were usually asked their opinion on either the proposed experience, the exemplars they encountered, or more generically scientists. Results usually reported a summary of the most frequently mentioned topics. In Jethwani et al. (2017), the main emerging theme was how the experience challenged girls' preconceptions of the computer science environment. Scott (2013) conducted a series of interviews with participants, seeking information on opportunities and constraints participants perceived during the program.

Considering opinions on mentors and role models, Rule et al. (2019) were interested in gathering specific information on why participants appreciated the presented professional. In Cherchiglia (2019), participants could customize the virtual mentor accompanying them during a STEM online game. Questions in the semi-structured interviews aimed at verifying whether participants recalled their virtual mentor and why they picked him/her with those particular features from those proposed.

Finally, opinions about scientists verified whether the exposure challenged stereotypical ideas of scientists. However, these studies varied in terms of the instruments used. Bennett et al. (1998) included an open-ended question. Howard (2015) asked participants to perform the Draw-A-Scientist test, which asks children to draw a scientist, letting them choose the gender, age, physical appearance and clothing of the drawn person. In other cases, a Likert-type scale was used, e.g., the Image of Scientists Scale (Granville, 1985).

5.6. Effects

The heterogeneity of populations, interventions, methods and variables of interest described previously, prevent from calculating an average effect and so performing a meta-analysis of the effects of these interventions. Nevertheless, I created a structured synthesis classifying findings by aim. The identified categories were grouped in two macro classes, i.e., those referring to the intervention and those to the exemplar type (see Table 5.3). Although some studies also included male participants, here we focus on the effect on girls. Note that here only studies using an experimental research design were included.

Table 5.3 *Effects*

Aim and type of effect	Number of studies	% of the total
Intervention		
Role model's presence (presence wrt absence)		
Positive	14	29%
Mixed	13	27%
Absent	2	4%
Reflection on role model (reflection wrt no reflection)		
Positive	3	6%
Seeing the role model (seeing wrt hearing/reading)		
Positive	1	2%
Mixed	1	2%
Feedback from role model (positive wrt negative)		
Positive	1	2%
Information on role model (job and private life wrt job only)		
Positive	1	2%
Role model		
Role model's gender (female wrt male)		
Positive	8	16%
Mixed	7	14%
Negative	1	2%
Absent	6	12%
Role model stereo-typicality (high wrt low/neutral)		
Positive	1	2%
Mixed	3	6%
Negative	5	10%
Role model's competence (high wrt low)		
Positive	2	4%
Mixed	1	2%
Similarity to role model (similar wrt non-similar)		
Positive	1	2%
Mixed	1	2%
Role model's success (hardworking wrt gifted)		

Positive	2	4%
Role model’s self-confidence (doubtful vs not doubtful)		
Mixed	1	2%

Most studies tested the effect of being exposed to a role model compared to not being exposed to a role model or rather being exposed to a non-STEM role model. Interestingly, in Bamberger (2014), after meeting with female scientists and visiting a high-tech company, fewer students perceived a woman scientist positively and therefore chose a STEM major. Furthermore, they also expressed less capability of dealing with STEM in the future and believed they could not deal with these courses.

Besides this main line of research, only a few studies tested variations in the way interventions were implemented. In three of these, participants were asked to reflect on role models after exposure. O’Brien et al. (2017) asked half of the sample to identify their favourite professional among those presented and to motivate their choice. They then compared them to the other half of the sample, who were asked to identify and write about their best friend. Van Camp et al. (2019) asked all participants to write a brief note after exposure, whose content varied depending on the treatment or control groups. One group was asked to write about how they identified with the role model, another one to write about the details of the woman’s biography and, finally, those in the control group wrote about the woman’s hobbies. Similarly, Gilbert (2015) required students in the treatment group to write about their reasons for identifying with the role model, while the other students summarized what they had read. In all three studies, results were positive, meaning that reflecting on identification with the role model was more efficient than reflecting on other aspects.

In two studies, researchers tested the difference between physically seeing the role models and only hearing their voices (Baylor et al., 2006) or reading about them (Pietri, Johnson, et al., 2021). In both studies, results showed that seeing them while they were talking was more effective. Pietri et al. (2021) found that students who saw an interview with a female computer scientist indicated a higher interest in computer science and perceived the scientist as warmer and more competent. This, in turn, strengthened their identification with STEM, compared to those who read the written transcripts of the interview. Baylor et al. (2006) compared a situation in which participants could see role models from the engineering sector with one in which they could only hear their voices. Results showed that in the former case, participants were significantly more likely to consider engineering as useful and reported higher self-efficacy, interest in engineering-related fields and STEM engagement.

LaCosse et al. (2016) tested the effect of receiving positive or negative feedback from role models. They found that negative feedback had detrimental effects on both intentions to continue in STEM after graduation and stereotypical attribution bias. Finally, Wessels (1987) tested the effect of receiving information on both the job and the private life of the role model compared to receiving information on the job alone. Girls assigned to the former condition held more positive opinions of women working in STEM, while those assigned to the job-only condition strengthened negative stereotyped images of women employed in STEM.

The second macro class collects studies in which role models' characteristics, especially gender, were manipulated. In eight out of twenty-two studies, female role models were more effective than male role models, in six studies there were no differences between male and female role models, whereas in seven studies the results were mixed. For instance, Dennehy and Dasgupta (2017) assigned female students majoring in engineering to either a male or a female mentor. At the end of the academic year, they compared their attitudes toward STEM and performance (grades) with that of a group not assigned to a mentor. Findings were inconclusive in determining whether it was better to assign students to a male or a female mentor. On the one hand, those assigned to a woman exhibited higher self-concept and self-efficacy, less stereotype threat, and a stronger intention to pursue advanced degrees in engineering compared to both those under the control condition and those assigned to a male mentor. On the other hand, female students assigned to men performed better than students in both the female mentor and control conditions.

In some studies, role models varied in the level of stereo-typicality (high compared to low or neutral) endorsed. For instance, Cheryan et al. (2011) manipulated stereo-typicality through clothing, hobbies, and preferences, while Lawner (2014) by reporting examples of the role model's communal behaviour (e.g., helping others). Van Loo and Rydell (2014) and Lewis et al. (2019) suggested stereo-typicality through dominance in the interaction between a man and a woman, i.e., a situation in which the men dominated the conversation compared to one in which the woman dominated the conversation. Other studies highlighted female role models' femininity. Surprisingly, the results were heterogeneous. For instance, Howard (2015) found that women's interest in physics increased after exposure to a feminine female role model such that they had equivalent interest compared to men. On the contrary, Betz and Sekaquaptewa

(2012) found that, compared to neutral models, feminine role models made middle school girls feel less capable and interested in maths.

Other manipulated role model characteristics were the level of competence and self-confidence, similarity to the participant and reason for success, i.e., being hardworking compared to gifted. In most studies, competent role models had a more positive influence on female participants than low competent ones, e.g., by increasing their performance (Marx et al., 2013; Marx & Roman, 2002), and self-efficacy (Thiem, 2016). As regards similarity to the role model, the hypothesis that more similarity would influence the target more obtained discording results in the two studies which tested it. Merritt et al. (2021) did not find any difference between girls asked to give feedback on their favourite role model and those who gave feedback on a randomly chosen role model. Marx and Ko (2012) on the other hand, found that participants' performance improved the closer the similarity to the role model, but that this similarity was irrelevant when the role model was highly competent. When presented with a hardworking role model, girls performed as well as boys on a math test, while they performed less well than boys when exposed to either a gifted role model or to the control condition (Bagès et al., 2016). Finally, Marx et al. (2013) found that female role models expressing doubt about their achievements undermined women's math performance. However, they found no difference between female participants exposed to either a doubtful or a typical male role model.

6. Discussion and Conclusions

Exposure to (counterstereotypical) role models is a widely diffused strategy used to enhance women's empowerment in the STEM sector with initiatives ranging from online-based resources (e.g., 'San Diego STEM Role Model Initiative', 'Techbridge Girls',) to mentoring programs (e.g., 'ACE Mentor Program', 'MentorNet') and school projects (e.g., Breda et al., 2020). The need for higher representation of women in male-dominated sectors has recently increased even beyond these examples. For instance, producers of toys, books or movies are trying to balance the representation of male and female characters, especially in STEM (e.g., Ignotofsky, 2016).

In this context, there is a need for academic research to inform these initiatives, assess the efficacy of various options and provide guidelines for their design and the implementation of various strategies. By applying a systematic approach to studies' selection, I have tried to summarize research on the effect of role models and mentors, by focusing on research type, targeted population, type of exposure, type of interventions, variables of interest and effects.

Findings showed significant heterogeneity in the type of exemplars, interventions, variable of interests and effects, while there was more similarity in research type and target population. The youngest students exposed to role models or mentors were children aged 9 years old, while there were no studies involving pre-schoolers. This is a pity, as gender stereotypes are said to develop by about two years of age (Martin & Ruble, 2010). Furthermore, while most studies investigated the effect of role models, only a minority concentrated on mentors. This is not surprising, as mentorship programmes require a long-term, one-to-one relationship, which makes these programmes' implementation more challenging. Finally, research varied in terms of types of interventions, including the type of action required by participants (i.e., reading, listening, seeing), the possibility to interact with the role model, the exposure's duration and participants' involvement in the activities.

As regards the target population's age, even if testing the interventions' effects on such young participants may be challenging, instruments such as the Draw-A-Scientist test (Farland-Smith, 2017) can be used even with children who have not developed verbal and written abilities. Future research could shed light on the efficacy of these interventions targeted at a younger population. On the one hand, changing attitudes toward STEM is easier when children have not yet fully developed their inclinations and interests. This is not the case for adolescents, the preferred target of research, who have probably already developed their own academic and occupational preferences (Gurres et al., 2021). On the other hand, children's not-fully-developed cognitive skills impose constraints on this type of intervention. As found in a study conducted in elementary schools, female scientists presented to children in the class were not perceived as scientists, but as teachers (Buck et al., 2002). Further research on this is crucial to understand whether role models are more effective at a younger stage of students' academic paths.

As regards mentors and role models, ascertaining whether these two figures have equivalent beneficial effects, or whether a more intimate relationship is key to developing a stronger link with STEM, would require more systematic and comparative research. In a study on physicians, Taylor et al. (2009) found that participants valued shorter and more focused mentoring relationships compared to more traditional, longitudinal experiences. Considering the possible implementation of these strategies on a large scale, initiatives involving role models should be preferred to more demanding mentorship programs because of their economies of scale, feasibility and replicability.

Finally, as regards heterogeneity in interventions, given that a role model is such if it inspires someone, anyone can potentially serve as a role model for someone else, even unwittingly. Thus, we could consider that heterogeneity and diversity are beneficial, even necessary. However, apart from more traditional means of exposure to role models, such as talks and articles, other more up-to-date and attractive strategies for young people should be investigated. It is worth mentioning here the idea of including a female virtual agent in an online game on STEM (Cherchiglia, 2019) or watching a movie with a female leading character (Ziegler & Stoeger, 2008). A study on women's representation in media content in STEM revealed that female characters account for 37% of the total (Geena Davis Institute, 2018). Further investigation is required to understand how this is influencing girls and whether a more equal representation of female and male characters in movies and TV shows could help to stimulate women's interest in studying STEM. Furthermore, studies inspired by real-life initiatives with children should also examine whether toys, such as dolls and action figures wearing scientists' clothing, could serve as role models for children.

Furthermore, this heterogeneity in the type of experimental intervention comes at a price: i.e., it is difficult to focus on a general understanding of the effectiveness of exposure to role models. Olsson and Martiny (2018) tried to circumvent this problem by assessing the efficacy of some of these interventions by targeting different age groups separately. Although a good starting point, a more systematic analysis of interventions' outcomes is required to assess the efficacy of these interventions and improve cumulativeness and replication of findings. The study's outcomes suggest that girls and women generally benefit from exposure to STEM role models. However, exceptions do exist, which in the worst-case scenario, could even have detrimental effects on female engagement in STEM. Further research should shed light on the causes underlying these negative and null effects.

This review is also subject to the limitations of any review of this kind, i.e., relevant sources of information may have been omitted, which is dependent on the availability of information. Given studies heterogeneity, I would recommend that researchers chose dependent variables that can be fully compared with previous research. This would help to create a systematic summary of results and provide a general picture of the effectiveness of role models' interventions for women in STEM. Without this coherence, it is impossible to make concrete conclusions on the beneficial effects of role models or identify the interventions' features, which, on the contrary, may have detrimental effects on women in this sector.

References

Alam, A., & Sanchez Tapia, I. (2020). *Mapping gender equality in STEM from school to work*. UNICEF Office of Global Insight and Policy.

Asgari, S., Dasgupta, N., & Stout, J. G. (2012). When do counterstereotypic ingroup members inspire versus deflate? The effect of successful professional women on young women's leadership self-concept. *Personality & Social Psychology Bulletin*, *38*(3), 370–383. https://doi.org/10.1177/0146167211431968

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*(1), 55–64. https://doi.org/10.1177/0361684315608842

Bamberger, Y. M. (2014). Encouraging Girls into Science and Technology with Feminine Role Model: Does This Work? *Journal of Science Education and Technology*, *23*(4), 549–561. https://doi.org/10.1007/s10956-014-9487-7

Bandura, A. (1977). *Social learning theory.* (pp. viii, 247). Prentice-Hall.

Baylor, A. L., Rosenberg-Kima, R. B., & Plant, E. A. (2006). *Interface agents as social models: The impact of appearance on females' attitude toward engineering*. 526. http://dl.acm.org/citation.cfm?doid=1125451.1125564

Bennett, D., Tsikalas, K., Hupert, N., Meade, T., & Honey, M. (1998). *The benefits of online mentoring for high school girls: Telementoring young women in science, engineering and computing project—Year 3 Evaluation*. Center for children and technology. http://cct.edc.org/sites/cct.edc.org/files/publications/telement_bomhsg98.pdf

Betz, D. E., & Sekaquaptewa, D. (2012). My fair physicist? Feminine math and science role models demotivate young girls. *Social Psychological and Personality Science*, *3*(6), 738–746. https://doi.org/10.1177/1948550612440735

Breda, T., Grenet, J., Monnet, M., & Effenterre, C. V. (2020). *Do Female Role Models Reduce the Gender Gap in Science? Evidence from French High Schools* (IZA Discussion Papers No. 13163). Institute of Labor Economics (IZA). http://hdl.handle.net/10419/216475

Buck, G. A., Leslie-Pelecky, D., & Kirby, S. K. (2002). Bringing Female Scientists into the Elementary Classroom: Confronting the Strength of Elementary Students' Stereotypical Images of Scientists. *Journal of Elementary Science Education*, *14*(2), 1–9.

Cherchiglia, L. L. (2019). The Creation of STEM Role Models: An Exploratory Study on the Design of Mentoring Characters for a STEM Gaming Website [Ph.D., Michigan State University]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/creation-stem-role-models-exploratory-study-on/docview/2300630345/se-2?accountid=12459

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*(6), 656–664. https://doi.org/10.1177/1948550611405218

Clark, E. K., Fuesting, M. A., & Diekman, A. B. (2016). Enhancing interest in science: Exemplars as cues to communal affordances of science. *Journal of Applied Social Psychology*, *46*(11), 641–654. https://doi.org/10.1111/jasp.12392

Dasgupta, N., & Greenwald, A. G. (2001). On the Malleability of Automatic Attitudes: Combating Automatic Prejudice With Images of Admired and Disliked Individuals. *Journal of Personality and Social Psychology*, *81*(5).

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming Images: How Television Commercials that Elicit Stereotype Threat Can Restrain Women Academically and Professionally. *Personality and Social Psychology Bulletin*, *28*(12), 1615–1628. https://doi.org/10.1177/014616702237644

Dennehy, T. C., & Dasgupta, N. (2017). Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences*, *114*(23), 5964–5969. https://doi.org/10.1073/pnas.1613117114

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*(4), 735–754. https://doi.org/10.1037/0022-3514.46.4.735

Eccles, J. S. (1987). Gender Roles and Women's Achievement-Related Decisions. *Psychology of Women Quarterly*, *11*(2), 135–172. https://doi.org/10.1111/j.1471-6402.1987.tb00781.x

Evans, M. A., Whigham, M., & Wang, M. C. (1995). The effect of a role model project upon the attitudes of ninth-grade science students. *Journal of Research in Science Teaching*, *32*(2), 195–204. https://doi.org/10.1002/tea.3660320208

Farland-Smith, D. (2017). The Evolution of the Analysis of the Draw-a-Scientist Test. In P. Katz (Ed.), *Drawing for Science Education* (pp. 171–178). SensePublishers. https://doi.org/10.1007/978-94-6300-875-4_15

Geena Davis Institute. (2018). *Portray Her: Representation of Women STEM Characters in Media*. Geena Davis Institute on Gender in Media. https://seejane.org/wp-content/uploads/portray-her-full-report.pdf

Gilbert, P. N. (2015). *The role of role models: How does identification with STEM role models impact women's implicit STEM stereotypes and STEM outcomes?* [Ph.D. Dissertation, Tulane University]. https://digitallibrary.tulane.edu/islandora/object/tulane%3A27945

Good, J. J., Woodzicka, J. A., & Wingfield, L. C. (2010). The Effects of Gender Stereotypic and Counter-Stereotypic Textbook Images on Science Performance. *The Journal of Social Psychology*, *150*(2), 132–147. https://doi.org/10.1080/00224540903366552

Granville, R. C. (1985). The influence of science role models on the attitudes of middle school students toward women in science [Educat.D., University of Miami]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/influence-science-role-models-on-attitudes-middle/docview/303419629/se-2?accountid=12459

Gurres, S., Dillmann, K.-U., Reith, W., & Krick, C. M. (2021). The Individual Inclination to an Occupation and its Neuronal Correlate. *Frontiers in Education*, *6*, 57. https://doi.org/10.3389/feduc.2021.633962

Hoffman, A. J., & Kurtz-Costes, B. (2018). Promoting science motivation in American Indian middle school students: An intervention. *Educational Psychology*, *39*(4), 448–469.

Howard, C. (2015). Do STEM fields need a makeover?: The effect of role model femininity on men and women's interest in STEM [M.A., San Diego State University]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/do-stem-fields-need-makeover-effect-role-model/docview/1726897936/se-2?accountid=12459

Ignotofsky, R. (2016). *Women in science: 50 fearless pioneers who changed the world* (First edition). Ten Speed Press.

Inzlicht, M., & Schmader, T. (2011). *Stereotype Threat Theory, Process, and Application*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199732449.001.0001

Jethwani, M. M., Memon, N., Seo, W., & Richer, A. (2017). "I Can Actually Be a Super Sleuth": Promising Practices for Engaging Adolescent Girls in Cybersecurity Education. *Journal of Educational Computing Research*, *55*(1), 3–25. https://doi.org/10.1177/0735633116651971

Kinch, J. W. (1963). A Formalized Theory of the Self-Concept. *American Journal of Sociology*, *68*(4), 481–486. https://doi.org/10.1086/223404

LaCosse, J., Sekaquaptewa, D., & Bennett, J. (2016). STEM Stereotypic Attribution Bias Among Women in an Unwelcoming Science Setting. *Psychology of Women Quarterly*, *40*(3), 378–397. https://doi.org/10.1177/0361684316630965

Lamers, D., & Mason, R. (2018). Advertising CS/IT degrees to female students in Australia. *Proceedings of the 20th Australasian Computing Education Conference*, 1–8. https://doi.org/10.1145/3160489.3160497

Lawner, E. K. (2014). *Impact of Role Model Gender and Communality on College Women's Math Performance and Interest in STEM* [University of Connecticut]. https://opencommons.uconn.edu/gs_theses/688

Lawner, E. K., Quinn, D. M., Camacho, G., Johnson, B. T., & Pan-Weisz, B. (2019). Ingroup role models and underrepresented students' performance and interest in STEM: A meta-analysis of lab and field studies. *Social Psychology of Education*. https://doi.org/10.1007/s11218-019-09518-1

Lenton, A. P., Bruder, M., & Sedikides, C. (2009). A Meta-Analysis on the Malleability of Automatic Gender Stereotypes. *Psychology of Women Quarterly*, *33*(2), 183–196. https://doi.org/10.1111/j.1471-6402.2009.01488.x

Lewis, N. A., Sekaquaptewa, D., & Meadows, L. A. (2019). Modeling gender counter-stereotypic group behavior: A brief video intervention reduces participation gender gaps on STEM teams. *Social Psychology of Education*. https://doi.org/10.1007/s11218-019-09489-3

Martin, C. L., & Ruble, D. N. (2010). Patterns of Gender Development. *Annual Review of Psychology*, *61*(1), 353–381. https://doi.org/10.1146/annurev.psych.093008.100511

Marx, D. M., & Ko, S. J. (2012). Superstars "like" me: The effect of role model similarity on performance under threat: Superstars "like" me. *European Journal of Social Psychology*, *42*(7), 807–812. https://doi.org/10.1002/ejsp.1907

Marx, D. M., Monroe, A. H., Cole, C. E., & Gilbert, P. N. (2013). No Doubt About It: When Doubtful Role Models Undermine Men's and Women's Math Performance Under Threat. *The Journal of Social Psychology*, *153*(5), 542–559. https://doi.org/10.1080/00224545.2013.778811

Marx, D. M., & Roman, J. S. (2002). Female Role Models: Protecting Women's Math Test Performance. *Personality and Social Psychology Bulletin*, *28*(9), 1183–1193. https://doi.org/10.1177/01461672022812004

Marx, D. M., Stapel, D. A., & Muller, D. (2005). We Can Do It: The Interplay of Construal Orientation and Social Comparisons Under Threat. *Journal of Personality and Social Psychology*, *88*(3), 432–446. https://doi.org/10.1037/0022-3514.88.3.432

Merritt, S. K., Hitti, A., Van Camp, A. R., Shaffer, E., Sanchez, M. H., & O'Brien, L. T. (2021). Maximizing the impact of exposure to scientific role models: Testing an intervention to increase science identity among adolescent girls. *Journal of Applied Social Psychology*, jasp.12774. https://doi.org/10.1111/jasp.12774

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*(4), 465–483. https://doi.org/10.1037/gpr0000059

Morin-Messabel, C., Ferrière, S., Martinez, F., Devif, J., & Reeb, L. (2017). Counter-stereotypes and images: An exploratory research and some questions. *Social Psychology of Education*, *20*(1), 1–13. https://doi.org/10.1007/s11218-017-9370-5

Nickerson, S., Bjorkman Sei, K., Ko, J., & Marx, D. M. (2017). Identification Matters: Effects of Female Peer Role Models Differ By Gender Between High and Low Mathematically Identified Students. *Proceedings of the 20th Annual Conference on Research in Undergraduate Mathematics Education*, 6. https://www.researchgate.net/publication/333489183_Identification_Matters_Effects_of_Female_Peer_Role_Models_Differ_By_Gender_Between_High_and_Low_Mathematically_Identified_Students

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*(3), 301–309. https://doi.org/10.1177/1948550616671997

Olsson, M., & Martiny, S. E. (2018). Does Exposure to Counterstereotypical Role Models Influence Girls' and Women's Gender Stereotypes and Career Choices? A Review of Social Psychological Research. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02264

Peters, M., Godfrey, C., McInerney, P., Munn, Z., Trico, A., & Khalil, H. (2020). Chapter 11: Scoping Reviews. In E. Aromataris & Z. Munn (Eds.), *JBI Manual for Evidence Synthesis*. JBI. https://doi.org/10.46658/JBIMES-20-12

Pietri, E. S., Drawbaugh, M. L., Johnson, I. R., & Colvin, V. E. (2021). Maybe He Is Relatable Too: Encouraging Women to Identify With Male Scientists by Highlighting Bias Against Fathers. *Psychology of Women Quarterly*, *45*(1), 79–105. https://doi.org/10.1177/0361684320972118

Pietri, E. S., Johnson, I. R., Majid, S., & Chu, C. (2021). Seeing What's Possible: Videos are more Effective than Written Portrayals for Enhancing the Relatability of Scientists and

Promoting Black Female Students' Interest in STEM. *Sex Roles*, *84*(1–2), 14–33. https://doi.org/10.1007/s11199-020-01153-x

Prieto-Rodriguez, E., Sincock, K., & Blackmore, K. (2020). STEM initiatives matter: Results from a systematic review of secondary school interventions for girls. *International Journal of Science Education*, *42*(7), 1144–1161. https://doi.org/10.1080/09500693.2020.1749909

Ramsey, L. R., Betz, D. E., & Sekaquaptewa, D. (2013). The effects of an academic environment intervention on science identification among women in STEM. *Social Psychology of Education*, *16*(3), 377–397. https://doi.org/10.1007/s11218-013-9218-6

Rosenberg, M. (1965). Rosenberg self-esteem scale (RSE). *Acceptance and Commitment Therapy. Measures Package*, *61*(52), 18.

Rule, A. C., Blaine, D. A., Edwards, C. M., & Gordon, M. (2019). Girls Explore Math Careers by Making 3D-Constructions of Diverse Women Mathematicians' Lives. *The Mathematics Enthusiast*, *16*(1). https://scholarworks.umt.edu/tme/vol16/iss1/27

Scott, J. R. (2013). *Einstein girls: Exploring STEM careers, interest, and identity in an online mentoring community* [Ph.D. Dissertation, University of Florida]. https://ui.adsabs.harvard.edu/abs/2013PhDT.......600S/abstract

Single, P. B., Muller, C. B., Cunningham, C. M., Single, R. M., & Carlsen, W. S. (2005). Mentornet: E-mentoring For Women Students In Engineering And Science. *Journal of Women and Minorities in Science and Engineering*, *11*(3), 295–310. https://doi.org/10.1615/JWomenMinorScienEng.v11.i3.60

Stoeger, H., Duan, X., Schirner, S., Greindl, T., & Ziegler, A. (2013). The effectiveness of a one-year online mentoring program for girls in STEM. *Computers & Education*, *69*, 408–418. https://doi.org/10.1016/j.compedu.2013.07.032

Stoeger, H., Schirner, S., Laemmle, L., Obergriesser, S., Heilemann, M., & Ziegler, A. (2016). A contextual perspective on talented female participants and their development in extracurricular STEM programs: Talented females in an extracurricular STEM program. *Annals of the New York Academy of Sciences*, *1377*(1), 53–66. https://doi.org/10.1111/nyas.13116

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*(2), 255–270. https://doi.org/10.1037/a0021385

Stout, J. G., Tamer, B., Wright, H. M., Clarke, L. A., Dwarkadas, S., & Howard, A. M. (2017). The Grad Cohort Workshop: Evaluating an Intervention to Retain Women Graduate Students in Computing. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.02071

Taylor, C. A., Taylor, J. C., & Stoller, J. K. (2009). The Influence of Mentorship and Role Modeling on Developing Physician–Leaders: Views of Aspiring and Established Physician–Leaders. *Journal of General Internal Medicine*, *24*(10), 1130–1134. https://doi.org/10.1007/s11606-009-1091-9

Thiem, K. (2016). Gender stereotypes and academic performance: The influence of salient role models on stereotype validation [Ph.D., The University of Iowa]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/gender-stereotypes-academic-performance-influence/docview/1878893949/se-2?accountid=12459

UNESCO. (2017). *Cracking the code girls' and women's education in science, technology, engineering and mathematics (STEM).* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000253479

U.S. Bureau of Labor Statistics. (2021). *Employment in STEM occupations*. https://www.bls.gov/emp/tables/stem-employment.htm

Van Camp, A. R., Gilbert, P. N., & O'Brien, L. T. (2019). Testing the effects of a role model intervention on women's STEM outcomes. *Social Psychology of Education*, *22*(3), 649–671. https://doi.org/10.1007/s11218-019-09498-2

Van Loo, K. J., & Rydell, R. J. (2014). Negative Exposure: Watching Another Woman Subjected to Dominant Male Behavior During a Math Interaction Can Induce Stereotype Threat. *Social Psychological and Personality Science*, *5*(5), 601–607. https://doi.org/10.1177/1948550613511501

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology*, *107*(2), 468–485. https://doi.org/10.1037/a0037461

Wessels, J. S. (1987). Effects Of Two Audiovisual Presentations On Motivating Ninth Grade Girls To Pursue Scientific, Technical Careers [Ph.D., University of Denver]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/effects-two-audiovisual-presentations-on/docview/303473817/se-2?accountid=12459

Ziegler, A., & Stoeger, H. (2008). Effects of Role Models from Films on Short-Term Ratings of Intent, Interest, and Self-Assessment of Ability by High School Youth: A Study of Gender-Stereotyped Academic Subjects. *Psychological Reports*, *102*(2), 509–531. https://doi.org/10.2466/pr0.102.2.509-531

Chapter 6. ***Exposing students to role models: Effects of an online experiment on gender-science stereotypes and the perceived gender gap in STEM and humanities***

1. Introduction

Research has long emphasized the importance of role modelling and the impact that exemplars may have at different stages of life, i.e., from childhood (Coto et al., 2019) to career development (Gibson, 2003; Scherer et al., 1989). Eventually, humans' ability to imitate and acquire information from others is 'the reason of our success': 'We are adaptive learners who, even as infants, carefully select when, what, and from whom to learn' (Henrich, 2016, p. 4). More specifically, role models exert their influence on goals and motivation by acting as behavioural models, representing viable paths and being inspirational (Morgenroth et al., 2015).

Role models play a pivotal role, especially for underrepresented minorities, e.g., women in male-dominated sectors (Murrell et al., 1999), who dramatically lack examples to aspire to. According to the 'social role theory' (Eagly & Wood, 2012), the division of labour between women and men determines the attribution of characteristics and roles based on gender. This attribution is internalized, eventually resulting in the endorsement of gender stereotypes, which affect choices and behaviours. The persistent gender gap in the STEM sector would, thus, derive from the observation of the underrepresentation of women engaged in STEM careers, i.e., since STEM professionals are more frequently male, people associate the STEM sector with being a man.

According to this theory, the provision of examples of women employed in male-dominated sectors could change the perception that those jobs are not suitable for women, eventually resulting in a decrease in the gender gap. The study conducted by Miller et al. (2015) on the correlation between women's participation in STEM and gender-science stereotypes in 66 nations confirmed this relationship. Countries with higher female participation in tertiary education in science were also characterized by weaker implicit and explicit gender-science stereotypes.

Initiatives and projects involving female role models have gained momentum, especially in the STEM sector. However, as reported in Chapter 5, research on the effect of role models on women in STEM did not come to univocal findings (Olsson & Martiny, 2018). As regards the effect on gender-science stereotypes, Van Camp et al. (2019) found that role models influenced the endorsement of explicit gender stereotypes and the strength of automatic women-science association only when participants were required to reflect and write about their experiences, whereas passive exposure to role models was ineffective. Despite numerous contributions, to my knowledge, the perception of the representativeness of women in the STEM sector has been investigated only in one study. After a role model intervention held in class for Grade 10 and Grade 12 French students, Breda et al. (2020) asked participants to what extent they agreed with the statement 'There are more men than women in science-related jobs'. Results revealed that the intervention increased both male and female students' awareness of the underrepresentation of women in STEM.

This study aimed to fill this gap by testing whether brief exposure to female and male professionals could affect not only explicit and implicit gender stereotypes, but also the perceived representativeness of women and men in sectors traditionally considered as being either masculine (STEM studies) or feminine (humanistic studies). This difference was relevant given that the mechanism identified by the 'social role theory' is the rationale for the effectiveness of role models' interventions. Unlike previous studies on role models, I did not focus only on female exemplars and STEM disciplines, but I included also male role models and humanistic studies. The gender gap in humanistic studies is usually neglected and, unlike that in STEM studies, not perceived as a problem. However, as stressed before in this thesis, we should consider the association between gender and STEM and humanities as complementary stereotypes.

2. Literature review

There is a large consensus on the pivotal role played by gender stereotypes in explaining the existence and persistence of the gender gap in STEM (Nosek et al., 2009). Empirical evidence suggests that female children aged six automatically associate males with mathematics (Tomasetto et al., 2012), albeit they are not explicitly aware of the stereotype until they are 8-9 years old (Andre et al., 1999). From elementary school to entry into the labour market, gender-science stereotypes shape girls' experience in STEM and drive their progressive abandonment of the field. A survey conducted by Microsoft on 6,000 girls (Microsoft Philantropies, 2018) revealed that the percentage of those perceiving as 'not for them' jobs requiring coding and programming increased from 31% in middle school to 40% in high school. In college, 58% of young women counted themselves out of these jobs.

The 'social role theory' proposed by Eagly and Wood (2012) is a cornerstone of this type of research. According to this theory, differences and similarities between men's and women's behaviour originate from the endorsement of gender stereotypes, which in turn stem from the observation of women and men's different social roles in society. In the context of STEM, the underrepresentation of women in STEM-related jobs would favour the belief that the STEM sector is 'a male thing'. The endorsement of this belief, in turn, determines women's abandonment of this sector. The traditional task specialization has produced a gendered division of labour that does not reflect modern societies anymore. However, gender role beliefs are difficult to eradicate. This is because gender roles 'seem to reflect innate attributes of the sexes' (Eagly & Wood, 2012, p. 459), thus appearing natural and inevitable. However, since gender stereotypes would derive from the observed division of labour, a change in the representation of women and men in gender-segregated occupations should drive also a change in gender stereotypes.

The hypothesis that stereotypes are not stable but would change over time depending on certain circumstances has been investigated for years. In her review of evidence on the malleability of automatic stereotypes, Blair (2002) argued that implicit stereotypes can be moderated by a wide variety of events, among which the exposure to counterstereotypical events and group members. In line with these theories, one of the most suggested and adopted strategies to increase female participation in STEM is to show girls and young women professionals or students of the same gender engaged in this sector. This is to provide a different picture of women's occupations in contemporary society or, using the words of Blair, certain positive counterstereotypical group members. These interventions are frequently sponsored and strongly encouraged by governments and international organisations. In this line, numerous initiatives have been recently implemented, including experimental studies that tried to assess the effectiveness of this strategy on female – sometimes also male – participants. However, as shown in Chapter 5, the type of intervention, the outcomes of interest and the instruments used to measure these effects varied considerably on a case-by-case basis.

Focusing on gender stereotypes, researchers have tested whether exposure to role models could change implicit and/or explicit gender-science stereotypes. Stout et al. (2011) briefly exposed undergraduate women to either male or female peer experts, but the intervention had no impact on the implicit association between maths and gender. On the contrary, a study conducted on French high-school students (Breda et al., 2020) found that those who met a female scientist in class were less likely to endorse explicit gender stereotypes on abilities compared to those who did not. However, the intervention had the opposite effect on explicit gender stereotypes on interest in STEM. In a study on middle-school students, Plant et al. (2009) found that boys reduced their endorsement of gender stereotypes on abilities after interacting with either a male or female agent, while girls' endorsement of gender stereotypes was not affected by the interaction with an agent. Finally, Lewis et al. (2019) found no effect of a brief video showing a mixed-gender team engaging in an engineering task on students' endorsement of gender stereotypes on abilities.

Coming back to the mechanism hypothesized by the 'social role theory', for a role model intervention to be effective, role models should be perceived as a representative sample of social roles in contemporary societies. In other words, any intervention would be condemned to failure if role models are classified as the 'exception that proves the rule'. This mechanism, known as 'subtyping' (Kunda & Oleson, 1995), may prevent stereotypes to change. Dasgupta and Greenwald (2001, p. 808) suggested that the absence of an effect of role models on explicit stereotypes could be related to subtyping and correction, which is more evident when 'perceivers have the cognitive resources to reflect on and re-categorize counterstereotypical exemplars'. This would also explain why they found an effect on implicit stereotypes, which are usually measured through time-constrained psychological tests.

To my knowledge, there is only one study that explicitly tested the effect of role models on the perceived distribution of women and men in STEM-related careers. After an intervention held in class for Grade 10 and Grade 12 French students, Breda et al. (2020) asked participants to what extent they agreed with the statement 'There are more men than women in science-related jobs'. Results showed that the intervention increased both male and female students' awareness of the underrepresentation of women in STEM.

3. Aim and hypotheses

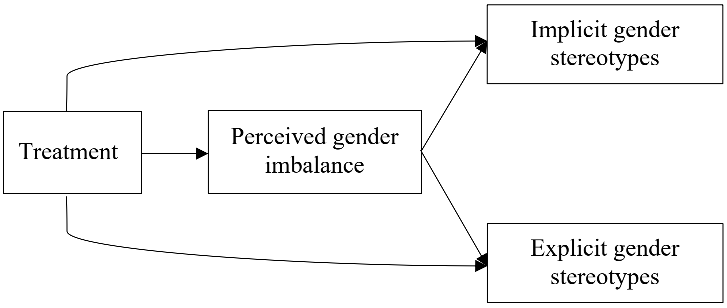
This study aimed to test the effects on high-school students of a brief video showing interviews with both male and female professionals working in the STEM or the humanities-related sectors. Two types of exposure were proposed. The first video showed only counterstereotypical examples (treatment A), i.e., female professionals for STEM studies and male professionals for humanistic studies. The second video showed both stereotypical and counterstereotypical examples (treatment B), i.e., male and female professionals for both the STEM and humanistic sectors. Finally, a control group was assigned to a video reporting the interviews' transcriptions. Those assigned to this group did not know whether the interviewee was a man or a woman.

Specifically, I tested the effect on four factors, distinguishing by participants' gender.

1. the perceived gender imbalance in STEM and humanistic studies.
2. implicit gender-science stereotypes on the association between STEM/humanities studies and gender.
3. explicit gender stereotypes on abilities in STEM and humanities.
4. explicit gender-science stereotypes' attribution, i.e., causes attributed to gender differences in STEM and humanistic studies.

Following the assumption of the social role theory, I also tested whether the effect of the treatments on gender stereotypes was mediated by the perceived gender imbalance, as shown in Figure 6.1.

Figure 6.1 *Indirect effects' model*



Informed by the social role theory I hypothesized a positive effect of both treatments on the four factors, higher for treatment A compared to treatment B, because of the presence of both stereotypical and counterstereotypical role models in the latter case. However, I hypothesized a similar, positive, effect of both treatments on the perceived gender imbalance, with students assigned to treatments more likely to believe that the two genders are equally represented. As found in previous studies on the effect of role models, I expected small effect size. As regards the mediating role of the perceived gender imbalance on gender stereotypes, I hypothesized that the perception of gender imbalance was associated with both implicit and explicit gender stereotypes, i.e., the higher the perceived unbalance, the stronger the endorsement of gender stereotypes is. Consequently, I expected the perceived gender imbalance to mediate the effect of treatments on the two outcomes.

4. Methodology

4.1. Trial design

This was a single-blind randomized control trial (RCT) with a parallel-group design comparing the effects of a brief intervention exposing high school students to role models coming from both the STEM and humanities domains. Participants were told they would have taken part in a research project aimed at studying the relevance of role models for major choices. They were informed that, as part of the project, they would have seen a video collecting interviews with professionals coming from several fields and that the researchers were interested in their opinion on the benefits of hearing about role models' experiences. They were, thus, blind to the experimental nature of the study, the true aim of the study and the existence of multiple interventions and that they were randomly assigned only to one of them. The study was approved by the Ethics committee of the University of Milan.

4.2. Participants

Given the aim of the project and how it was presented to schools, the ideal participants were students not yet enrolled in university but already interested in choosing where to major. While the initial idea was to restrict the sample only to students in the last year of high school and attending a *liceo*, a discussion with schools' guidance counsellors before the beginning of the project suggested that students begin to choose or think about their choice already in the second-last year. Furthermore, while in Italy there is a distinction between academic – *liceo* – and vocational – technical and professional – high schools, according to counsellors most students attending a technical school choose to enrol in college rather than enter the job market right after high school.

Eligible participants were thus students attending either the fourth or the fifth year of high school and enrolled in either a *liceo* or a technical school. Finally, participants were required to accept and sign the informed consent. Participants who drop the study before completion of the post-treatment questionnaire and those who did not pass the attention test (more details below) were excluded. The study was conducted online using the SoSci Survey platform (Leiner, 2019). The sample was unselected within the schools that adhered to the project, including all students who opted in for participation on the online platform.

4.3. Interventions and randomization

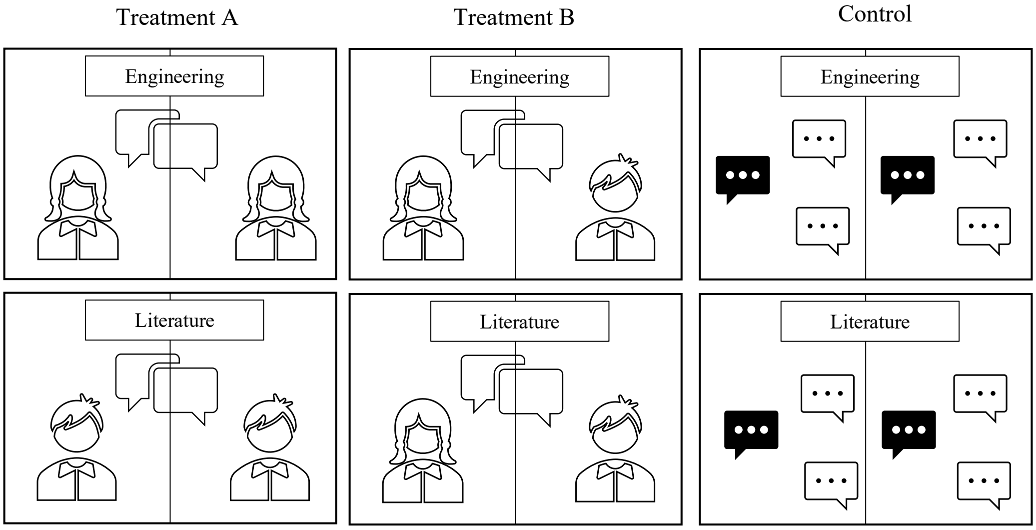
I contacted young professionals who got a degree in one of the following sectors, three related to STEM, i.e., engineering, IT and physics, and three related to humanistic studies, i.e., literature, modern languages and psychology. Those who agreed were video interviewed and asked to briefly answer a few questions about their job and their experience in their field (more details in the Appendix). These interviews were edited into a video collecting double interviews, i.e., for each of the six sectors, two professionals were shown in the video side by side. An off-screen voice asked the questions and interviewees, alternately, briefly answered them. This was repeated for all the sectors. Three different videos were created to reflect the two treatments and the control condition (a frame of each video is reported in the Appendix), lasting around ten minutes.

- Treatment A video showed only counterstereotypical exemplars, i.e., female professionals when the sector was STEM-related, male professionals when it was humanities-related.

- Treatment B video showed both a male and a female professional for each sector.
- The control video showed the transcription of the interviews as an online chat between the interviewees and the interviewer. Participants did not have any clue about professionals' gender, as this could not be inferred by the answers of the interviewees.

Figure 6.2 shows two illustrative frames of the video presented. In the control group, white boxes represent the interviewer's questions and black boxes the interviewees' answers. The two, questions and answers, appeared on the screen sequentially, mimicking a normal conversation on a chat platform. Participants were randomly assigned to one of the three interventions, with a normal equally distributed type of draw, which ensures that every intervention is drawn equally often, available on the SoSci Survey platform.

Figure 6.2 *Illustrative frames of the videos used in the experiment*



4.4. Procedure

I contacted a large number of Italian high schools and asked them to advertise the project among their students, by circulating a leaflet with a brief description of the project and a link to the experiment's platform. Participation was on a voluntary basis. Those interested in participating were invited to click on the link provided in the leaflet and follow the instructions. The website on which the study was conducted was organized in sequential pages, guiding the participant through the phases of the experiment, i.e., (1) informed consent, (2) pre-treatment questionnaire, (3) intervention, (4) attention test, (5) post-treatment questionnaire. Students could abandon the project at any moment. Those who did not accept the informed consent were not allowed to proceed. The experiment took place between March and May 2021. In October 2021, students who were in the last year of high school at the time of the experiment were contacted again via email and asked about the major they eventually chose or intended to choose if they were still in high school. However, due to the insufficient number of answers, I did not present here the results on major choice.

4.5. Outcomes

Perceived gender gap in STEM and humanities: Students were asked to indicate the representativeness of women and men in both four STEM (physics, engineering, maths, IT) and four humanistic (literature, modern languages, education and psychology) majors. They could choose among five options, i.e., ‘Almost all men’, ‘More men than women’, ‘Equal number of women and men’, ‘More women than men’, ‘Almost all women’ (Barth et al., 2018). The Cronbach alpha was too low to construct a single instrument ($\alpha_{\mathrm{STEM}} = .57$; $\alpha_{\mathrm{Hum}} = .49$), thus results are presented separately for each major.

Implicit gender-science stereotypes: Automatic association of gender with STEM and humanistic majors was tested using the Implicit Association Test (Greenwald et al., 1998). The IAT measures the difference in the time needed to do an association between compatible constructs (e.g., women and humanistic majors, men and STEM majors) and the time needed to do an association with incompatible constructs (e.g., women and STEM majors, men and humanistic majors), where compatibility reflects stereotypical beliefs. Being aware of the debate on the IAT (Fiedler et al., 2006; Jost, 2019), I decided to use this test as it is still the most adopted psychological test to measure implicit stereotypes. The version of the IAT used here required participants to associate male and female names with STEM (Physics, Mathematics, Engineering, Astronomy, Geology, Information Technology, Chemistry) and humanistic majors (Art, Literature, Philosophy, History, Modern Languages, Italian, History of Art). Results report the D score, suggested by Greenwald et al. (2003). This indicator ranges from -2 to +2, where negative values mean that it is easier for the participant to associate incompatible rather than compatible constructs, a value around 0 indicates that the participant is indifferent between compatible and incompatible constructs, while a positive value means that it is easier for him/her to associate compatible rather than incompatible constructs. Thus, the higher the value of the D score, the stronger automatic stereotypical beliefs are (see Table 4.7 in the Appendix for more details on the IAT).

Explicit gender stereotypes on abilities: Students were asked to rate on a 5-point Likert scale the extent to which they agreed with statements on gendered abilities in STEM and humanistic studies, i.e., ‘Men are generally more inclined to scientific studies’, ‘Women are generally more inclined to humanistic studies’ (Galdi et al., 2017).

Causes attributed to the gender gap: Students were asked to rate on a 5-point Likert scale their agreement on statements regarding the reason for the observed gender gap, in both the STEM and humanities areas. Items were adapted from Nosek et al. (1998). Almost half of these items suggested a cause ascribable to individual aptitude and biological characteristics, e.g., ‘Women are usually better than men in humanistic studies because they are by nature more sensitive’; ‘If there are more men than women in STEM studies is because men are more interested in this field’. The remaining statements suggested a cause ascribable to social pressure, e.g., ‘Men are encouraged more than women to choose STEM-related majors’; ‘If there are more women than men in humanistic studies is because men are hampered and discriminated in this field’ (see Table 6.8 in the Appendix for the complete list of items).

I created two indicators for each field of study (STEM and humanities), one called ‘Attribution to biological characteristics’ (Cronbach alpha 0.67 for STEM, 0.68 for humanities) and the other called ‘Attribution to social pressure’ (Cronbach alpha 0.81 for STEM, 0.64 for humanities). I built an indicator of the difference in the propensity to attribute the gender gap to biological rather than social causes by measuring for both sectors the difference between the two scales. The indicator was zero when there was no difference in the attribution of gender differences to any of the two explanations. A positive value indicated that biological characteristics counted more than social pressure, while a negative value indicated that social pressure counted more than biological characteristics in explaining the gender gap.

4.6. Statistical methods

The analysed sample consisted of students who completed both the pre-treatment and the post-treatment questionnaire and that passed the attention test. The latter consisted of a set of questions asked right after the video, asking participants about the jobs of the professionals interviewed in the video. Those who answered at least half of the questions correctly passed the test. I also used a second criterion based on the time spent on the page of the video. Those who spent less than the length of the video and more than one hour were dropped from the final sample.

The effect on the outcomes was tested using ordered logistic regression models, except for the implicit gender-science stereotype, which was tested using a linear regression model. Differences among the three groups were verified using a post-estimation Chi-square test, with Holm's correction for multiple comparisons (Holm, 1979). In the case of statistically significant results, marginal effects were computed to assess the entity and direction of the difference between groups. Given that previous research mostly highlighted that the response differed depending on the participant's gender (Cheryan et al., 2011; Marx & Roman, 2002), I tested the effect separately for male and female participants.

To verify the mediating role of the perceived gender imbalance on treatment's effect, I performed a mediation analysis for implicit and explicit stereotypical beliefs. Indirect effects' significance was determined using percentile bootstrap 95% confidence intervals (Bollen & Stine, 1990), while the direct association between perceived gender imbalance and the outcomes of interest was estimated using an ordinal regression model in the case of explicit gender stereotypes and a linear regression model in case of implicit gender stereotypes.

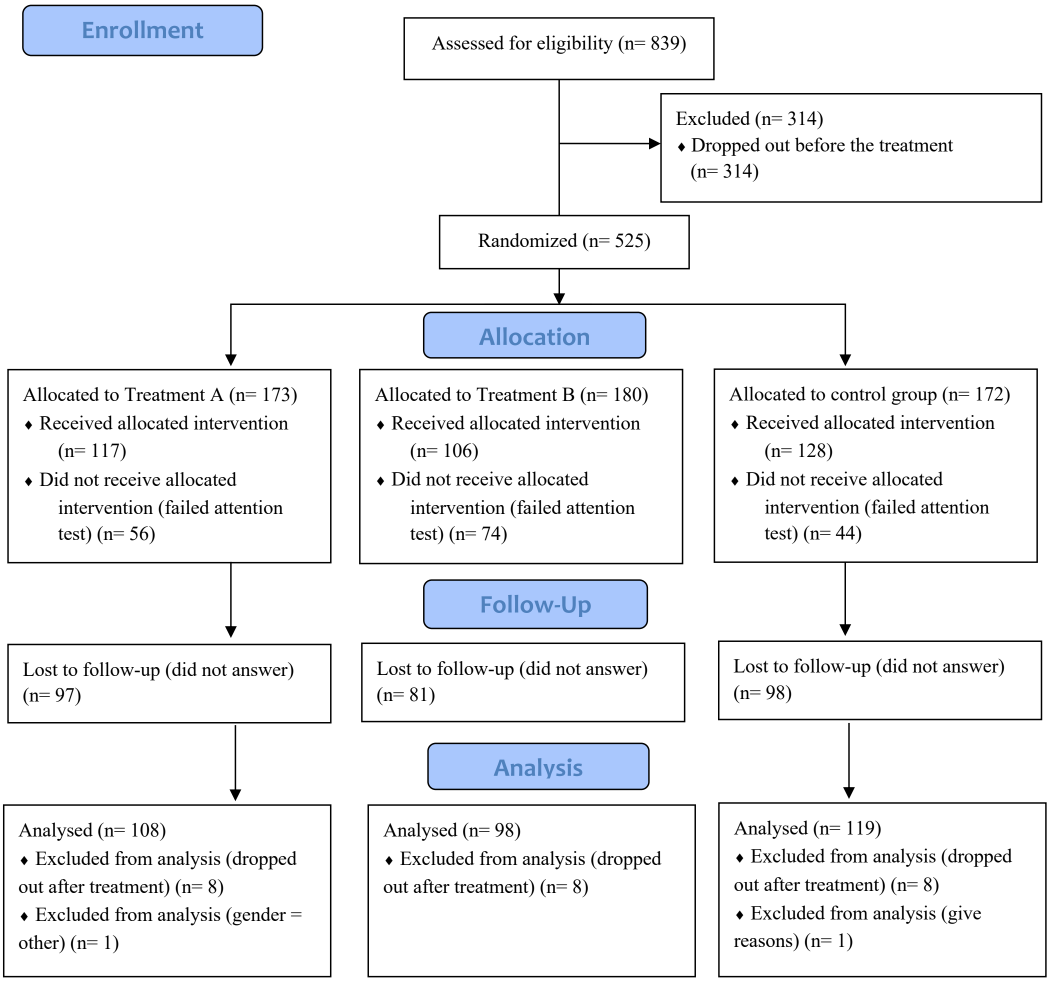
I did not set a fixed sample size before the data collection as students' participation was on a voluntary basis. A sensitivity power analysis using G*Power (Faul et al., 2007) suggests that, for the female sample ($N = 236$, $\alpha = .05$, power = .8, predictors = 2) the minimum detectable effect is $f^2 = .04$, while for the male sample ($N = 89$, $\alpha = .05$, power = .8, predictors = 2) the minimum detectable effect size is $f^2 = .11$, assuming a continuous dependent variable.

5. Results

5.1. Participant flow and recruitment

Figure 6.3 summarizes the phases of the randomised trial of the three groups (Schulz et al., 2010). The sample analysed consisted of 325 participants, of which 108 were assigned to treatment A, 98 to treatment B and 119 to the control group.

Figure 6.3 *CONSORT participant flow diagram*



5.2. Demographic characteristics

Most participants were female (73%) and Italian (95%). 62% of them were attending the last year of high school at the time of the project, while 38% of them were attending the penultimate year. Almost all came from a *liceo*, while a minority from a technical school. In the pre-treatment questionnaire, 32% of the sample expressed the intention of majoring in STEM, while 39% of the sample expressed the intention of majoring in humanities. There were no differences in the characteristics of the students assigned to the three treatment groups, as shown in Table 6.1.

Table 6.1 *Participants' demographic characteristics by treatment*

Characteristic	Treatment A		Treatment B		Control	
	Count	%	Count	%	Count	%
Gender [Female]	76	70%	75	77%	85	71%
Nationality [Italian]	106	98%	93	95%	111	93%
Academic year [last]	64	59%	64	65%	72	61%
Type of school [*liceo*]	103	95%	94	96%	109	91%
Interested in STEM-related majors	37	34%	23	24%	43	36%
Interested in humanities-related majors	48	44%	35	36%	42	35%

5.3. Outcomes

5.3.1. Perceived gender imbalance

On average, there was a small difference in the perceived gender imbalance between those assigned to treatments and those assigned to the control group, in line with the hypothesis in the case of male participants, in the opposite direction in the case of female participants. Young women were more likely to believe that there are almost all women in humanities-related majors and less likely to believe that there is gender parity, while the opposite occurred for male participants. As regards STEM majors, regardless of gender, there was a positive difference between those assigned to treatment B and those in the control group, i.e., more participants believed that there is gender parity in scientific studies, while the pattern was the opposite when comparing those assigned to treatment A and those in the control group. However, data were insufficient to confirm that the difference was statistically significant in most cases, with some exceptions, shown below.

Table 6.2 reports the results of the Chi-square tests conducted after regression models to test the difference in the perceived gender imbalance between the two treatments and the control group.

Table 6.2 *Chi-square test on perceived gender imbalance*

Null hypothesis	Female		Male	
	Chi2	p-value	Chi2	p-value
	Humanities			
Literature				
Treatment A = Control	.02	.881	.00	.947
Treatment A = Treatment B	.29	.591	.06	.799
Treatment B = Control	.49	.485	.09	.758
Languages				
Treatment A = Control	.15	.697	.44	.509
Treatment A = Treatment B	.91	.339	.97	.324
Treatment B = Control	.33	.566	.193	.164
Psychology				
Treatment A = Control	4.90	**.027**	3.52	.061
Treatment A = Treatment B	4.76	**.029**	1.54	.214
Treatment B = Control	.01	.941	8.55	**.003**
Education				
Treatment A = Control	5.13	**.024**	.43	.509
Treatment A = Treatment B	.37	.542	.01	.928
Treatment B = Control	3.01	.083	.42	.518
	STEM			
Engineering				
Treatment A = Control	5.26	**.022**	.04	.842
Treatment A = Treatment B	3.84	**.050**	2.97	.085
Treatment B = Control	.52	.471	2.73	.098
Physics				
Treatment A = Control	.00	.979	1.87	.172
Treatment A = Treatment B	1.42	.233	4.41	**.036**
Treatment B = Control	1.69	.193	1.51	.219
IT				
Treatment A = Control	.57	.451	.03	.861
Treatment A = Treatment B	1.26	.261	3.02	.082
Treatment B = Control	.14	.709	1.65	.199
Math				
Treatment A = Control	.15	.696	.60	.439
Treatment A = Treatment B	.01	.923	.10	.754

Treatment B = Control	.11	.740	.81	.368

Note. N_{female} = 236; N_{male} = 87; df = 1

Holm-adjusted p-values for multiple comparison

As regards humanistic studies, in the case of female participants, there was a statistically significant difference in the perceived gender imbalance in psychology between those assigned to treatment A and those assigned to either treatment B or the control group. Contrary to the initial hypothesis, compared to the control group, young women in treatment A were more likely to believe that in psychology there are almost all women (CME[1] = .12, z = 2.17, p-value = .030) and less likely to believe that there is an equal number of women and men (CME = -.09, z = -2.20, p-value = .028). Similarly, compared to those assigned to treatment B, those in treatment A were more likely to believe that in psychology there are almost all women (CME = .12, z = 2.18, p-value = .030) and less likely to believe that there is an equal number of women and men (CME = -.09, z = -2.11, p-value = .035). There was also a statistically significant difference in the perception on educational studies, i.e., female participants assigned to treatment A were more likely to believe there are almost all women in educational studies (CME = .16, z = 2.30, p-value = .021) and less likely to believe that there are more women than men (CME = -.10, z = -2.17, p-value = .030) or an equal number of women and men (CME = -.06, z = -2.06, p-value = .039).

In the case of male students, there was a statistically significant difference in the perceived gender imbalance in psychological studies. In this case, the difference is coherent with the hypotheses, i.e., those assigned to treatment B were more likely to believe there is an equal number of men and women (CME = .29, z = 2.99, p-value = .003) and less likely to believe that there are almost all women (CME = -.21, z = -2.74, p-value = .006) compared to those in the control group.

As regards STEM-related studies, I found a statistically significant difference among young women for studies in engineering, and among young men for studies in physics. In both cases, the difference disconfirms the hypotheses. Female participants assigned to treatment A were less likely to believe there is an equal number of men and women in engineering (CME = -.11, z = -2.28, p-value = .022) and more likely to believe that there are almost all men (CME = .12,

---

[1] CME = Conditional marginal effect

z = 2.24, p-value = .025) compared to those in the control group. A similar difference is observed comparing treatment A group and treatment B group, i.e., the former was less likely to believe there is an equal number of men and women in engineering (CME = -.08, z = -2.08, p-value = .038). Finally, male participants assigned to treatment A were more likely to believe there are almost all men in physics-related studies (CME = .13, z = 2.14, p-value = .033) and less likely to believe there is an equal number of men and women (CME = -.23, z = -2.01, p-value = .044) compared to those in the control group.

5.3.2. Implicit gender-science stereotypes

On average, students assigned to either treatment A or treatment B, regardless of gender, exhibited weaker implicit gender-science stereotypes compared to those in the control group. However, data were insufficient to confirm that the difference was statistically significant (see Table 6.9 in the Appendix).

5.3.3. Explicit gender stereotypes

On average, female students assigned to either treatment A or treatment B exhibited stronger explicit gender stereotypes in both STEM and humanities compared to those in the control group. On the contrary, male students assigned to either treatment A or B exhibited lower explicit gender stereotypes in both STEM and humanities compared to those in the control group. The only exception is represented by those assigned to treatment B who exhibited stronger explicit gender stereotypes on STEM compared to those in the control group. However, data were insufficient to confirm that the difference was statistically significant (see Table 6.9 in the Appendix).

5.3.4. Causes attributed to the gender gap

On average, compared to those in the control group, participants believed that social pressure counts more than biological characteristics in explaining the gender gap in STEM studies, except for male participants assigned to treatment B. As regards the gender gap in humanities, differences were mixed, with young women assigned to treatment A and young men assigned to treatment B perceiving biological characteristics as more relevant, while young women assigned to treatment B and young men assigned to treatment A perceiving social pressure as more relevant. However, data were sufficient to confirm the differences are statistically significant only in some cases, as shown in Table 6.3.

Table 6.3 *Chi-square test on stereotypes' attribution*

Null hypothesis	Female		Male	
	Chi2	p-value	Chi2	p-value
Difference in stereotypes’ attribution [gender gap in humanities]				
Treatment A = Control	.25	.621	.84	.359
Treatment A = Treatment B	.97	.384	2.66	.103
Treatment B = Control	.26	.614	.58	.447
Difference in stereotypes’ attribution [gender gap in STEM]				
Treatment A = Control	.09	.767	1.89	.169
Treatment A = Treatment B	2.76	.097	5.49	**.019**
Treatment B = Control	4.01	**.045**	1.10	.294

Note. N_{female} = 236; N_{male} = 87; df = 1

Holm-adjusted p-values for multiple comparison

As regards female participants, there was a statistically significant difference between those assigned to treatment B and those in the control group. Young women in the former group were more likely to believe social pressure counts more in explaining gender differences in STEM studies (CME = .15, z = 2.05, p-value = .041), while they were less likely to believe biological characteristics counts more (CME = -.07, z = -2.03, p-value = .042). As regards male participants, there was a statistically significant difference between those assigned to treatment B and those assigned to treatment A. Young men in the former group were more likely to consider biological characteristics more important in explaining gender differences in STEM studies (CME = .25, z = 2.39, p-value = .017), while they were less likely to consider social pressure as more relevant (CME = -.25, z = -2.43, p-value = .015).

5.4. Mediation analysis

I, finally, tested whether the effect of the treatments on implicit and explicit gender stereotypes was mediated by the perceived gender gap in STEM and humanities. Table 6.4 summarizes the direct effects of perceived gender imbalance on implicit and explicit gender-science stereotypes for female participants. On average, results confirm the hypothesis that perceiving the sector as more unequal is associated with stronger gender stereotypes. However, data were insufficient to confirm that the association was statistically significant for the perceived gender imbalance in STEM and in the case of explicit gender stereotypes in STEM.

Compared to those who believe that there is an equal number of women and men in the humanities sector, participants who believe there are almost all women exhibited stronger implicit and explicit stereotypes. Results suggest that data are not sufficient to reject the null hypothesis of no effect on the perceived gender imbalance in STEM. As regards male participants, results were not statistically significant, with one exception. Contrary to expectations, male participants who believe that in the STEM sector there are almost all men exhibited weaker implicit gender stereotypes compared to those who believe there is an equal number of women and men (see Table 6.10 in the Appendix).

Table 6.4 *Direct effect of perceived gender imbalance on stereotypes (females)*

Predictor	Estimate	SE	t	p-value	95% CI	
					LB	UB
Implicit gender stereotypes						
Perceived gender imbalance[1] (hum)						
More women than men	.17	.11	1.58	.117	-.04	.38
Almost all women	.32	.12	2.69	.008	.09	.56
Perceived gender imbalance[1] (STEM)						
Almost all men	.16	.09	1.90	.059	-.01	.33
More men than women	-.02	.06	-.32	.751	-.15	.11
Explicit gender stereotypes (humanities)						
Perceived gender imbalance[1] (hum)						
More women than men	.72	.54	1.34	.180	-.33	1.78
Almost all women	1.17	.64	1.82	.069	-.09	2.43
Explicit gender stereotypes (STEM)						
Perceived gender imbalance[1] (STEM)						
Almost all men	1.54	.70	2.19	.028	.16	2.92
More men than women	.94	.39	2.43	.015	.18	1.70

Note. SE = Robust standard errors; CI = Confidence interval; LB = Lower bound; UB = Upper bound

[1] Reference category: 'Equal number of women and men'

However, indirect effects were almost equal to zero and, based on percentile bootstrap 95% confidence intervals, only the indirect effect of treatment A on implicit gender stereotypes through the perceived gender imbalance in humanities was statistically significant (see Table 6.11 in the Appendix).

6. Discussion and conclusions

This study aimed to test the effect of a role models' intervention on female and male high school students. In particular, it verified whether being exposed only to counterstereotypical examples – women working in STEM, men working in humanities – or to both stereotypical and counterstereotypical models could affect the perception of the representativeness of women and men in STEM and humanities sectors. It also explored the influence of these interventions on both implicit and explicit gender stereotypes. Due to the insufficient number of answers to the follow-up, it was not possible to test the effect on major choice six months after exposure.

Results from the experiment are mixed and differ depending on the students' gender. Compared to those assigned to either stereotypical and counterstereotypical role models or the control group, female participants exposed only to counterstereotypical examples were less likely to believe that there is an equal number of women and men in psychological and education-related studies. They were also more likely to believe that there are almost all men in engineering studies, thus disconfirming the initial hypothesis. As regards male participants, those assigned to both types of role models were more likely than those in the control group to believe that there is an equal number of women and men in psychological studies, as hypothesized. Conversely, those assigned to counterstereotypical role models were more likely to believe that there are almost all men in physics-related studies compared to those assigned to both types of role models.

As regards implicit and explicit gender stereotypes, I did not find any statistically significant difference between the two treatment groups and the control group. Finally, as regards stereotypes' attribution, I found that female participants assigned to both types of role models considered social pressure as more relevant in explaining the gender gap in STEM than biological characteristics, compared to those assigned to the control group. The opposite effect was found between male participants assigned to both types of role models and those assigned to counterstereotypical models. The former group was more likely to believe social pressure counts less in explaining the gender gap in STEM.

Results from the mediation analysis suggest that, for female participants, implicit gender stereotypes were associated with the perception of the gender gap in humanities, while explicit gender stereotypes in STEM were associated with the perception of the gender gap in STEM.

The association confirms the initial hypotheses, i.e., a more equal perception of the sector is associated with weaker gender stereotypes.

The mixed nature of the study's results prevents us from either confirming or disconfirming the mechanism hypothesized by the social role theory, i.e., that exposure to counterstereotypical examples would induce a more equal perception of the male- and female-dominated sectors that, in turn, would favour the decrease in the endorsement of gender stereotypes. Nevertheless, the experiment provides some interesting information.

First of all, the negative effect of the exposure to counterstereotypical role models on young women's perception of the gender gap could be explained by the mechanism known as subtyping, mentioned in the introduction. Role models may be seen as an exception rather than representative of workers in those sectors. This finding is coherent with Breda et al. (2020)'s outcomes. In their study, students attending Grade 10 and Grade 12 met female researchers or professionals employed in the STEM sector, who explicitly talked about the underrepresentation of women in this area. After the exposure, students were more aware of the gender gap in the STEM sector. While in this case, the gender gap issue was not explicitly mentioned in the interviews, still, the exposure seems to have strengthened the idea of a wide difference in the representation of women and men in the two sectors. Interestingly, a difference was found between both female students assigned to both counterstereotypical and stereotypical role models and those in the control group.

Furthermore, the difference was confirmed only for some of the majors included. While this could be related to the sample size, it also suggests that we should not treat majors as equivalent, even when related to the same, broad, domain. As discussed in Chapter 2 regarding instruments, and in Chapter 3 when discussing the models' results, the gender gap is not homogenous among scientific domains. Further research would be needed on the heterogeneity of women's representation in these sectors and on whether gender stereotypes differently affect women's attitudes toward them.

It is interesting to note that the perceived representativeness of women and men in humanities was associated with the endorsement of implicit gender stereotypes of female participants, while the perceived representativeness in STEM was associated with females' endorsement of explicit gender stereotypes. In particular, as predicted by the 'social role theory', those who believe that sectors were more gender-balanced endorsed a weaker explicit and implicit association between gender and the two sectors. Unfortunately, the fact that the implicit association test did not allow us to disentangle the weight of the men/STEM/ vs. women/humanities associations prevent us from understanding the different contribution of the awareness of the gender gap on the automatic association of gender and STEM/humanities. Further research would be needed on the difference between the two complementary stereotypes activated at the implicit level.

While results are mixed and did not allow us to draw any general conclusions, the experimental findings clearly indicate that the issue of the underrepresentation of men in humanities deserves careful attention and would require a contextualised approach. While participants were aware of the gender gap in both sectors, female students were convinced that the gap was wider in the humanities field compared to the STEM field (see Tables 6.5 – 6.7 in the Appendix). However, the width of the gender gap in Italy is similar in the two sectors. Women graduating in 2020 in Italy represented 62% of grad students in the humanities field, 84% of those in modern languages and 81% of those in psychology. On the contrary, men represented 86% of those graduating in IT, 50% in math, 70% in physics and 74% in industrial engineering. Interestingly, whenever asked about what they believed were the causes underlying these gender differences, 48% of students gave more credit to social-related factors in the case of differences in STEM, whereas only 14% shared this opinion for differences in humanities. In the latter case, regardless of their gender, students mostly believed that biological-related characteristics were responsible for such a difference.

This is relevant as, while women suffer from a social penalization when entering the STEM field, men's underrepresentation is rather attributed to biological – so immutable – causes. As suggested by Croft et al. (2015, p. 361),

> ‘the threat of identity misclassification or risk of losing status might mean that the costs of behaving counterstereotypically are even more pronounced for men than they are for women’.

This difference in the gender gap’s perception could contribute to explaining why female students assigned to the gender parity condition gave more credit to social pressure as a cause of the gender differences in STEM, while we did not observe anything similar for the gender differences in humanities. Here, further research would be needed on whether exposure to role models may also affect men’s attitudes toward female-dominated sectors.

This said this study has various limitations. First, the sample size is relatively small given the number of treatment and control groups with all problems in detecting a small effect size, especially for the male sample. Furthermore, the experiment was conducted online, a setting that limits researchers' control over participants. However, I tried to limit the risk of including participants who did not see the videos by using an attention test. On the other hand, performing well-controlled lab or field experiments during a global pandemic was impossible. Online experiments do have some advantages, e.g., higher ecological validity, the possibility to reach a sample with more demographic diversity, and reduced logistical constraints (van Steenbergen & Bocanegra, 2016). Finally, results could reflect instruments' limitations. As previously mentioned, there are well-known arguments against the validity of the Implicit Association Test, which were partly solved by the use of an improved algorithm for the final score (Greenwald et al., 2003). Unfortunately, this test did not allow us to disentangle the two associations – men/ STEM and women/humanities.

To conclude, while policymakers and the public usually see role models as an effective solution for the underrepresentation of women in STEM, results from academic research are controversial. As found in this study, the risk of a subtyping effect after exposure is possible, with detrimental and unintended consequences on participants. Future research should focus on understanding when and how role models' interventions are beneficial to women, but also men in female-dominated sectors. Furthermore, this study did not clarify whether these interventions can have also medium- and long-term effects on intentions and behaviour.

References

Andre, T., Whigham, M., Hendrickson, A., & Chambers, S. (1999). Competency beliefs, positive affect, and gender stereotypes of elementary students and their parents about science versus other school subjects. *Journal of Research in Science Teaching*, *36*(6), 719–747. https://doi.org/10.1002/(SICI)1098-2736(199908)36:6<719::AID-TEA8>3.0.CO;2-R

Barth, J. M., Kim, H., Eno, C. A., & Guadagno, R. E. (2018). Matching Abilities to Careers for Others and Self: Do Gender Stereotypes Matter to Students in Advanced Math and Science Classes? *Sex Roles*, *79*(1–2), 83–97. https://doi.org/10.1007/s11199-017-0857-5

Blair, I. V. (2002). The Malleability of Automatic Stereotypes and Prejudice. *Personality and Social Psychology Review*, *6*(3), 242–261. https://doi.org/10.1207/S15327957PSPR0603_8

Bollen, K. A., & Stine, R. (1990). Direct and Indirect Effects: Classical and Bootstrap Estimates of Variability. *Sociological Methodology*, *20*, 115. https://doi.org/10.2307/271084

Breda, T., Grenet, J., Monnet, M., & Effenterre, C. V. (2020). *Do Female Role Models Reduce the Gender Gap in Science? Evidence from French High Schools* (IZA Discussion Papers No. 13163). Institute of Labor Economics (IZA). http://hdl.handle.net/10419/216475

Cheryan, S., Meltzoff, A. N., & Kim, S. (2011). Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computers & Education*, *57*(2), 1825–1835. https://doi.org/10.1016/j.compedu.2011.02.004

Coto, J., Pulgaron, E. R., Graziano, P. A., Bagner, D. M., Villa, M., Malik, J. A., & Delamater, A. M. (2019). Parents as Role Models: Associations Between Parent and Young Children's Weight, Dietary Intake, and Physical Activity in a Minority Sample. *Maternal and Child Health Journal*, *23*(7), 943–950. https://doi.org/10.1007/s10995-018-02722-z

Croft, A., Schmader, T., & Block, K. (2015). An Underexamined Inequality: Cultural and Psychological Barriers to Men's Engagement With Communal Roles. *Personality and Social Psychology Review*, *19*(4), 343–370. https://doi.org/10.1177/1088868314564789

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of*

Personality and Social Psychology, *81*(5), 800–814. https://doi.org/10.1037/0022-3514.81.5.800

Eagly, A. H., & Wood, W. (2012). Social Role Theory. In *Handbook of Theories of Social Psychology: Volume 2* (Vol. 1–2, pp. 458–476). SAGE Publications Ltd. https://doi.org/10.4135/9781446249222

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I", the "A", and the "T": A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology*, *17*(1), 74–147. https://doi.org/10.1080/10463280600681248

Galdi, S., Mirisola, A., & Tomasetto, C. (2017). On the relations between parents' and children's implicit and explicit academic gender stereotypes. *Psicologia sociale*, *2*, 215–238. https://doi.org/10.1482/87248

Gibson, D. E. (2003). Developing the Professional Self-Concept: Role Model Construals in Early, Middle, and Late Career Stages. *Organization Science*, *14*(5), 591–610. https://doi.org/10.1287/orsc.14.5.591.16767

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Henrich, J. P. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton university press.

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. JSTOR. http://www.jstor.org/stable/4615733

Jost, J. T. (2019). The IAT Is Dead, Long Live the IAT: Context-Sensitive Measures of Implicit Attitudes Are Indispensable to Social and Political Psychology. *Current Directions in Psychological Science*, *28*(1), 10–19. https://doi.org/10.1177/0963721418797309

Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, *68*(4), 565–579. https://doi.org/10.1037//0022-3514.68.4.565

Leiner, D. J. (2019). *SoSci Survey* (3.1.06) [Computer software]. https://www.soscisurvey.de

Lewis, N. A., Sekaquaptewa, D., & Meadows, L. A. (2019). Modeling gender counter-stereotypic group behavior: A brief video intervention reduces participation gender gaps on STEM teams. *Social Psychology of Education*. https://doi.org/10.1007/s11218-019-09489-3

Marx, D. M., & Roman, J. S. (2002). Female Role Models: Protecting Women's Math Test Performance. *Personality and Social Psychology Bulletin*, *28*(9), 1183–1193. https://doi.org/10.1177/01461672022812004

Microsoft Philantropies. (2018). *Closing the STEM gap: Why STEM classes and careers still lack girls and waht we can do about it*. Microsoft. https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE1UMWz

Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology*, *107*(3), 631–644. https://doi.org/10.1037/edu0000005

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*(4), 465–483. https://doi.org/10.1037/gpr0000059

Murrell, A. J., Crosby, F. J., & Ely, R. J. (1999). *Mentoring Dilemmas: Developmental Relationships Within Multicultural Organizations*. Taylor & Francis. http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=425237

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., … Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593–10597. https://doi.org/10.1073/pnas.0809921106

Nosek, B., Banaji, M. R., & Greenwald, A. G. (1998). *Project Implicit*. https://implicit.harvard.edu/implicit/index.jsp

Olsson, M., & Martiny, S. E. (2018). Does Exposure to Counterstereotypical Role Models Influence Girls' and Women's Gender Stereotypes and Career Choices? A Review of Social Psychological Research. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02264

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*(2), 209–215. https://doi.org/10.1016/j.compedu.2009.01.013

Scherer, R. F., Adams, J. S., Carley, S. S., & Wiebe, F. A. (1989). Role Model Performance Effects on Development of Entrepreneurial Career Preference. *Entrepreneurship Theory and Practice*, *13*(3), 53–72. https://doi.org/10.1177/104225878901300306

Schulz, K.F., Altman, D.G., Moher, D., CONSORT Group. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ (Clinical research ed.), 340*, c332. https://doi.org/10.1136/bmj.c332

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*(2), 255–270. https://doi.org/10.1037/a0021385

Tomasetto, C., Galdi, S., & Cadinu, M. (2012). Quando l'implicito precede l'esplicito: Gli stereotipi di genere sulla matematica in bambine e bambini di 6 anni. *Psicologia sociale*, *2*, 169–186. https://doi.org/10.1482/37693

Van Camp, A. R., Gilbert, P. N., & O'Brien, L. T. (2019). Testing the effects of a role model intervention on women's STEM outcomes. *Social Psychology of Education*, *22*(3), 649–671. https://doi.org/10.1007/s11218-019-09498-2 van Steenbergen, H., & Bocanegra, B. R. (2016). Promises and pitfalls of Web-based experimentation in the advance of replicable psychological science: A reply to Plant (2015). *Behavior Research Methods*, *48*(4), 1713–1717. https://doi.org/10.3758/s13428-015-0677-x

Chapter 7 ***Conclusion***

This dissertation aimed to contribute to the lively discussion on gender stereotypes and their link to the persistent gender gap in sectors traditionally considered as either masculine or feminine, i.e., STEM and humanities. Despite the numerous and long-standing studies on this theme (Kanny, Sax, and Riggers-Pieh 2014; Wang and Degol 2017), this dissertation essentially proved that if we consider these issues carefully, there is still room for refinement, debate and hopefully valuable new contributions.

Starting from the most basics contribution of this thesis, the review of the instruments used to measure gender stereotypes revealed that we are far from agreeing on the definition of stereotypical beliefs on gender and STEM. As observed by Zosuls et al. (2011:6) in their review on the trends in research on gender development:

> ‘the term “gender stereotyping” was used without indication of whether gender stereotypes were assessed in terms of personal stereotype beliefs, knowledge of cultural stereotypes, stereotyped judgments or the enactment of stereotype-consistent behaviours. […] As such, applying the general term “gender stereotyping” without explicit indication of whether gender-stereotyped beliefs, knowledge, or behaviours are being measured can confuse and more importantly, conflate conceptually distinct constructs’.

Moving to the studies on real-world settings, results from both the network analysis on ability attribution and the association between implicit gender-science stereotypes and major intentions confirmed that, despite all interventions and policies, gender stereotypes still seem to play an active role in shaping attitudes and behaviours. On the one hand, expectations on performance reflected gender constructs even when information on abilities was available. On the other hand, female students' major intentions were negatively correlated with their endorsement of implicit gender stereotypes.

Finally, the two studies on the role models contributed to the literature tackling the issue of the gender gap in STEM. One reported a review conducted with a systematic approach on all existing studies. The other provided insights on the effect of an online experiment on the perception of the gender gap and the endorsement of both implicit and explicit gender stereotypes. Results suggest that the effect of role models is mixed showing that further research is needed to identify the conditions necessary for a beneficial and effective influence of role models on students.

Beyond the specific contributions of each study, this dissertation highlighted two major issues. The first is the need to include men and the humanities sector in the discussion. While this perspective has been essentially neglected in previous research, the two associations, men with STEM and women with humanities should be considered together as they are complementary stereotypes. While research has already pointed out the existence of gender differences in facing non-traditional careers and educational paths (e.g., Croft, Schmader, and Block 2015; Simpson 2005), this issue has still not gained sufficient attention and empirical evidence is too scarce to provide a clear picture. As shown by Croft et al. (2015:356), any increase in the shift of men to female-dominated sectors could be beneficial in several ways. Eventually,

> 'the scarcity of men in communal roles establishes strong behavioural norms about what constitutes appropriate behaviour for men and women' thus 'as more men move into communal roles, the rigidity of gender stereotypes should also be diminished'.

Furthermore, breaking the association between women and humanities could also serve the scope of increasing female participation in the STEM sector. An interesting insight on this comes from a longitudinal study conducted by Wang et al. (2013) on 12th grade students, who were interviewed again fifteen years after the first wave, at the age of 33, and asked about their occupations. The study shows that those who, in school, were high proficient both in maths and verbal tests were less likely to pursue a career in STEM, than those who had high math skills but moderate verbal skills. As suggested by the authors, this could be explained by the fact that individuals are more likely to choose a career in which they think they can succeed.

Apparently, abilities may not be enough to convince girls to abandon a path that seems to be safer for them.

The second point is the need to overcome the focus on math whenever referring to women facing stereotypical contexts. This review of instruments highlighted a strong tendency of studies on STEM gender stereotypes to refer uniquely to math rather than to the whole STEM sector or other scientific fields of study. Unfortunately, this limits the extent to which instruments are measuring science-related beliefs. As previously mentioned in this thesis, trends in the performance of girls and boys suggest that the gender gap in math is narrowing (OECD 2020). Indeed, in Italy female bachelor graduates in math were 50% in 2020, while they were 30% in physics, 26% in engineering and 14% in IT (Almalaurea 2021). Coherently with these trends, the network analysis on ability attribution found that the tendency to nominate more frequently men for the math competition was not statistically significant, while the difference in nominations for the science competition was higher and significant. More attention should be given to all science-related fields, as the risk of current research is to underestimate the endorsement of gender stereotypes by concentrating only on people's opinions on math.

As claimed by Cheryan et al. (2017), this difference among STEM fields derives from a difference in their masculine culture, insufficient early experience and differences in self-efficacy. By reviewing studies on the theme, they found that contrary to other scientific fields, i.e., computer science, engineering and physics, biology and chemistry are also associated with men but to a lower extent, as they are perceived to have a higher proportion of women. Being the underrepresentation of women less pronounced in maths, biology and chemistry, there are also differences in the number of role models available compared to other STEM fields. Finally, while some of these fields are school subjects that all students are required to attend, e.g., maths, others are either not taught in school or available only as an optional choice. This, thus, reduces the possibility for female students of familiarizing with the field when in school.

A third issue was introduced in Chapter 2 and marginally emerged in other chapters but was not deeply discussed. However, I think it is a theme worth mentioning before concluding this thesis. When discussing the limitations of current instruments measuring gender stereotypes, I mentioned the lack of a distinction between descriptive and prescriptive stereotypes. Descriptive stereotypes can be defined as the description of what group members are typically like – e.g., STEM is a masculine sector – while prescriptive stereotypes are the description of the behaviour group members should uphold to avoid derision – e.g., STEM should be a masculine sector (Gill, 2004). The difference between the two is relevant since the former would lead to an unintentional form of discrimination, which may be modified when information on the inaccuracy of the gender bias is provided, while the latter would lead to a stronger form of discrimination and prejudice, which is not dented by any information (Burgess and Borgida 1999).

If we claim that the observed gender gap in STEM derives from the erroneous belief that men are naturally inclined to science, maths and logical reasoning and we are interested in verifying whether this belief is still endorsed, it is necessary to be more specific when designing gender stereotypes' instruments. People's agreement with the statement 'In general, men may be better than women at math' does not allow to distinguish whether the individual agrees because of what he/she observes in the real world or because he/she thinks that one gender is by nature more talented than the other. As mentioned before, this distinction deserves to be made. As a matter of fact, data suggest that a gender gap do exists, thus we should not be surprised if people are aware of it, nor we should blame them for this. As stated by Nelson, earlier definitions of stereotypes also required 'inaccuracy, negativity, and overgeneralization. It is unfortunate that we have let those original requirements go – after all, they really are the heart of why we care about the topic at all.' (Nelson, 2009, p. 2)

It would, thus, be necessary, as done in some studies, to specify the cause to which we attribute the difference between women and men, e.g., 'Girls have fewer natural abilities than men for STEM issues'. The studies in this thesis partially contribute to this issue, but further research is needed to gain insights into the consequences of the endorsement of descriptive versus prescriptive gender stereotypes in STEM. As reported in the Appendix, a SEM analysis of the Project Implicit data suggests that answers to questions on the masculinity of science domains were associated with the attribution of the gender gap to social factors, e.g., discrimination, in the case of women and to personal characteristics, e.g., interests, in the case of men. Furthermore, the descriptive statistics of the data collected for the experimental study in Chapter 6 showed that half of the female sample considered social pressure as more relevant than biological characteristics in explaining the gender gap in STEM. Conversely, only 30% of the male sample shared this opinion. More should be done to understand how this difference in stereotypes' attribution is shaping young women's attitudes toward STEM.

References

AlmaLaurea. (2021). XXIII Indagine Profilo dei Laureati. AlmaLaurea. https://www.almalaurea.it/universita/indagini/laureati/profilo

Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. Psychology, Public Policy, and Law, 5(3), 665–692. https://doi.org/10.1037/1076-8971.5.3.665

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, 143(1), 1–35. https://doi.org/10.1037/bul0000052

Croft, A., Schmader, T., & Block, K. (2015). An Underexamined Inequality: Cultural and Psychological Barriers to Men's Engagement With Communal Roles. Personality and Social Psychology Review, 19(4), 343–370. https://doi.org/10.1177/1088868314564789

Gill, M. J. (2004). When information does not deter stereotyping: Prescriptive stereotyping can foster bias under conditions that deter descriptive stereotyping. Journal of Experimental Social Psychology, 40(5), 619–632. https://doi.org/10.1016/j.jesp.2003.12.001

Kanny, M. A., Sax, L. J., & Riggers-Pieh, T. A. (2014). Investigating forty years of stem research: How explanations for the gender gap have evolved over time. Journal of Women and Minorities in Science and Engineering, 20(2), 127–148. https://doi.org/10.1615/JWomenMinorScienEng.2014007246

Nelson, T. D. (Ed.). (2009). *Handbook of prejudice, stereotyping, and discrimination*. Psychology Press.

OECD. (2020). Girls' and boys' performance in PISA. In PISA 2018 Results (Volume II): Where All Students Can Succeed.

Simpson, R. (2005). Men in Non-Traditional Occupations: Career Entry, Career Orientation and Experience of Role Strain. Gender, Work & Organization, 12(4), 363–380. https://doi.org/10.1111/j.1468-0432.2005.00278.x

Wang, M.-T., & Degol, J. L. (2017). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future

Directions. Educational Psychology Review, 29(1), 119–140. https://doi.org/10.1007/s10648-015-9355-x

Wang, M.-T., Eccles, J. S., & Kenny, S. (2013). Not Lack of Ability but More Choice: Individual and Gender Differences in Choice of Careers in Science, Technology, Engineering, and Mathematics. *Psychological Science*, 24(5), 770–775. https://doi.org/10.1177/0956797612458937

Zosuls, K. M., Miller, C. F., Ruble, D. N., Martin, C. L., & Fabes, R. A. (2011). Gender Development Research in Sex Roles: Historical Trends and Future Directions. Sex Roles, 64(11–12), 826–842. https://doi.org/10.1007/s11199-010-9902-3

Appendix

Chapter 2 Supplementary material

Summary of measures of gender-science stereotypes

Table 2.3 *List of explicit instruments*

Scale type	Instrument	Construct	Reference	Reliability
11-point Likert scale	Are men or women more suited to - Science - Maths -Physics - Chemistry - Computing - Engineering	*Suitability*	(Farrell et al., 2015)	Items analyzed separately
			(Farrell & McHugh, 2017)	Not reported
	Are men or women more suited to - Arts - English - Drama - Music - French - History		(Farrell et al., 2020)	Time 1: $\alpha_{\text{STEM}}=.81$; $\alpha_{\text{Arts}}=.69$ Time 2: $\alpha_{\text{STEM}}=.82$ $\alpha_{\text{Arts}}=.68$
Visual analog scale	- I think that in Math boys (girls) do this well... (from not well at all to very well) - I think that boys (girls) find Math... (from very hard to very easy) - I think that in language arts boys (girls) do this well... (from not well at all to very well) - I think that boys (girls) find language arts... (from very hard to very easy)	*Skills*	(Vuletich et al., 2020)	$.53<r<.61$

7-point Likert scale	- Please rate how much you associate mathematics with males or females - Please rate how much you associate language with males or females	*Gendered domain*	(del Rio et al., 2019) (del Rio et al., 2020)	Not reported Not reported
4-point Likert scale	To what extent do you agree with the following statements: - I doubt that a woman could excel in computing courses - Men are more capable than women at solving computing problems - Computing is an appropriate subject for both men and women to study - It is not appropriate for men to study computing - Women are more capable than men at solving computing problems - Women are more likely to excel in careers that involve computing than men are - Women produce higher quality work in computing than men - Women and men can both excel in careers that involve computing - I doubt that a man could excel in computing courses - It is not appropriate for women to study computing - Men produce higher quality work in computing than women - Men are more likely to excel in careers that involve computing than women are - Women produce the same quality work in computing as men - Men and women are equally capable of solving computing problems - Men and women can both excel in computing courses	*Skills* *Suitability*	(Forssen et al., 2011)	α = .85

Visual analog scale	- I think that in math boys (girls) do this well... (from not well at all to very well) - I think that in science boys (girls) do this well... (from not well at all to very well) - I think that in reading boys (girls) do this well... (from not well at all to very well) - I think that in writing boys (girls) do this well... (from not well at all to very well)	*Skills*	(Swinton, 2012)	α = .76 for boys' perception of boys' English competence α=.67 for girls' perception of girls' English competence
7-point Likert scale	- Even when possessing the same education and credentials, men's brains probably make them better at programming - Even when possessing the same education, men seem like they are naturally better at jobs requiring technical skills and analysis - Men's dispositions probably make them better in careers technology careers	*Skills*	(Martin & Phillips, 2019)	α = .91
5-point Likert scale 7-point Likert scale	To what extent do you believe the following traits are more or less characteristic of the average man or woman in society (1=more characteristic of women, 5=more characteristic of men): - Analytic - Mathematical - Good with numbers - Good at reasoning *Stereotype endorsement = average male score – average female score	*Skills*	(Martin & Phillips, 2019)	Study 1 α = .77 Study 3a: $\alpha_{\text{male}} = .88$ $\alpha_{\text{female}} = .85$ Study 3b: $\alpha_{\text{male}} = .82$ $\alpha_{\text{female}} = .90$
7-point Likert scale	To what extent you think that in general each trait describes Mark (Karen): - Analytic - Mathematical - Good with numbers - Good at reasoning	*Skills*	(Martin & Phillips, 2019)	Study 3a: $\alpha_{\text{Mark}} = .75$ $\alpha_{\text{Karen}} = .61$

	*Stereotype endorsement = average Mark score – average Karen score			
5-point Likert scale	E.g., “Boys (girls) are good at math problem solving” E.g., “Boys (girls) are good at language”	*Skills*	(Zhao et al., 2018)	α_{math} = .80 α_{lang} = .80
7-point Likert scale	- It is possible that men have more math ability than do women - In general, men may be better than women at math - I don’t think that there are any real gender differences in math ability	*Skills*	(Schmader et al., 2004)	α= .88
			(Kiefer & Sekaquaptewa, 2007a)	α= .93
			(Ramsey & Sekaquaptewa, 2011)	1st wave: α = .87 2nd wave: α = .71
			(Ramsey et al., 2013)	α_{study1} = .79 α_{study2} = .80
			(Cundiff et al., 2013) *Only 2 items	α = .93
			(Wille et al., 2018) *”boys” and “girls” instead of “men” and “women”	α = .93
			(Van Camp et al., 2019)	α_{time1} = .88 α_{time2} = .90
5-point Likert scale	E.g., “How likely are most boys (girls) to be good at math?” E.g., “How likely are most boys (girls) to be good at spelling?”	*Skills*	(Heyman & Legare, 2004)	Not reported

3-point Likert scale	(mostly boys=0, mostly girls=1, about the same=.5) E.g., "Who is best at math?" E.g., "Who is best at spelling?"	*Skills*	(Heyman & Legare, 2004)	Not reported
Rate	Here are some words that describe how children do in mathematics. Please put a check mark in the grey box (column) to show who each word belongs to. The word can belong to just boys, just girls, both boys and girls, or no one. - Positive adjectives: pass, perfect, good, fast, right, clever, brilliant, smart, excellent - Negative adjectives: weak, poor, hard, fail, mistakes, wrong, dumb, slow, bad, stupid *Separate male and female gender-mathematics stereotype computed as the difference between positive and negative traits Here are some words that describe how children do in reading. Please put a check mark in the grey box (column) to show who each word belongs to. The word can belong to just boys, just girls, both boys and girls, or no one. - Positive adjectives: pass, perfect, good, fast, right, clever, brilliant, smart, excellent - Negative adjectives: weak, poor, hard, fail, mistakes, wrong, dumb, slow, bad, stupid *Separate male and female gender-reading stereotype computed as the difference between positive and negative traits	*Skills*	(Nowicki & Lopata, 2017)	$\alpha_{math} = .78$ $\alpha_{reading} = .78$

5-point Likert scale	- Girls (Boys) are more suited than boys (girls) to work in engineering branches. - Girls (Boys) are more willing than boys (girls) to work in mathematically-related areas - Girls (Boys) are more successful than boys (girls) in predicting how to solve mathematical problems - Girls (Boys) are more likely than boys (girls) to believe they can be successful in mathematics. - Girls (Boys) like solving mathematics problems that their classmates are not able to more than boys (girls) do. - Girls (Boys) are more successful than boys (girls) in describing the situation given in mathematical problems with mathematical symbols - Girls (Boys) use mathematical tools such as rulers, number blocks etc., more effectively than boys (girls) do. - Girls (Boys) are more successful than boys (girls) in using a calculator in mathematics classes. - Girls (Boys) have higher mathematical thinking abilities than boys (girls) have. - Girls (Boys) are more successful than boys (girls) in modelling mathematical relationships by drawings. - Compared to boys (girls), girls (boys) mostly increase their mathematical achievement, because of the support of their teachers. - Compared to boys (girls), girls (boys) mostly increase their mathematics scores when the examination is too easy - Compared to boys (girls), girls (boys) mostly increase their mathematics scores because their	*Skills* *Interest* *Suitability* *Attribution*	(Nurlu, 2017)	α_{male}=.88 α_{female}=.91 *Computed on a set including other items

	parents provide them with mathematical support.			
Rate	Among boys, girls or both girls and boys: Who do you think is usually good at - Science - Engineering - Technology - Math Who do you think should be good at - Science - Engineering - Technology - Math Who do you think can be good at - Science - Engineering - Technology - Math	*Skills*	(McGuire et al., 2020)	Performed PCA
5-point Likert scale	Please rate how much you associate the following domains with males or females: - ... Mathematics courses (not specified in the text) - ... Language courses (not specified in the text) *The final scale is the difference between the math scale and the language scale	*Gendered domain*	(Morrissey et al., 2019)	Not reported
9-point Likert scale	Overall, do you think boys or girls are better at - English - Mathematics - Science - Computing Who works in jobs that use a lot of - English - Mathematics - Science - Computing	*Skills*	(Pennington et al., 2021)	Items analyzed separately

10-point Likert scale	Who is better suited for science?	*Suitability*	(Sanchis-Segura et al., 2018)	
5-point Likert scale	- Males are not naturally better than females in math. - It's hard to believe a female could be a genius in mathematics. - When a woman has to solve a math problem, she should ask a man for help. - Women can do just as well as men in math. - I would have more faith in the answer for a math problem solved by a man than a woman. - Women who enjoy studying math are a little strange. - Females are as good as males in geometry. - Women certainly are smart enough to do well in math. - I would expect a woman mathematician to be a forceful type of person. - Studying math is just as good for women as for men. - I would trust a female just as much as I would trust a male to solve important math problems.	*Skills Conformance*	(Doepken et al., 2004) (Song et al., 2016) (Song et al., 2017)	α = .92 α = .82 α = .82
5-point Likert scale	Math is a male domain	*Gendered domain*	(Bieg et al., 2015)	
5-point Likert scale	I think that in ... (1="boys are much better than girls", 3="boys and girls are the same", 5="girls are much better than boys") - Reading and writing - Math - Science	*Skills*	(Kurtz-Costes et al., 2014)	Items analyzed separately

7-point Likert scale	Rate to what extent the following majors are masculine or feminine - Mathematics - Physics - Chemistry - Computer science - Engineering - French literature - Cultural studies - English literature - Fine Arts - Education	*Gendered domain*	(Abouchedid & Nasser, 2000)	Items analzyed separately
5-point Likert scale	(1=boys definitely more likely than girls, 2=boys probably more likely than girls, 3=no difference between boys and girls, 4=girls probably more likely than boys, 5=girls definitely more likely than boys) *Who and Mathematics scale* - Mathematics is their favourite subject - Think it is important to understand the work in mathematics - Are asked more questions by the mathematics teacher - Give up when they find a mathematics problem too difficult - Have to work hard in mathematics to do well - Enjoy mathematics - Care about doing well in mathematics - Think they did not work hard enough if they do not do well in mathematics - Parents would be disappointed if they did not do well in mathematics - Need mathematics to maximise future employment opportunities - Like challenging mathematics problems - Are encouraged to do well by the mathematics teacher - Mathematics teachers think	*Skills* *Interest* *Attrribution* *Relevance*	(Leder & Forgasz, 2002; Forgasz et al., 2004; Forgasz & Mittelberg, 2008)	Not reported In Forgasz et al. (2004) validity and reliability were tested (but not reported)

	they will do well - Think mathematics will be important in their adult life - Expect to do well in mathematics - Distract other students from their mathematics work - Get the wrong answers in mathematics - Find mathematics easy - Parents think it is important for them to study mathematics - Need more help in mathematics - Tease boys if they are good at mathematics - Worry if do not do well in mathematics - Are not good at mathematics - Like using computers to work on mathematics problems - Mathematics teachers spend more time with them - Consider mathematics to be boring - Find mathematics difficult - Get on with their work in class - Think mathematics is interesting - Tease girls if they are good at mathematics

5-point Likert scale	*Mathematics as a male domain (MD) scale* - Boys understand mathematics better than girls do - Mathematics is easier for men than its for women - Men are mathematically more intelligent than women - Career choices make the study of mathematics more important for boys than for girls - Boys have more use for mathematics than girls do when they leave school - Mathematics is liked more by boys than by girls - More boys than girls care about doing well at mathematics - Girls are less interested in mathematics than are boys - Boys are encouraged more than girls to do well in mathematics - Boys, more than girls, want to do well in mathematics to please their parents - There are more popular boys than popular girls who are good at mathematics - It is more acceptable for a man than a woman to be good at mathematics - Boys are more determined than girls to do well in mathematics - Compared to boys, girls do less work in mathematics classes - Boys, more than girls, like challenging mathematics problems - The mathematical task done in class suit boys more than they suit girls *Mathematics as a female domain(FD) scale* - Girls are more likely than boys to believe they are good at mathematics - Girls have more natural mathematical ability than do	*Skills* *Suitability* *Interest* *Conformance* *Attribution* *Relevance*	(Fennema & Sherman, 1977; Leder & Forgasz, 2002; Forgasz et al., 2004)	$\alpha_{MD} = .90$ $\alpha_{FD} = .89$ $\alpha_{ND} = .84$
			(Forgasz & Mittelberg, 2008)	Not reported
			(Plante et al., 2009)	Grade 6: $\alpha_{MD} = .83$ $\alpha_{FD} =. 88$ Grade 8: $\alpha_{MD} = .87$ $\alpha_{FD} = .91$ Grade 10: $\alpha_{MD} = .86$ $\alpha_{FD} = .91$

boys
- The weakest mathematics students are more often boys than girls
- When they leave school, girls will have more use for mathematics than boys will
- Girls are more suited than boys to a career in a mathematically-related area
- Girls, more than boys, care about doing well at mathematics
- Girls enjoy mathematics more than boys do
- Girls are more likely than boys to say mathematics is their favourite subject
- Girls are encouraged more than boys to do well in mathematics
- In a mathematics class with both boys and girls, girls tend to speak up more than boys
- Parents believe mathematics is more important for their daughters than for their sons
- Boys are distracted from their work in mathematics classes more than are girls
- Girls are more careful than boys when doing mathematics
- Compared to girls, boys give up more easily when they have difficulty with a mathematics problem
- Boys, more than girls, say the mathematics test was too hard if they do not do well
- Explaining answers in mathematics is harder for boys than for girls

Mathematics as a neutral (ND) domain
- Women are more equally likely to be good mathematics teacher
- Being good at mathematics comes as naturally to girls as to boys
- Men and women are equally suited to careers in the computer

industry
- It is just as difficult for girls as it is for boys to get a job in a mathematically-related profession
- Girls and boys are equally likely to believe that mathematics is important for their career
- Boys are just as likely as girls to enjoy mathematics
- Girls are just as likely as boys to say they want to excel in mathematics
- Students who say mathematics is their favourite subject are equally likely to be girls or boys
- Parents are as likely to help their daughters as their sons with mathematics
- Parents think that getting high grades in mathematics is as important for their daughters as for their sons
- Girls and boys who do well in a mathematics test are just as likely to be congratulated
- Boys are just as likely as girls to help friends with their mathematics
- Girls and boys are just as likely to be lazy in mathematics classes
- Girls are just as likely to work hard in mathematics as boys
- Students who get poor marks on mathematics tests are just as likely to be boys as girls
- Boys and girls are equally good at using calculators in mathematics

5-point Likert scale	*Mathematics as a male domain scale from the Fennema-Sherman Mathematics Attitudes Scales* - Females are as good as males in mathematics - Studying mathematics is just as appropriate for women as for men - I would trust a woman just as much as I would trust a man to figure out important calculations - Girls can do just as well as boys in mathematics - Males are not naturally better than females in mathematics - Women certainly are logical enough to do well in mathematics - It's hard to believe a female could be a genius in mathematics - When a woman has to solve a math problem, it is feminine to ask a man for help - I would have more faith in the answer for a math problem solved by a man than a woman - Girls who enjoy studying math are a bit peculiar - Mathematics is for men, arithmetic is for women - I would expect a woman mathematician to be a masculine type of person	*Skills* *Suitability* *Conformance* *Relevance*	(Fennema & Sherman, 1976)	α = .87
			(Fennema & Sherman, 1977)	Not reported
			(Wortel, 1997)	Not reported

5-point Likert scale	*Language as a male domain (MD) scale* - Boys understand language better than girls do - Language is easier for men than its for women - Men are more intelligent than women in language - Career choices make the study of language more important for boys than for girls - Boys have more use for language than girls do when they leave school - Language is liked more by boys than by girls - More boys than girls care about doing well at language - Girls are less interested in language than are boys - Boys are encouraged more than girls to do well in language - Boys, more than girls, want to do well in language to please their parents - There are more popular boys than popular girls who are good at language - It is more acceptable for a man than a woman to be good at language - Boys are more determined than girls to do well in language - Compared to boys, girls do less work in language classes - Boys, more than girls, like challenging language problems - The language task done in class suit boys more than they suit girls *Language as a female domain(FD) scale* - Girls are more likely than boys to believe they are good at language - Girls have more natural language ability than do boys	*Skills* *Suitability* *Interest* *Conformance* *Attribution*	(Plante et al., 2009)	Grade 6: $\alpha_{MD} = .84$ $\alpha_{FD} = .87$ Grade 8: $\alpha_{MD} = .85$ $\alpha_{FD} = .88$ Grade 10: $\alpha_{MD} = .89$ $\alpha_{FD} = .89$

	- The weakest language students are more often boys than girls - When they leave school, girls will have more use for language than boys will - Girls are more suited than boys to a career in a language-related area - Girls, more than boys, care about doing well at language - Girls enjoy language more than boys do - Girls are more likely than boys to say language is their favourite subject - Girls are encouraged more than boys to do well in language - In a language class with both boys and girls, girls tend to speak up more than boys - Parents believe language is more important for their daughters than for their sons - Boys are distracted from their work in language classes more than are girls - Girls are more careful than boys when doing language - Compared to girls, boys give up more easily when they have difficulty with a language problem - Boys, more than girls, say the language test was too hard if they do not do well - Explaining answers in language is harder for boys than for girls			
5-point Likert scale	- Usually, boys are more talented than girls at math - Usually, girls are better at arts and language than at math	*Skills*	(Tomasetto et al., 2011)	$r_{\text{mothers}} = .62$ $r_{\text{fathers}} = .61$
5-point Likert scale	E.g., "In general, boys/men are better at math than girls/women	*Skills*	(Greene et al., 1999)	$\alpha = .87$

7-point Likert scale	- I believe to be true the stereotype regarding females as poorer in math than males. - Males are better at math than females. - The stereotype about females being poorer at math than males is not true. - I believe in the stereotype that females are not as capable as males in the math arena. - I do not believe the stereotype that females are not as capable as males in the math arena. - I endorse the stereotype that females are not as capable as males in the math arena.	*Skills*	(Fabert, 2015)	α = .87
Analog feeling thermometer	How masculine/man-oriented or feminine/woman-oriented do you view STEM?	*Gendered domain*	(Starr, 2018)	
5-point Likert scale	- Men are just better at science than women - If I were having trouble with a math problem, I would go to a man instead of a woman To what extent genders differ in skill at (1= Men much better, 5= women much better): - sciences - humanities	*Skills*	(Lane et al., 2012)	α = .72
4-point Likert scale	- Men are better at math than women are - Women can achieve as much as men in math	*Skills*	(Nosek & Smyth, 2011)	r = .60
5-point Likert scale	- Boys (Girls) are often talented for doing math - Boys (Girls) are often talented for doing German	*Skills*	(Nurnberger et al., 2016)	r = .47
5-point Likert scale	- Boys (Girls) are often talented for doing math - Math is rather a typical subject for boys (Girls) - Boys (Girls) are often talented for doing German - German is rather a typical subject for boys (Girls)	*Skills* *Gendered domain*	(Steffens et al., 2010; Steffens & Jelenec, 2011)	Not reported

	*math-boys stereotype=math items on girls – math items on boys language-girls stereotype=language items on girls – language items on boys			
5-point Likert scale	Female students with the same math grade of males are less likely to attend a scientific track during high school. According with your experience, how much can these factors influence the choice of females toward alternative tracks? - Low interest for scientific subjects - Low inclination for scientific subjects - Low self-esteem - Encouragement of the family toward alternative paths - Influence of gender predicament ("women are bad at math")	*Attribution*	(Carlana, 2019)	Not reported
5-point Likert scale	There are innate biological differences in math abilities of women and men	*Skills*	(Carlana, 2019)	
7-point Likert scale	How much you associate Science with males or females? How much you associate Liberal Arts with males or females?	*Gendered domain*	(Nosek et al., 1998; Nosek et al., 2009; Zitelny et al., 2017) (Cai et al., 2016)	 $r_{test} = .4$ $r_{retest} = .56$
5-point Likert scale	How many men and women usually do this work (1= almost all men, 3=about equal men and women, 5=almost all women) - computer software programmer - environmental engineer - bio-mechanical engineer - computer designer - chemical engineer - forensic chemist - engineer for a relief agency - astronomer	*Gender imbalance*	(Barth et al., 2018)	α = 0.69

5-point Likert scale	Please rate how much you associate the following domains with males or females: - Math - Arts *Explicit measure is the difference in self-reported belief that math is male versus math is female (compared with arts)	*Gendered domain*	(Nosek et al., 2002b)	
5-point Likert scale	From -2=definitely girls to 2=definitely boys - Who is better in solving the mental rotation items? - Who is better in mathematics? - Who is better in first language?	*Skills*	(Moe, 2018)	Items analyzed separately
Rate	- These are a boy and a girl. They are 6 years old and they are good at school. Is the boy better at math, the girl better at math, or are they the same? - These are a boy and a girl. They are 6 years old and they are good at school. Is the boy better at language, the girl better at language, or are they the same?	*Skills*	(Galdi et al., 2014, 2017)	Not reported
9-point Likert scale	- I believe that generally males are more talented than females at math - I believe that generally females are more talented than males at language - I believe that generally males have more difficulty with math than with language - I believe that generally females have more facility with language than with math	*Skills*	(Galdi et al., 2017)	Not reported
9-point Likert scale	- I believe that generally males are more talented than females at math - I believe that generally females are more talented than males at language - I believe that generally females are more talented at language than at math - I believe that generally males	*Skills*	(Tomasetto et al., 2015)	α = 0.96

	are more talented at math than at language			
Likert scale	*Gender Ideology scale* - Women have the same technical skills as men - In the ICT field, a man’s performance will be better than a woman’s - Women are capable of developing useful software - Women and men have equal employment opportunities in ICT careers - Boys prefer STEM-related hobbies - There are more boys than girls in the STEM studies as they are more freaks - Women working in STEM areas have to be/act like men - To have a successful career in STEM you need to think and act like a man - Girls are not as good as boys in STEM issues - Girls are not as interested as boys in STEM issues - STEM themes are more masculine than others - Girls have fewer natural abilities than men for STEM issues - Most girls are better at other things (such as letters/languages) and choose studies in which they are better - Career in STEM are not associated with the traditional role of women - University studies in STEM are generally more attractive to boys	*Skills* *Gendered domain* *Suitability* *Interest* *Conformance*	(Verdugo-Castro et al., 2020)	Not reported

Rate	This is a boy (girl): - Which one do you think he (she) likes better: the language worksheet or the math worksheet? - In which of these does he (she) get better grades, in the language or in the math homework? - Do you believe he (she) is better at the language or the math homework? - Which one do you think he (she) likes less: language or math? - Which homework do you think is harder for him (her): math or language? - In which one do you think he (she) gets worse grades: language or math? * Two scales, one on male math affinity and the other on female math affinity	*Skills* *Interest*	(del Río & Strasser, 2013)	$\alpha_{male} = .88$ $\alpha_{female} = .79$
5-point Likert scale	To what extent the following subjects are generally more suitable for girls or for boys? (1=more suitable for girls, 5=more suitable for boys) - Nature - Geography - Maths - Technical Education	*Suitability*	(Blazev et al., 2017)	Not reported
7-point Likert scale	Males perform better in science because of greater natural ability	*Skills*	(Zitelny et al., 2017)	
6-point Likert scale	- In general, it is true that men are smarter in math than women - Suppose there is a team competition to see which team is smarter in math. I would prefer to have men on my team, not women - Of all the people who are very smart in math, most of these people are men - Of all the people who are NOT smart in math, most of these people are women	*Skills*	(Good, 2001)	$\alpha = .94$

5-Likert scale	Women hold a smaller portion of the science and engineering faculty positions at top research universities than do men. The following factors are sometimes offered as reasons for this difference. Please rate how important you think each factor is for explaining this difference - Different proportions of men and women are found among people with the very highest levels of math ability - On average, men and women differ in their willingness to devote the time required by such high-powered positions - On average, men and women differ naturally in their scientific interest - On average, men and women differ in their willingness to spend away from their families - Directly or indirectly, boys and girls tend to receive different levels of encouragement for developing scientific interest - On average, whether consciously or unconsciously, men are favoured in hiring and promotion	*Attribution*	(Nosek et al., 1998; Zitelny et al., 2017)	Not reported
7-point Likert scale	- Girls better at English - Boys better at Science - Girls better at Maths - Boys better at Art	*Skills*	(Lawrie & Brown, 1992)	Items analyzed separately
5-point Likert scale	- Please rate how much you associate "Math/Science" with males or females? (1=strongly male, 5=strongly female) - Please rate how much you associate "Liberal Arts" with males or females? (1=strongly male, 5=strongly female) *Final scale as a combination of the two items	*Gendered domain*	(Rentas, 2015)	Not reported

5-point Likert scale	- Men are naturally betterat advanced math than women - Men are naturally better at mechanical things than women - Men find math more useful than women	*Skills Relevance*	(Riegle-Crumb et al., 2017)	α = .90
5-point Likert scale	- Boys are able to do practical things better than girls - Boys know more about technology than girls do this - Boys are more capable of doing technological jobs than girls	*Skills*	(Boeve-de Pauw et al., 2020) (Ardies et al., 2013)	α = .84 α = .80
7-point Likert scale	Men are better in mathematics than women and therefore, usually do better in mathematics classes	*Skills*	(Kapitanoff & Pandey, 2017)	
5-point Likert scale	Stereotypes about interests (7 items): E.g., “Girls show less interest in STEM subjects than boys”	*Interest*	(Ertl et al., 2017)	α = .73
5-point Likert scale	Stereotypes about ability (5 items): E.g., “Girls have lower skills in STEM subjects than boys”	*Skills*	(Ertl et al., 2017)	α = .70
5-point Likert scale	Stereotypes about conformance (2 items): E.g., “Females that are working in the field of STEM have to be like men”	*Conformance*	(Ertl et al., 2017)	r = .77
Rate	- How would you compare males’ and females’ ability in physics? (male is better, female is better, both are equally good) - How would you compare males’ and females’ ability in biology? (male is better, female is better, both are equally good)	*Skills*	(Lerdpornkulrat et al., 2012)	Items analyzed separately
7-point Likert scale	Rate on a scale ranging from 1 (masculine) to 7 (feminine): - Science - Chemistry - Engineering - Physics	*Gendered domain*	(Young et al., 2013)	α = .77

Rate	In your opinion, in this kind of task [mental rotation task performed by participants] 1. Women score higher than men 2. Men score higher than women 3. Men and women score equally	*Skills*	(Guizzo et al., 2019)	
5-point Likert scale	Does the characteristic or behaviour described apply to all students or is it more typical for boys or girls (1=mainly to girls, 5=mainly to boys): - They have an aptitude in whatever concerns computers and ICTs. - They need to try hard to learn things about computers and ICTs. - Studying information or computer science is in agreement with their general interests. - Studying information or computer science is in agreement with their personality. - They discover easily new things about ICTs. - They are interested in the computers and the ICTs in general. - They can help someone older to solve simple problems that can emerge during the use of ICTs. - It is important for their future career to have enough knowledge about ICTs.	*Skills* *Interest* *Relevance*	(Vekiri, 2013)	α = .81
Rate	After being shown two photographs, one of a boy and the other of a girl, participants were asked: - Who is better in math, the boy, the girl, or they are equally good? - Who is better in Italian, the boy, the girl, or they are equally good?	*Skills*	(Tomasetto et al., 2012)	
10-point Likert scale	(+5=masculine, 0=neutral, -5=feminine) - Technology high school suggests.....to me - Natural and Health Science high school suggests.....to me	*Gendered domain*	(Lopez-Saez et al., 2011)	Not reported

	- Humanities and Social Science high school suggests.....to me			
7-point Likert scale	To what extent are the following occupations masculine or feminine: - Engineer - Elementary school teacher	*Gendered domain*	(Heyden et al., 2016; White & White, 2006)	Items analyzed separately
3-point Likert scale	Who works at jobs that use a lot of... (1=Mostly men, 2=Men and women, 3=Mostly women): - Math - Life science - Physical science - Reading	*Gender imbalance*	(Andre et al., 1999)	Not reported
5-point Likert scale	Who works at jobs that use a lot of... (1=Mostly women, 3=Equal numbers, 5=Mostly men) - Math - Life science - Physical science - Computer skills - Reading - Social sciences - Language arts - Arts	*Gender imbalance*	(Andre et al., 1999)	Not reported
3-point Likert scale (younger students) 5-point Likert scale (older students)	Are boys better than girls at mathematics, are girls better than boys, or are boys and girls equally good?	*Skills*	(Ambady et al., 2001; Muzzatti & Agnoli, 2007)	Not reported
5-point Likert scale	Please rate how much you associate the following domains with males or females - Science - Liberal Arts	*Gendered domain*	(Greenwald et al., 2003; Miller et al., 2015; Smyth & Nosek, 2015)	Not reported

5-point Likert scale	- Males are much more talented in math than females vs. females are much more talented in math than males - Mathematics is rather a boys' discipline vs. mathematics is rather a girls' discipline - Mathematics is more important for boys' later occupation vs. mathematics is more important for girls' later occupation - Males are much better in logical thinking ability than females vs. females are much more better in logical thinking ability	*Skills* *Gendered domain* *Relevance*	(Tiedemann, 2002)	α = .72
11-point Likert scale	- How much truth is there to the stereotype that "Men typically have better math skills than women"? - How much truth is there to the stereotype that "Men typically have better spatial skills than women"?	*Skills*	(Blanton et al., 2002)	r = .87
5-point Likert scale	- Technology is as difficult for boys as it is for girls - A girl can very well have a technological job - A girl can become a car mechanic - Boys are able to do practical things better than girls - Girls are able to operate a computer - Boys are more capable of doing technological jobs than girls - More girls should work in technology	*Skills*	(Bame et al., 1993)	Not reported
7-point Likert scale	Men outperform women on mathematics tests	*Skills*	(McIntyre et al., 2003)	Not applicable
5-point Likert scale	It is more important for boys to understand physical science than girls.	*Relevance*	(Buck et al., 2002)	
5-point Likert scale	Who do you think works in the jobs that involve physical science? (1=Almost all men, 3=Equal numbers of men and women, 5=Almost all women)	*Gender imbalance*	(Buck et al., 2002)	Not reported

4-point Likert scale	Among the following statements, which seem true to you, and which seem false? [Possible answers: True/Somewhat true/Somewhat false/False] - There are more men that women in science-related jobs - Men are more gifted than women in mathematics - Women and men are born with different brains - Women don't really like science - Women face discrimination in science-related jobs	*Skills* *Gender imbalance* *Interest* *Attribution*	(Breda et al., 2020)	Not reported
5-point Likert scale	- When I think of people who are very good at math and science, I am more likely to think of men than women - When I think of people who are scientists/mathematicians, I am more likely to think of men than women - When it comes to math and science, I believe that on average, men are better at math and science than women	*Skills* *Gender imbalance*	(Scott & Martin, 2014)	α = .63
7-point Likert scale	(1=Mostly men, 4=Equal number of men and women, 7=Mostly women) - When I think of people who are very good at engineering, I think of... - When I think of people who have careers in engineering, I think of…	*Skills* *Gender imbalance*	(Dasgupta et al., 2015)	r = 0.47
9-point Likert scale	To what extent each of the following characteristics apply to female versus male students at your university (1=It applies only to female students, 9=It applies only to male students): - math skills - spatial abilities - analytical thinking - interest in science - scientific thinking	*Skills* *Interest*	(Jasko et al., 2019)	α = .18 −.17<r<.26

	-artistic skills - ...			
3-point Likert scale	(1=girls are better than boys, 2=girls and boys are equally good, 3=boys are better than girls) - Who is better at science and math, girls or boys?	*Skills*	(Riegle-Crumb & Morton, 2017)	
Likert scale	To what extent you associate science with males or females? (From strongly male to strongly female) - To what extent you associate humanities with males or females *Explicit stereotypes=humanities item – science item	*Gendered domain*	(Liu et al., 2010)	Not reported
9-point Likert scale	- It is possible that men have more STEM ability than do women - In general, men may be better than women in STEM fields - I don't think that there are any real gender differences in STEM ability - Men typically have better STEM skills than women - Men typically have better spatial skills than women - Men are just naturally better at STEM compared to women - Men and women are equally good in STEM fields	*Skills*	(Gilbert, 2014)	α_{pre} = .88 α_{post} = .90
Rate 0-100	- How many women are “good at math”? - How many men were “good at math”? *Math-gender stereotype= women item - men item	*Skills*	(Kiefer & Sekaquaptewa, 2007b)	Not reported
9-point Likert scale	E.g., “I think that in general, men are better at math, science, and engineering than are women”	*Skills*	(Park et al., 2001)	Not reported
5-point Likert scale	(-2=Girls are definitely better, +2= Boys are definitely better) - In your opinion, who is better at Math between girls and boys?	*Skills*	(Passolunghi et al., 2014)	

7-point Likert scale	Do being good at math and being girly go together?	*Conformance*	(Betz & Sekaquaptewa, 2012)	
Rate (0-10)	Suppose that ten 9th-grade boys were picked at random from a typical U.S. high school. How many would you predict will complete a calculus course before finishing high school? Suppose that ten 9th-grade girls were picked at random from a typical U.S. high school. How many would you predict will complete a calculus course before finishing high school?	*Skills*	(Nosek et al., 1998)	
11-point Likert scales	E.g., "When I think of people who are very good at math, I think of…" E.g., "When I think of people who are very good at English, I think of…"	*Skills*	(Stout et al., 2011)	$\alpha_{math} = .74$ $\alpha_{English} = .80$
Rate	Select the proportion of women among United States: - engineers (1, 4, 7, 13, 16, or 19%) - computer systems analysts (9, 16, 23, 37, 44, or 51%)	*Gender imbalance*	(Ramsey et al., 2013)	$r_{study1}=.27$, $p = .03$ $r_{study2} = .46$, $p = .001$
Choose 1 answer	Which statement best describes you? - I strongly associate liberal arts with females and science with males - I moderately associate liberal arts with females and science with males - I associate males and females with science and liberal arts equally - I moderately associate science with females and liberal arts with males - I strongly associate science with females and liberal arts with males	*Gendered domain*	(Greenwald et al., 2003)	

Choose 1 answer	For each question, children were shown two pictures of a child (a girl and a boy) and responded to the following questions: - Which one (the boy or the girl) like to do math more? And to what extent (a little more or a lot more) - Which one (the boy or the girl) like reading more? And to what extent (a little more or a lot more) *The math-gender stereotype scale is the difference between the two items	*Skills*	(Cvencek et al., 2011)	$\alpha = .03$
			(Cvencek et al., 2014)	$\alpha = .31$
			(Cvencek et al., 2015)	$r = .11$
			(Paz-Albo Prieto et al., 2017)	
			(del Rio et al., 2019, 2020)	
5-point Likert scale	E.g., “Women are worse at math than men”		(Jackson et al., 2014)	$\alpha_{\text{pre-test}} = .83$ $\alpha_{\text{post-test}} = .87$

Table 2.4. *List of implicit instruments*

Test	Target and categories	Reference
AMP	Target: men/women (photos) Categories: Good/bad at math Target: men/women (photos) Categories: Good/bad at language arts	(Vuletich et al. 2020)
IRAP	Target: "women more suited to" Categories: Arts /STEM subjects Target: "men more suited to" Categories: Arts /STEM subjects	(Farrell, Cochrane, and McHugh 2015; Farrell, Nearchou, and McHugh 2020)
IRAP	Target: men/women (photos) Categories: science-/arts-related careers Response option: similar/opposite	(Fleming, Foody, and Murphy 2020)
GNAT	Target: men/women Categories: Science-related majors Target: men/women Categories: Humanities-related majors	(Gilbert 2014; Van Camp, Gilbert, and O'Brien 2019)
GNAT	Target: men/women Categories: math-related features Target: men/women Categories: English-related features	(Gilbert et al. 2015)
GNAT	Target: men/women Categories: math-related features Target: men/women Categories: language-related features	(Steffens, Jelenec, and Noack 2010; Steffens and Jelenec 2011)
SPF	Target: men/women (photos) Categories: math/language (words)	(Nurnberger et al. 2016)
(Child) IAT	Target: boys/girls (pictures) Categories: math-related or reading-related objects (pictures)	(Tomasetto, Galdi, and Cadinu 2012; Galdi, Cadinu, and Tomasetto 2014; Galdi, Mirisola, and Tomasetto 2017; del Río et al. 2019, 2020)
(Child) IAT	Target: female/male dolls (picture) Categories: negative/positive adjectives about ability in math	(Nowicki and Lopata 2017)

	Target: female/male dolls (picture) Categories: negative/positive adjectives about ability in reading	
SA-IAT	Target: men/women Category: Space-related concepts	(Guizzo et al. 2019)
IAT	Target: men/women Categories: STEM/liberal arts majors or academic fields	(Nosek et al., 1998; Nosek et al., 2002a; Greenwald et al., 2003; Nosek et al., 2009; Sriram & Greenwald, 2009; Liu et al., 2010; Stamm, 2010; Nosek & Smyth, 2011; Ramsey & Sekaquaptewa, 2011; Lane et al., 2012; Ramsey et al., 2013; Young et al., 2013; Farrell et al., 2015; Smyth & Nosek, 2015; Rentas, 2015; Miller et al., 2015; Cai et al., 2016; Zitelny et al., 2017; Starr, 2018; Sanchis-Segura et al., 2018; Martin & Phillips, 2019; Carlana, 2019)
IAT	Target: men/women Categories science/people-oriented careers	(Dunlap and Barth 2019)
IAT	Target: men/women Categories: Math-/reading-related words	(Stout et al. 2011; Cvencek, Meltzoff, and Greenwald 2011; Cvencek, Meltzoff, and Kapur 2014; Cvencek, Kapur, and Meltzoff 2015; Galdi et al. 2017)
IAT	Target: men/women Categories: Math-/language-related concepts	(Nosek et al., 2002b; Smeding, 2012; Passolunghi et al., 2014; Morrissey et al., 2019)
IAT	Target: men/women Categories: Math-/humanities-related concepts	(Park et al., 2001; Nosek et al., 2002b; Kiefer & Sekaquaptewa, 2007a)
IAT	Target: men/women Categories: Science-/arts-related concepts	(Nosek et al., 2002b)
IAT	Target: men/women Categories: STEM/liberal arts features	(Cundiff et al. 2013)
IAT	Target: boys/girls Categories: spatial/language words	(Heyden et al. 2016)
IAT	Target: men/women	(White and White 2006)

	Categories: Engineer/Elementary school teacher	
IAT	Target: men/women Categories: engineering-/English-related features	(Dasgupta, Scircle, and Hunsinger 2015)
IAT	Target: men/women (photos) Categories: Math/science-/liberal arts-related words	(Kiefer and Sekaquaptewa 2007; Reuben, Sapienza, and Zingales 2014)

Table 2.5 *List of indirect instruments*

Instrument	Reference
Participants were told: "There were many good students in my high school, but one of my classmates stood out from the rest. This student got As in every subject, but was especially good at math. This student could figure out problems that even the teachers couldn't solve. One time, the student entered a state math contest and got a perfect score!". The participant was then asked to repeat the story, and the experimenter noted whether she used the pronoun "he" or "she"	(Ambady et al. 2001)
After been shown 5 photos of boys and 5 photos of girls, participants were presented with questions such as "Which kid can figure out how to do the hardest math problems?" or "Which child is the best at math" and asked to select one photograph. Similarly, they were asked questions about spelling abilities.	(Heyman and Legare 2004)
Participants were presented with two pictures, one of a boy (Carlo) and the other of a girl (Francesca) and asked - Is Carlo better at math, Francesca better at math, or are they the same?	(Tomasetto, Alparone, and Cadinu 2011)
After a story about an island where inhabitants would not consider boys and girls equally good in school subjects, participants are asked whether, in their opinion, the inhabitants of the island considered boys or girls better in math and the same question for reading	(Tomasetto et al. 2012)
Among a list of eight characters, differing in age and gender: - Who can be the best student in reading? - Who can be the best student in math? - Who can do long studies in literature? - Who can do long studies in math? - Who can teach reading? - Who can teach math? - Who can use reading abilities on the job? - Who can be a mathematician? - Who is the most serious student in the class?	(Martinot, Bages, and Desert 2012)
After been shown two pairs of images, one depicting a woman with a lab coat, the other depicting a man with a lab coat, participants were asked "Please, look carefully at the following images. Is it	(Christidou, Bonoti, and Kontopoulou 2016)

more possible that one of these two people is a scientist, could they both be scientists, or none of them?”	
After a vignette showing a boy or a girl, participants were asked to rate on a 7-point scale ranging fromfully agree to fully disagree regarding the students’ interests and giftedness in math/science and language (four items) and on how well they judged each student’s fit to a math/science- or a language-oriented school E.g., - The student appears to be very interested in math and science - I assume the student is very talented in the field of languages	(Nurnberger et al. 2016)

References

Abouchedid, K., & Nasser, R. (2000). External and internal social barriers in stereotyping university majors. Current Research in Social Psychology, 5(9).

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype Susceptibility in Children: Effects of Identity Activation on Quantitative Performance. Psychological Science, 12(5), 385–390. https://doi.org/10.1111/1467-9280.00371

Andre, T., Whigham, M., Hendrickson, A., & Chambers, S. (1999). Competency beliefs, positive affect, and gender stereotypes of elementary students and their parents about science versus other school subjects. Journal of Research in Science Teaching, 36(6), 719–747. https://doi.org/10.1002/(SICI)1098-2736(199908)36:6<719::AID-TEA8>3.0.CO;2-R

Ardies, J., De Maeyer, S., & Gijbels, D. (2013). Reconstructing the Pupils Attitude Towards Technology-survey. Design and Technology Education: An International Journal, 18, 8–19. https://ojs.lboro.ac.uk/DATE/article/view/1796

Bame, E. A., Dugger, W. E., de Vries, M., & McBee, J. (1993). Pupils' attitude toward technology – PATT-USA. The Journal of Technology Studies, 19(1), 40–48.

Barth, J. M., Kim, H., Eno, C. A., & Guadagno, R. E. (2018). Matching abilities to careers for others and self: Do gender stereotypes matter to students in advanced math and science classes? Sex Roles: A Journal of Research, 79(1–2), 83–97. https://doi.org/10.1007/s11199-017-0857-5

Betz, D. E., & Sekaquaptewa, D. (2012). My Fair Physicist? Feminine Math and Science Role Models Demotivate Young Girls. Social Psychological and Personality Science, 3(6), 738–746. https://doi.org/10.1177/1948550612440735

Bieg, M., Goetz, T., Wolter, I., & Hall, N. C. (2015). Gender stereotype endorsement differentially predicts girls' and boys' trait-state discrepancy in math anxiety. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.01404

Blanton, H., Christie, C., & Dye, M. (2002). Social Identity versus Reference Frame Comparisons: The Moderating Role of Stereotype Endorsement. Journal of Experimental Social Psychology, 38(3), 253–267. https://doi.org/10.1006/jesp.2001.1510

Blazev, M., Karabegovic, M., Burusic, J., & Selimbegovic, L. (2017). Predicting gender-STEM stereotyped beliefs among boys and girls from prior school achievement and interest in STEM school subjects. Social Psychology of Education: An International Journal, 20(4), 831–847. https://doi.org/10.1007/s11218-017-9397-7

Boeve-de Pauw, J., Ardies, J., Hens, K., Wullemen, A., Van de Vyver, Y., Rydant, T., De Spiegeleer, L., & Verbraeken, H. (2020). Short and long term impact of a high-tech STEM intervention on pupils' attitudes towards technology. International Journal of Technology and Design Education. Scopus. https://doi.org/10.1007/s10798-020-09627-5

Breda, T., Grenet, J., Monnet, M., & Effenterre, C. V. (2020). Do Female Role Models Reduce the Gender Gap in Science? Evidence from French High Schools (IZA Discussion Papers No. 13163). Institute of Labor Economics (IZA). http://hdl.handle.net/10419/216475

Buck, G. A., Leslie-Pelecky, D., & Kirby, S. K. (2002). Bringing female scientists into the elementary classroom: Confronting the strength of elementary students' stereotypical images of scientists. Journal of Elementary Science Education, 14(2), 1–9. https://doi.org/10.1007/BF03173844

Cai, H., Luo, Y. L. L., Shi, Y., Liu, Y., & Yang, Z. (2016). Male = science, female = humanities: Both implicit and explicit gender-science stereotypes are heritable. Social Psychological and Personality Science, 7(5), 412–419. https://doi.org/10.1177/1948550615627367

Carlana, M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias*. The Quarterly Journal of Economics, 134(3), 1163–1224. https://doi.org/10.1093/qje/qjz008

Christidou, V., Bonoti, F., & Kontopoulou, A. (2016). American and Greek Children's Visual Images of Scientists: Enduring or Fading Stereotypes? Science and Education, 25(5–6), 497–522. Scopus. https://doi.org/10.1007/s11191-016-9832-8

Cundiff, J. L., Vescio, T. K., Loken, E., & Lo, L. (2013). Do gender-science stereotypes predict science identification and science career aspirations among undergraduate science majors? Social Psychology of Education: An International Journal, 16(4), 541–554. https://doi.org/10.1007/s11218-013-9232-8

Cvencek, D., Kapur, M., & Meltzoff, A. N. (2015). Math achievement, stereotypes, and math self-concepts among elementary-school students in Singapore. Learning and Instruction, 39, 1–10. https://doi.org/10.1016/j.learninstruc.2015.04.002

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math-Gender Stereotypes in Elementary School Children: Gender Stereotypes. Child Development, 82(3), 766–779. https://doi.org/10.1111/j.1467-8624.2010.01529.x

Cvencek, D., Meltzoff, A. N., & Kapur, M. (2014). Cognitive consistency and math-gender stereotypes in Singaporean children. Journal of Experimental Child Psychology, 117, 73–91. https://doi.org/10.1016/j.jecp.2013.07.018

Dasgupta, N., Scircle, M. M., & Hunsinger, M. (2015). Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. Proceedings of the National Academy of Sciences, 112(16), 4988–4993. https://doi.org/10.1073/pnas.1422822112

del Río, M. F., & Strasser, K. (2013). Preschool Children's Beliefs about Gender Differences in Academic Skills. Sex Roles, 68(3–4), 231–238. Scopus. https://doi.org/10.1007/s11199-012-0195-6

del Rio, M. F., Strasser, K., Cvencek, D., Susperreguy, M. I., & Meltzoff, A. N. (2019). Chilean kindergarten children's beliefs about mathematics: Family matters. Developmental Psychology, 55(4), 687–702. https://doi.org/10.1037/dev0000658

del Rio, M. F., Susperreguy, M. I., Strasser, K., Cvencek, D., Iturra, C., Gallardo, I., & Meltzoff, A. N. (2020). Early sources of children's math achievement in chile: The role of parental beliefs and feelings about math. Early Education and Development. https://doi.org/10.1080/10409289.2020.1799617

Doepken, D., Lawsky, E., & Padwa, L. (2004). Modified Fennema-Sherman Attitude Scales. http://www.pearweb.org/atis/tools/55

Dunlap, S. T., & Barth, J. M. (2019). Career stereotypes and identities: Implicit beliefs and major choice for college women and men in STEM and female-dominated fields. Sex Roles: A Journal of Research, 81(9–10), 548–560. https://doi.org/10.1007/s11199-019-1013-1

Ertl, B., Luttenberger, S., & Paechter, M. (2017). The Impact of Gender Stereotypes on the Self-Concept of Female Students in STEM Subjects with an Under-Representation of Females. Frontiers in Psychology, 8, 703. https://doi.org/10.3389/fpsyg.2017.00703

Fabert, N. S. (2015). Growth mindset training to increase women's self-efficacy in science and engineering: A randomized-controlled trial. Dissertation Abstracts International: Section B: The Sciences and Engineering, 75(12-B(E)).

Farrell, L., Cochrane, A., & McHugh, L. (2015). Exploring attitudes towards gender and science: The advantages of an IRAP approach versus the IAT. Journal of Contextual Behavioral Science, 4(2), 121–128. https://doi.org/10.1016/j.jcbs.2015.04.002

Farrell, L., & McHugh, L. (2017). Examining gender-STEM bias among STEM and non-STEM students using the Implicit Relational Assessment Procedure (IRAP). Journal of Contextual Behavioral Science, 6(1), 80–90. https://doi.org/10.1016/j.jcbs.2017.02.001

Farrell, L., Nearchou, F., & McHugh, L. (2020). Examining the effectiveness of brief interventions to strengthen a positive implicit relation between women and STEM across two timepoints. Social Psychology of Education, 23(5), 1203–1231. https://doi.org/10.1007/s11218-020-09576-w

Fennema, E., & Sherman, J. (1977). Sex-Related Differences in Mathematics Achievement, Spatial Visualization and Affective Factors. American Educational Research Journal, 14(1), 51–71. https://doi.org/10.3102/00028312014001051

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman Mathematics Attitudes Scales: Instruments Designed to Measure Attitudes Toward the Learning of Mathematics by females and Males. Journal for Research in Mathematics Education, 7(5), 324–326.

Fleming, K., Foody, M., & Murphy, C. (2020). Using the Implicit Relational Assessment Procedure (IRAP) to Examine Implicit Gender Stereotypes in Science, Technology, Engineering and Maths (STEM). The Psychological Record, 70(3), 459–469. https://doi.org/10.1007/s40732-020-00401-6

Forgasz, H. J., Leder, G. C., & Kloosterman, P. (2004). New Perspectives on the Gender Stereotyping of Mathematics. Mathematical Thinking and Learning, 6(4), 389–420. https://doi.org/10.1207/s15327833mtl0604_2

Forgasz, H. J., & Mittelberg, D. (2008). Israeli Jewish and Arab students' gendering of mathematics. ZDM - International Journal on Mathematics Education, 40(4), 545–558. Scopus. https://doi.org/10.1007/s11858-008-0139-3

Forssen, A. V., Moskal, B. M., & Harriger, A. R. (2011). Measuring the impact of a high school intervention on students' attitudes in information technology: Validation and use of an attitude survey. ASEE Annual Conference and Exposition, Conference Proceedings, Vancouver, BC. Scopus. https://peer.asee.org/measuring-the-impact-of-a-high-school-intervention-on-students-attitudes-in-information-technology-validation-and-use-of-an-attitude-survey.pdf

Galdi, S., Cadinu, M., & Tomasetto, C. (2014). The Roots of Stereotype Threat: When Automatic Associations Disrupt Girls' Math Performance. Child Development, 85(1), 250–263. https://doi.org/10.1111/cdev.12128

Galdi, S., Mirisola, A., & Tomasetto, C. (2017). On the relations between parents' and children's implicit and explicit academic gender stereotypes. Psicologia Sociale, 12(2), 215–238.

Gilbert, P. N. (2014). How Does Identification with STEM Role Models Impact Women's Implicit STEM Stereotypes and STEM Outcomes? [PhD Dissertation, Tulane University]. https://digitallibrary.tulane.edu/islandora/object/tulane%3A27945

Gilbert, P. N., O'Brien, L. T., Garcia, D. M., & Marx, D. M. (2015). Not the sum of its parts: Decomposing implicit academic stereotypes to understand sense of fit in math and English. Sex Roles: A Journal of Research, 72(1–2), 25–39. https://doi.org/10.1007/s11199-014-0428-y

Good, C. D. (2001). The development of stereotype threat and its relation to theories of intelligence: Effects on elementary school girls' mathematics achievement and task choices. Dissertation Abstracts International: Section B: The Sciences and Engineering, 62(3-B), 1641.

Greene, B. A., DeBacker, T. K., Ravindran, B., & Krows, A. J. (1999). Goals, values, and beliefs as predictors of achievement and effort in high school mathematics classes. Sex Roles: A Journal of Research, 40(5–6), 421–458. https://doi.org/10.1023/A:1018871610174

Greenwald, A. G., Nosek, B., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. Journal of Personality and Social Psychology, 85(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Guizzo, F., Moe, A., Cadinu, M., & Bertolli, C. (2019). The role of implicit gender spatial stereotyping in mental rotation performance. ACTA PSYCHOLOGICA, 194, 63–68. https://doi.org/10.1016/j.actpsy.2019.01.013

Heyden, K. M. V., van Atteveldt, N. M., Huizinga, M., & Jolles, J. (2016). Implicit and explicit gender beliefs in spatial ability: Stronger stereotyping in boys than girls. Frontiers in Psychology, 7. https://doi.org/10.3389/fpsyg.2016.01114

Heyman, G., & Legare, C. (2004). Children’s beliefs about gender differences in the academic and social domains. SEX ROLES, 50(3–4), 227–239. https://doi.org/10.1023/B:SERS.0000015554.12336.30

Jackson, S. M., Hillard, A. L., & Schneider, T. R. (2014). Using implicit bias training to improve attitudes toward women in STEM. Social Psychology of Education: An International Journal, 17(3), 419–438. https://doi.org/10.1007/s11218-014-9259-5

Jasko, K., Dukala, K., & Szastok, M. (2019). Focusing on gender similarities increases female students’ motivation to participate in STEM. Journal of Applied Social Psychology, 49(8), 473–487. https://doi.org/10.1111/jasp.12598

Kapitanoff, S., & Pandey, C. (2017). Stereotype threat, anxiety, instructor gender, and underperformance in women. Active Learning in Higher Education, 18(3), 213–229. https://doi.org/10.1177/1469787417715202

Kiefer, A. K., & Sekaquaptewa, D. (2007a). Implicit Stereotypes, Gender Identification, and Math-Related Outcomes: A Prospective Study of Female College Students. Psychological Science, 18(1), 13–18. https://doi.org/10.1111/j.1467-9280.2007.01841.x

Kiefer, A. K., & Sekaquaptewa, D. (2007b). Implicit stereotypes and women’s math performance: How implicit gender-math stereotypes influence women’s susceptibility to stereotype threat. Journal of Experimental Social Psychology, 43(5), 825–832. https://doi.org/10.1016/j.jesp.2006.08.004

Kurtz-Costes, B., Copping, K. E., Rowley, S. J., & Kinlaw, C. R. (2014). Gender and age differences in awareness and endorsement of gender stereotypes about academic abilities. European Journal of Psychology of Education, 29(4), 603–618. https://doi.org/10.1007/s10212-014-0216-7

Lane, K. A., Goh, J. X., & Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. Sex Roles: A Journal of Research, 66(3–4), 220–234. https://doi.org/10.1007/s11199-011-0036-z

Lawrie, L., & Brown, R. (1992). Sex stereotypes, school subject preferences and career aspirations as a function of single mixed-sex schooling and presence absence of an opposite sex sibling. British Journal of Educational Psychology, 62, 132–138. https://doi.org/10.1111/j.2044-8279.1992.tb01006.x

Leder, G., & Forgasz, H. (2002). Two new instruments to probe attitudes about gender and mathematics (p. 29). La Trobe University. https://www.academia.edu/20451640/Two_New_Instruments_To_Probe_Attitudes_about_Gender_and_Mathematics

Lerdpornkulrat, T., Koul, R., & Sujivorakul, C. (2012). The influence of ability beliefs and motivational orientation on the self-efficacy of high school science students in Thailand. Australian Journal of Education, 56(2), 163–181. Scopus. https://doi.org/10.1177/000494411205600205

Liu, M., Hu, W., Jiannong, S., & Adey, P. (2010). Gender Stereotyping and Affective Attitudes Towards Science in Chinese Secondary School Students. International Journal of Science Education, 32(3), 379–395. https://doi.org/10.1080/09500690802595847

Lopez-Saez, M., Puertas, S., & Sainz, M. (2011). Why don't girls choose technological studies? Adolescents' stereotypes and attitudes towards studies related to medicine or engineering. The Spanish Journal of Psychology, 14(1), 74–87. https://doi.org/10.5209/rev_SJOP.2011.v14.n1.6

Martin, A. E., & Phillips, K. W. (2019). Blind to bias: The benefits of gender-blindness for STEM stereotyping. Journal of Experimental Social Psychology, 82, 294–306. https://doi.org/10.1016/j.jesp.2018.11.003

Martinot, D., Bages, C., & Desert, M. (2012). French children's awareness of gender stereotypes about mathematics and reading: When girls improve their reputation in math. Sex Roles: A Journal of Research, 66(3–4), 210–219. https://doi.org/10.1007/s11199-011-0032-3

McGuire, L., Mulvey, K. L., Goff, E., Irvin, M. J., Winterbottom, M., Fields, G. E., Hartstone-Rose, A., & Rutland, A. (2020). STEM gender stereotypes from early childhood through adolescence at informal science centers. Journal of Applied Developmental Psychology, 67. https://doi.org/10.1016/j.appdev.2020.101109

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(1), 83–90. https://doi.org/10.1016/S0022-1031(02)00513-9

Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. Journal of Educational Psychology, 107(3), 631–644. https://doi.org/10.1037/edu0000005

Moe, A. (2018). Mental rotation and mathematics: Gender-stereotyped beliefs and relationships in primary school children. Learning and Individual Differences, 61, 172–180. https://doi.org/10.1016/j.lindif.2017.12.002

Morrissey, K., Hallett, D., Bakhtiar, A., & Fitzpatrick, C. (2019). Implicit math-gender stereotype present in adults but not in 8th grade. Journal of Adolescence, 74, 173–182. https://doi.org/10.1016/j.adolescence.2019.06.003

Muzzatti, B., & Agnoli, F. (2007). Gender and mathematics: Attitudes and stereotype threat susceptibility in Italian children. Developmental Psychology, 43(3), 747–759. https://doi.org/10.1037/0012-1649.43.3.747

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002a). Harvesting implicit group attitudes and beliefs from a demonstration web site. Group Dynamics: Theory, Research, and Practice, 6(1), 101–115. https://doi.org/10.1037/1089-2699.6.1.101

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002b). Math = male, me = female, therefore math = me. Journal of Personality and Social Psychology, 83(1), 44–59. https://doi.org/10.1037/0022-3514.83.1.44

Nosek, B. A., & Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. American Educational Research Journal, 48(5), 1125–1156. https://doi.org/10.3102/0002831211410683

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., … Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. PNAS Proceedings of the National Academy of Sciences of the United States of America, 106(26), 10593–10597. https://doi.org/10.1073/pnas.0809921106

Nosek, B., Banaji, M. R., & Greenwald, A. G. (1998). Project Implicit. https://implicit.harvard.edu/implicit/index.jsp

Nosek, B., & Smyth, F. L. (2011). Implicit Social Cognitions Predict Sex Differences in Math Engagement and Achievement. American Educational Research Journal, 48(5), 1125–1156. https://doi.org/10.3102/0002831211410683

Nowicki, E. A., & Lopata, J. (2017). Children’s implicit and explicit gender stereotypes about mathematics and reading ability. Social Psychology of Education: An International Journal, 20(2), 329–345. https://doi.org/10.1007/s11218-015-9313-y

Nurlu, Ö. (2017). Developing a teachers’ gender stereotype scale toward mathematics. International Electronic Journal of Elementary Education, 10(2), 287–299. Scopus. https://doi.org/10.26822/iejee.2017236124

Nurnberger, M., Nerb, J., Schmitz, F., Keller, J., & Sutterlin, S. (2016). Implicit gender stereotypes and essentialist beliefs predict preservice teachers’ tracking recommendations. Journal of Experimental Education, 84(1), 152–174. https://doi.org/10.1080/00220973.2015.1027807

Park, L. E., Cook, K. E., & Greenwald, A. G. (2001). Implicit Indicators of Women’s Persistence in Math, Science, and Engineering. Psi Chi Journal of Psychological Research, 145–152. https://doi.org/10.24839/1089-4136.JN6.4.145

Passolunghi, M. C., Rueda Ferreira, T. I., & Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. Learning and Individual Differences, 34, 70–76. https://doi.org/10.1016/j.lindif.2014.05.005

Paz-Albo Prieto, J., Cvencek, D., Llacer, C. V. H., Escobar, A. H., & Meltzoff, A. N. (2017). Preschoolers' mathematical play and colour preferences: A new window into the development of gendered beliefs about math. Early Child Development and Care, 187(8), 1273–1283. https://doi.org/10.1080/03004430.2017.1295234

Pennington, C., Kaye, L., Qureshi, A., & Heim, D. (2021). Do gender differences in academic attainment correspond with scholastic attitudes? An exploratory study in a UK secondary school. Journal of Applied Social Psychology, 51(1), 3–16. https://doi.org/10.1111/jasp.12711

Plante, I., Théorêt, M., & Favreau, O. E. (2009). Student gender stereotypes: Contrasting the perceived maleness and femaleness of mathematics and language. Educational Psychology, 29(4), 385–405. https://doi.org/10.1080/01443410902971500

Ramsey, L. R., Betz, D. E., & Sekaquaptewa, D. (2013). The effects of an academic environment intervention on science identification among women in STEM. Social Psychology of Education, 16(3), 377–397. https://doi.org/10.1007/s11218-013-9218-6

Ramsey, L. R., & Sekaquaptewa, D. (2011). Changing stereotypes, changing grades: A longitudinal study of stereotyping during a college math course. Social Psychology of Education: An International Journal, 14(3), 377–387. https://doi.org/10.1007/s11218-010-9150-y

Rentas, C. A. (2015). The effects of a perceived causal relationship on the strength of stereotypes. Dissertation Abstracts International Section A: Humanities and Social Sciences, 76(6-A(E)).

Reuben, E., Sapienza, P., & Zingales, L. (2014). How stereotypes impair women's careers in science. Proceedings of the National Academy of Sciences, 111(12), 4403–4408. https://doi.org/10.1073/pnas.1314788111

Riegle-Crumb, C., Moore, C., & Buontempo, J. (2017). Shifting STEM stereotypes? Considering the role of peer and teacher gender. Journal of Research on Adolescence, 27(3), 492–505. https://doi.org/10.1111/jora.12289

Riegle-Crumb, C., & Morton, K. (2017). Gendered Expectations: Examining How Peers Shape Female Students' Intent to Pursue STEM Fields. Frontiers in Psychology, 8. https://doi.org/10.3389/fpsyg.2017.00329

Sanchis-Segura, C., Aguirre, N., Cruz-Gomez, A. J., Solozano, N., & Forn, C. (2018). Do gender-related stereotypes affect spatial performance? Exploring when, how and to whom using a chronometric two-choice mental rotation task. Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.01261

Schmader, T., Johns, M., & Barquissau, M. (2004). The Costs of Accepting Gender Differences: The Role of Stereotype Endorsement in Women's Experience in the Math Domain. Sex Roles, 50(11/12), 835–850. https://doi.org/10.1023/B:SERS.0000029101.74557.a0

Scott, A., & Martin, A. (2014). Perceived barriers and endorsement of stereotypes among adolescent girls of color in STEM. 19. https://www.kaporcenter.org/wp-content/uploads/2017/05/perceived_barriers_and_endorsement_of_stereotypes_among_adolescent_girls_of_color_in_stem.pdf

Smeding, A. (2012). Women in Science, Technology, Engineering, and Mathematics (STEM): An investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. Sex Roles: A Journal of Research, 67(11–12), 617–629. https://doi.org/10.1007/s11199-012-0209-4

Smyth, F. L., & Nosek, B. A. (2015). On the gender–science stereotypes held by scientists: Explicit accord with gender-ratios, implicit accord with scientific identity. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.00415

Song, J., Zuo, B., Wen, F., & Yan, L. (2017). Math-gender stereotypes and career intentions: An application of expectancy-value theory. British Journal of Guidance & Counselling, 45(3), 328–340. https://doi.org/10.1080/03069885.2017.1283680

Song, J., Zuo, B., & Yan, L. (2016). Effects of gender stereotypes on performance in mathematics: A serial multivariable mediation model. Social Behavior and Personality: An International Journal, 44(6), 943–952. https://doi.org/10.2224/sbp.2016.44.6.943

Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. Experimental Psychology, 56(4), 283–294. https://doi.org/10.1027/1618-3169.56.4.283

Stamm, K. E. (2010). Stereotype threat and implicit attitudes: Implications for the leaky pipeline of women in science. Dissertation Abstracts International: Section B: The Sciences and Engineering, 71(3-B), 2095.

Starr, C. R. (2018). “I’m not a science nerd!”: STEM stereotypes, identity, and motivation among undergraduate women. Psychology of Women Quarterly, 42(4), 489–503. https://doi.org/10.1177/0361684318793848

Steffens, M. C., & Jelenec, P. (2011). Separating Implicit Gender Stereotypes regarding Math and Language: Implicit Ability Stereotypes are Self-serving for Boys and Men, but not for Girls and Women. Sex Roles, 64(5–6), 324–335. https://doi.org/10.1007/s11199-010-9924-x

Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. Journal of Educational Psychology, 102(4), 947–963. https://doi.org/10.1037/a0019920

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(2), 255–270. https://doi.org/10.1037/a0021385

Swinton, A. D. (2012). African American boys’ and girls’ causal attributions about math, English, and science are shaped by gender stereotypes, influence classroom engagement, and change across the high school years. Dissertation Abstracts International: Section B: The Sciences and Engineering, 72(8-B), 5012.

Tiedemann, J. (2002). Teacher's Gender Stereotypes as Determinants of Teacher Perceptions in Elementary School Mathematics. Educational Studies in Mathematics, 50(1), 49–62. https://doi.org/10.1023/A:1020518104346

Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. Developmental Psychology, 47(4), 943–949. https://doi.org/10.1037/a0024047

Tomasetto, C., Galdi, S., & Cadinu, M. (2012). Quando l'implicito precede l'esplicito: Gli stereotipi di genere sulla matematica in bambine e bambini di 6 anni. Psicologia sociale, 2, 169–186. https://doi.org/10.1482/37693

Tomasetto, C., Mirisola, A., Galdi, S., & Cadinu, M. (2015). Parents' math–gender stereotypes, children's self-perception of ability, and children's appraisal of parents' evaluations in 6-year-olds. Contemporary Educational Psychology, 42, 186–198. https://doi.org/10.1016/j.cedpsych.2015.06.007

Van Camp, A., Gilbert, P., & O'Brien, L. (2019). Testing the effects of a role model intervention on women's STEM outcomes. Social Psychology of Education, 22(3), 649–671. https://doi.org/10.1007/s11218-019-09498-2

Vekiri, I. (2013). Users and experts: Greek primary teachers' views about boys, girls, ICTs and computing. Technology, Pedagogy and Education, 22(1), 73–87. https://doi.org/10.1080/1475939X.2012.753779

Verdugo-Castro, S., Sánchez-Gómez, M. C., García-Holgado, A., & Bakieva, M. (2020). Pilot study on university students' opinion about STEM studies at higher education. Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality, 158–165. https://doi.org/10.1145/3434780.3436616

Vuletich, H. A., Kurtz-Costes, B., Cooley, E., & Payne, B. K. (2020). Math and language gender stereotypes: Age and gender differences in implicit biases and explicit beliefs. PLoS ONE, 15(9). https://doi.org/10.1371/journal.pone.0238230

White, M. J., & White, G. B. (2006). Implicit and explicit occupational gender stereotypes. Sex Roles: A Journal of Research, 55(3–4), 259–266. https://doi.org/10.1007/s11199-006-9078-z

Wille, E., Gaspard, H., Trautwein, U., Oschatz, K., Scheiter, K., & Nagengast, B. (2018). Gender stereotypes in a children's television program: Effects on girls' and boys' stereotype endorsement, math performance, motivational dispositions, and attitudes. Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.02435

Wortel, K. E. (1997). Teacher gender-role identity and gender attribution of student mathematic ability. Dissertation Abstracts International: Section B: The Sciences and Engineering, 57(8-B), 5354.

Young, D., Rudman, L., Buettner, H., & McLean, M. (2013). The Influence of Female Role Models on Women's Implicit Science Cognitions. Psychology of Women Quarterly, 37(3), 283–292. https://doi.org/10.1177/0361684313482109

Zhao, F., Zhang, Y., Alterman, V., Zhang, B., & Yu, G. (2018). Can math-gender stereotypes be reduced? A theory-based intervention program with adolescent girls. Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 37(3), 612–624. https://doi.org/10.1007/s12144-016-9543-y

Zitelny, H., Shalom, M., & Bar-Anan, Y. (2017). What is the implicit gender-science stereotype? Exploring correlations between the gender-science IAT and self-report measures. Social Psychological and Personality Science, 8(7), 719–735. https://doi.org/10.1177/1948550616683017

Attribution to the gender gap and gender-science association

Introduction

As mentioned in Chapter 2, the freedom in the interpretation of masculinity and femininity in questions on the 'gendered domain' construct is a serious limitation of this type of studies. To provide an example, I analysed data from Project Implicit on a subsample of U.S. residents aged from 18 to 30 years old.

Project Implicit is a non-profit organisation of researchers investigating implicit social cognition (Nosek et al., 1998). The website hosts several Implicit Association Tests, including the IAT on gender and science. There is no restriction on access to the test, which can be performed on the website for free at any time. Along with the IAT, a questionnaire is proposed, including sociodemographic questions and measurements of explicit gender stereotypes on the theme of the IAT.

I exploited the fact that questions on explicit gender stereotypes included both 'gendered domain' and 'attribution' constructs to verify whether different causes addressed to gender bias in science were reflected in different answers to the question on masculinity and femininity of this domain.

Research questions

In particular, the objective here was twofold:

> O1: To test whether two hypothesized components could be distinguished among the causes attributed to gender differences in science, i.e., causes related to personal characteristics and causes related to social or contextual factors.

The two components differ somehow in the locus of attribution. In the first case, the locus is internal to the individual, while in the second case the locus is external. This distinction is also reflected in a different role assumed by the individual. If the causes are internal, individuals can exercise control over them and so the consequences can be attributed to the individual. On the contrary, if the causes are external, they are not under the control of any individual who thus cannot be responsible for any consequence.

> O2: To verify whether the two identified latent components are associated with implicit and explicit gender stereotypes, the latter being measured using the construct of the gendered domain.

Elaborating a hypothesis on the relationships between the two latent components and explicit and implicit measures was impossible given that this is, to the best of my knowledge, the first study addressing this issue.

Methodology

Sample

The sample was retrieved from data collected by Project Implicit from 2007 to 2019 on the gender-science IAT (Xu et al., 2020). Only cases from the U.S., who (1) both performed the test and answered self-reported questions in the survey and (2) had at least 18 years old and no more than 30 years old, were retained. The final sample consisted of 150,749 individuals (M_{age}=22, SD_{age}=3.36), for the majority female (70%). Most participants identified as White (68%) and the majority attended at least some years of college (42%).

Instruments

Explicit gender stereotypes. There are multiple instruments aimed at measuring explicit gender stereotypes. Here, the interest was in the question ‘Please rate how much you associate science with males or females’, measured on a 7-point Likert scale (1 = ‘Strongly female’, 4 = ‘Neither female nor male’, 7 = ‘Strongly male’).

Implicit gender stereotypes. Implicit gender stereotypes were inferred by the D-score resulting from the IAT (Greenwald et al., 2003). The version proposed by Project Implicit includes gender (e.g., man, son, woman, daughter) as target and science and humanistic majors (e.g., astronomy, math, history, arts) as categories.

Attribution to gender bias in science. The question on attribution was stated as follows: ‘Women hold a smaller portion of the science and engineering faculty positions at top research universities than do men. The following factors were typically included to explore possible explanations for these differences’. Participants were asked to rate each of the following six items on a 5-point Likert scale (1 = ‘Not at all important’, 5 = ‘Extremely important’).

- Different proportions of men and women are found among people with the very highest levels of math ability (item *abilities*).
- On average, men and women differ in their willingness to devote the time required by such high-powered positions (item *power*).
- On average, men and women differ naturally in their scientific interest (item *interest*).
- On average, men and women differ in their willingness to spend time away from their families (item *family*).
- Directly or indirectly, boys and girls tend to receive different levels of encouragement for developing scientific interest (item *encouragement*).
- On average, whether consciously or unconsciously, men are favoured in hiring and promotion (item *discrimination*).

Analytical methods

Structural equation modelling (SEM) was applied to answer the two research questions. SEM is a family of statistical techniques used to estimate the relationships among constructs, as it is a combination of factor analysis and path analysis (Weston & Gore, 2006). It consists of two components: a measurement model describing the relationship between observed variables and latent constructs and a structural model describing the interrelationship among constructs (Weston & Gore, 2006). In this case, the measurement model captures the relationship between the six attribution items and the hypothesized two latent constructs, i.e., personal and social causes for gender differences in science (O1). The structural model considers the relationship between the two latent constructs and the implicit and explicit indicators of gender stereotypes (O2).

Following Weston and Gore (2006), the measurement model was tested using factor analysis and the structural model using path analysis. After randomly splitting the sample into two subsamples, an exploratory factor analysis (EFA) was performed on the first subsample to assess the number of latent components. Then, the measurement model was tested on the other subsample through confirmatory factor analysis (CFA). Finally, a path analysis tested the fit of the structural model. The analysis was conducted in R (R Core Team, 2017), using the package psych (Revelle, 2020) to perform the EFA and the package lavaan (Rosseel, 2012) to perform the CFA and the path analysis.

Results

I performed an exploratory factor analysis and compared a one-factor, two-factor and three-factor solution. Scree tests and parallel analysis determined the number of factors to retain. Coherently with the hypotheses, results suggested that the six items could be grouped into two distinct components, the former including ‘discrimination’ and ‘encouragement’, the latter including ‘power’, ‘family’, ‘interest’ and ‘abilities’. Table 2.5 shows the factor loadings greater than 0.4 for each item in the two-factor solution.

Table 2.5 *Factor loadings from the two-factor solution in EFA*

Item	Factor 1	Factor 2
Ability	.59	
Interest	.73	
Family	.69	
Power	.83	
Discrimination		.91
Encouragement		.72

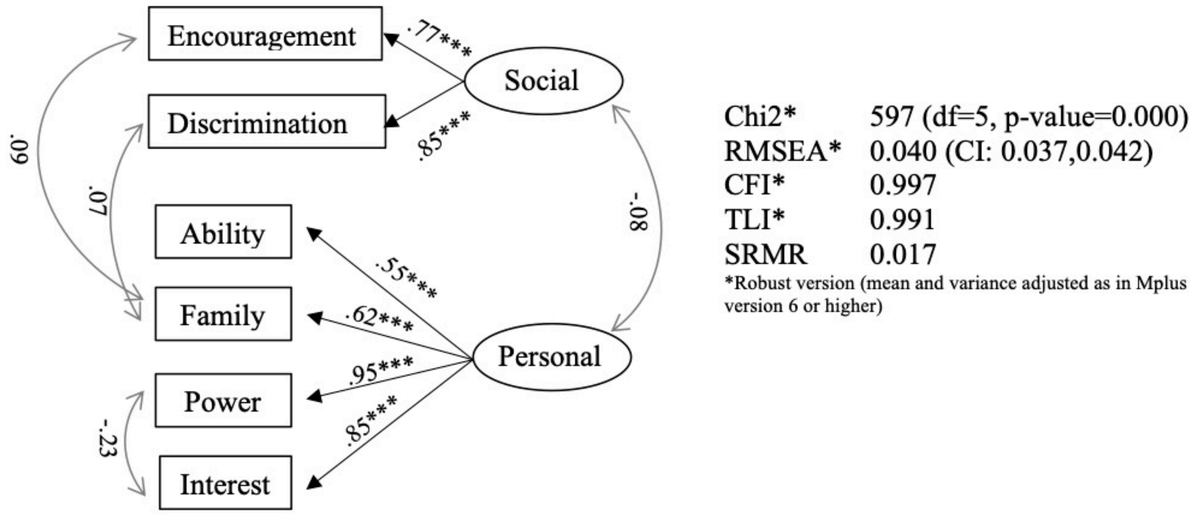
Note. Extraction method: weighted least square

Rotation: orthogonal varimax

Correlation matrix: polychoric

The fit of the two-factor model was tested in confirmatory factor analysis. The CFA confirmed the result of the EFA, as the two-factor model had a better fit compared to the one-factor model ($SRMR_{\text{one-factor}}=0.158$, $SRMR_{\text{two-factor}}=0.017$). The fit of the model as indicated by the SRMR is acceptable (Hu and Bentler 1999). Figure 2.2 shows the factor loadings and the correlation between the latent constructs from the CFA. Other fit indices besides SRMR are reported, even if usual cut-offs are not reliable in the case of categorical data (Xia & Yang, 2019). The decision on the goodness of fit was thus based on the SRMR, as suggested by Shi and Maydeu-Olivares (2020). Note that the social and personal latent constructs seem uncorrelated, while there is a negative and weak correlation between the item ‘interest’ and the item ‘power’.

Figure 2.2 *Measurement model from CFA*



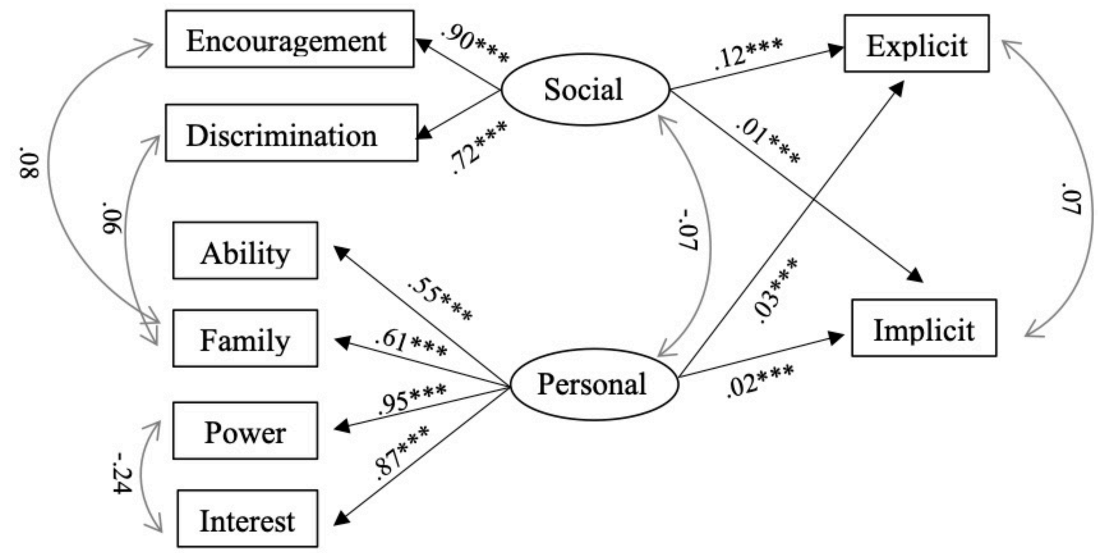
Chi2*	597 (df=5, p-value=0.000)
RMSEA*	0.040 (CI: 0.037,0.042)
CFI*	0.997
TLI*	0.991
SRMR	0.017

*Robust version (mean and variance adjusted as in Mplus version 6 or higher)

Before adding the structural part of the model, I checked whether there were differences in the model related to the respondents' gender, a property known as 'measurement invariance' (Van De Schoot et al., 2015). Measurement invariance was not reached, meaning that the association between the items and the latent factors (i.e., factor loadings, item intercepts and item residual variances) depended on respondents' gender. The path analysis was, therefore, applied to sampled men and women separately.

Figures 2.3 and 2.4 show the structural and measurement models for the female group and the male group, respectively. For both groups, personal and social latent factors had a statistically significant direct effect on implicit gender stereotypes regardless of their strength. The direct effect on explicit gender stereotypes was statistically significant and greater in size for both groups. While in the case of women, the explicit indicator was positively associated with the social component of attribution, for men the explicit indicator was positively associated with the personal component.

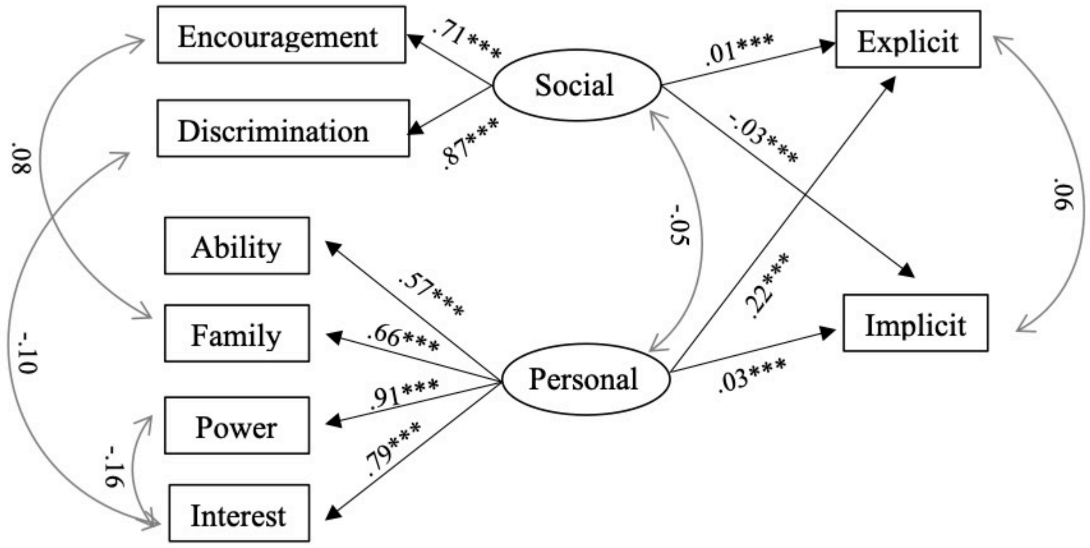
Figure 2.3 *Structural and measurement models for women*



Chi2*: 365.049 (df=13, p-value=0.000); RMSEA*: 0.023 (CI: 0.021-0.025); CFI*: 0.997, TLI*: 0.994; SRMR: 0.011

*Robust version (mean and variance adjusted as in Mplus version 6 or higher)

Figure 2.4 *Structural and measurement models for men*



Chi2*: 477.128 (df=13, p-value=0.000); RMSEA*: 0.04 (CI: 0.037-0.043); CFI*: 0.993, TLI*: 0.984; SRMR: 0.019

*Robust version (mean and variance adjusted as in Mplus version 6 or higher)

Discussion and conclusions

To sum up, the results confirmed the initial hypothesis that attributions to the gender gap in science are related to two distinct factors, which I called ‘personal’ and ‘social’. The personal component of attribution includes causes pertaining to individual characteristics, attitudes or choices, while the social component includes those causes referring to behaviour or choices of others that affect the individual.

However, this configuration was not equivalent for men and women. Indeed, when looking at the association between the two factors and explicit and implicit indicators of gender stereotypes, there was an important difference. On the one hand, in both groups, neither the social nor the personal component had a strong (although statistically significant) association with implicit stereotypes. On the other hand, in the female group, the social component had a small and significant association with explicit gender stereotypes, whereas in the male group the personal component had a small and significant association with explicit gender stereotypes.

This means that whenever being asked questions on the masculinity of science domains, men mainly referred to differences on which women (and men) have control, thus linked to choices and characteristics. Conversely, women mainly referred to differences that women (and men) undergo because of the choices and behaviours of others. To conclude, when questions on the association between gender and science are ambiguous, the deriving scales measure different constructs of gender stereotypes, at least in part depending on the respondent’s gender.

References

Greenwald, A. G., Nosek, B., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. Journal of Personality and Social Psychology, 85(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Nosek, B., Banaji, M. R., & Greenwald, A. G. (1998). Project Implicit. https://implicit.harvard.edu/implicit/index.jsp

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Revelle, W. (2020). psych: Procedures for Psychological, Psychometric, and Personality Research (R package version 2.0.9) [Computer software]. Northwestern University. https://CRAN.R-project.org/package=psych

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Shi, D., & Maydeu-Olivares, A. (2020). The Effect of Estimation Methods on SEM Fit Indices. Educational and Psychological Measurement, 80(3), 421–445. https://doi.org/10.1177/0013164419885164

Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement Invariance. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.01064

Weston, R., & Gore, P. A. (2006). A Brief Guide to Structural Equation Modeling. The Counseling Psychologist, 34(5), 719–751. https://doi.org/10.1177/0011000006286345

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. Behavior Research Methods, 51(1), 409–428. https://doi.org/10.3758/s13428-018-1055-2

Xu, K., Nosek, B., Greenwald, A., Ratliff, K., Bar-Anan, Y., Umansky, E., Banaji, M., Lofaro, N., Smith, C., Axt, J., & Lai, C. (2020). Project Implicit Demo Website Datasets. https://doi.org/10.17605/OSF.IO/Y9HIQ

Chapter 3: Supplementary material

Instruments

Competition

Imagine that your class is called to take part in a competition and that you are asked to form a team of four students who will represent your class. Note that in case of victory, the class as a whole would win the price (a trip to a European city), not only those who actually took part in the competition.

1. Who among your classmates do you think should be part of the team if the competition were about math? Select four students. Note that you can include yourself.
2. Who among your classmates do you think should be part of the team if the competition were about reading? Select four students. Note that you can include yourself.
3. Who among your classmates do you think should be part of the team if the competition were about science? Select four students. Note that you can include yourself.

Identification with the subject

Items:

- How important to you are your math abilities?
- How important to you are your science abilities?
- How important to you are your reading abilities?
- How important to you is that others believe you are good at math?
- How important to you is that others believe you are good at science?
- How important to you is that others believe you are good at reading?

- How important do you feel that math skills and abilities are to your success in college?
- How important do you feel that science skills and abilities are for your success in college?
- How important do you feel that reading skills and abilities are for your success in college?
- How important do you feel that math skills and abilities are for your future career?
- How important do you feel that science skills and abilities are for your future career?
- How important do you feel that reading skills and abilities are for your future career?
- If you were to fail a test in a math class, how bothered would you be by your failure?
- If you were to fail a test in a science class, how bothered would you be by your failure?
- If you were to fail a test in a reading class, how bothered would you be by your failure?

Descriptive statistics

Table 3.3 shows mean and standard deviations of performance, identification with the subject and self-nominations by gender.

Table 3.3 *Variables mean and standard deviation*

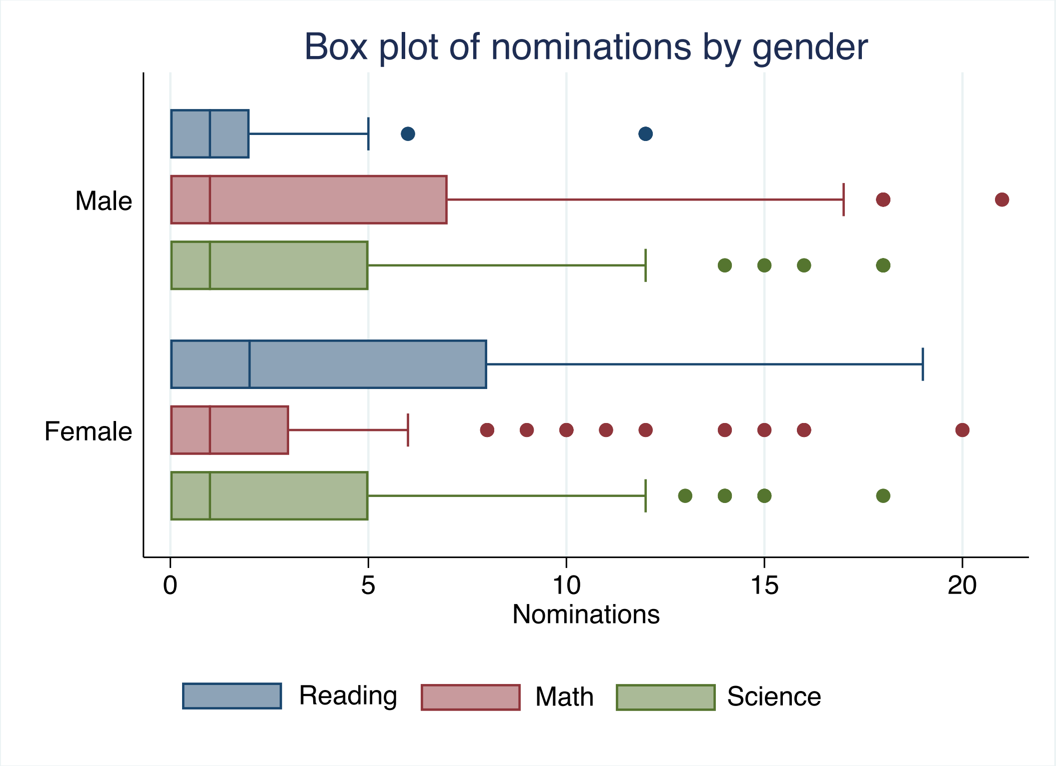
Variable	Total		Male		Female		t	Prob (T>t)	Prob (T<t)
	Mean	SD	Mean	SD	Mean	SD			
Performance									
Reading	7.08	.85	6.94	.90	7.23	.95	-2.42		.008
Math	6.72	.97	6.69	1.05	6.74	.93	-.29		.383
Science	7.10	.94	6.92	.69	7.21	.93	-2.14		.017
Identification									
Reading	3.52	.80	3.34	.83	3.65	.76	-2.63		.005
Math	3.07	.91	3.16	.86	3.02	.95	1.02	.156	
Science	3.05	.95	3.18	.89	2.96	.98	1.51	.067	
Self-nominations									
Reading	.39	.49	.41	.49	.36	.48	.68	.248	
Math	.35	.48	.45	.50	.27	.45	2.56	.006	
Science	.33	.47	.46	.50	.24	.43	3.34	.001	

Note. SD = Standard deviation; t = t statistics

On average, female students had slightly higher grades in both science and reading, while no difference in maths was found. As regards identification with the subject, females had higher identification with reading compared to male students. Finally, males nominated themselves more frequently than female students in all three competitions, but the difference was statistically significant only in case of math and science.

Figure 3.1 shows the box plot of nominations by gender. For the math competition, males generally preferred female candidates, while in case of reading, this preference was reversed. In case of science, no relevant difference in nominations of male and female students was found.

Figure 3.1 *Box plot of nominations by gender*



Graphical representation of networks

Each node represents a student in the class (females are represented as round shapes, males as squares). The number inside the node indicates the student's final grade in the corresponding school subject (if empty, the value was missing). Nominations are represented by links, which point to the nominated candidate.

Figure 3.2 *Network 1*

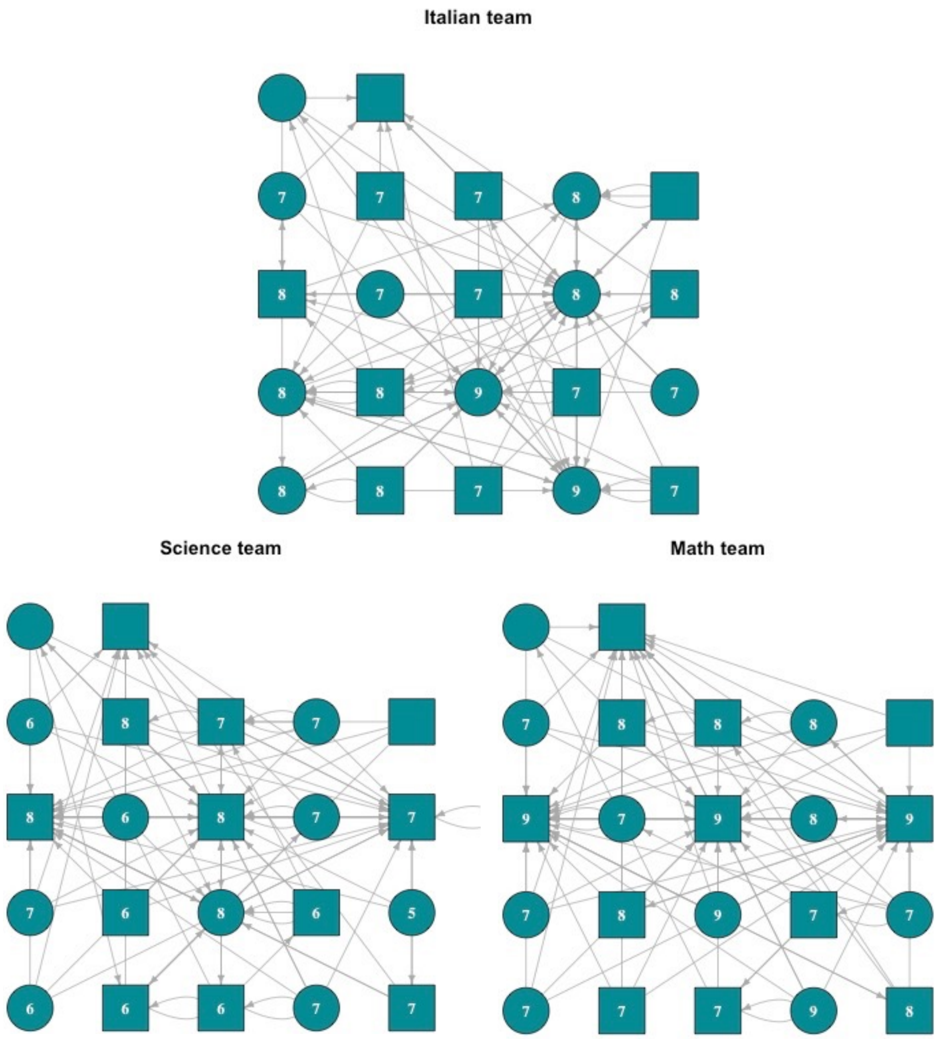
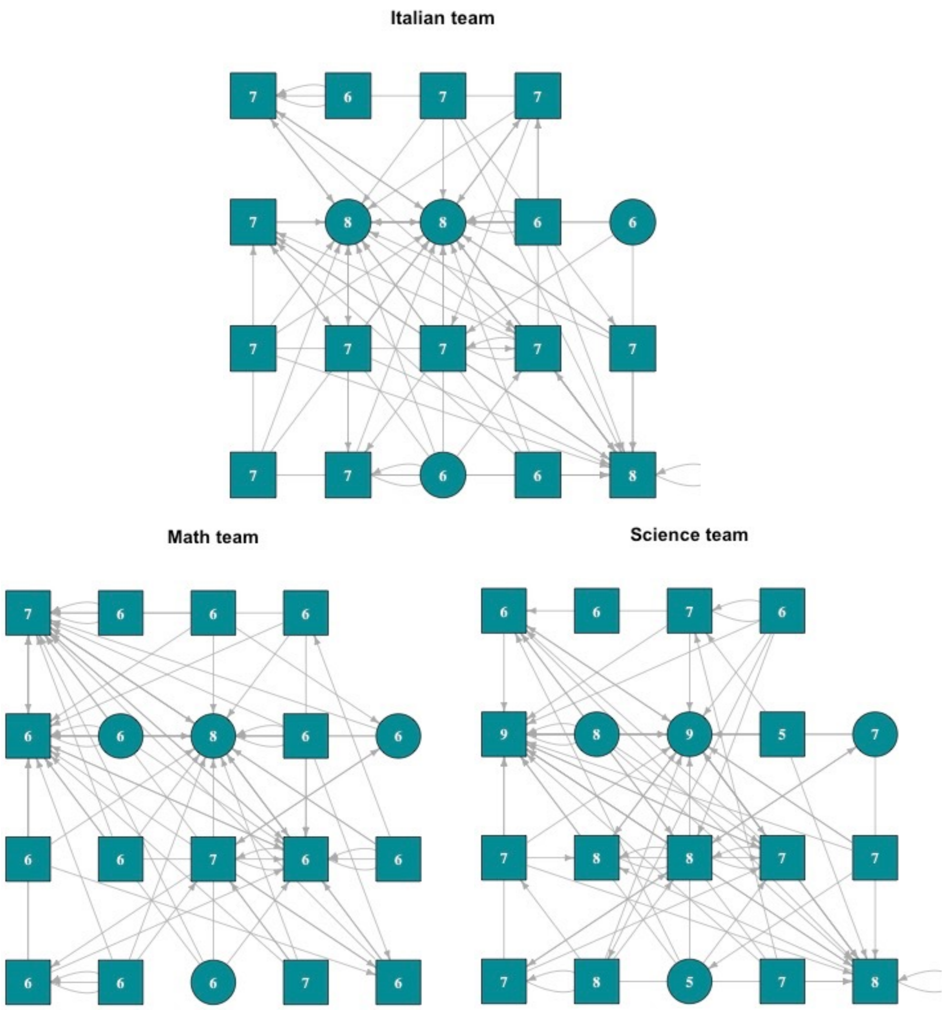


Figure 3.3 *Network 2*

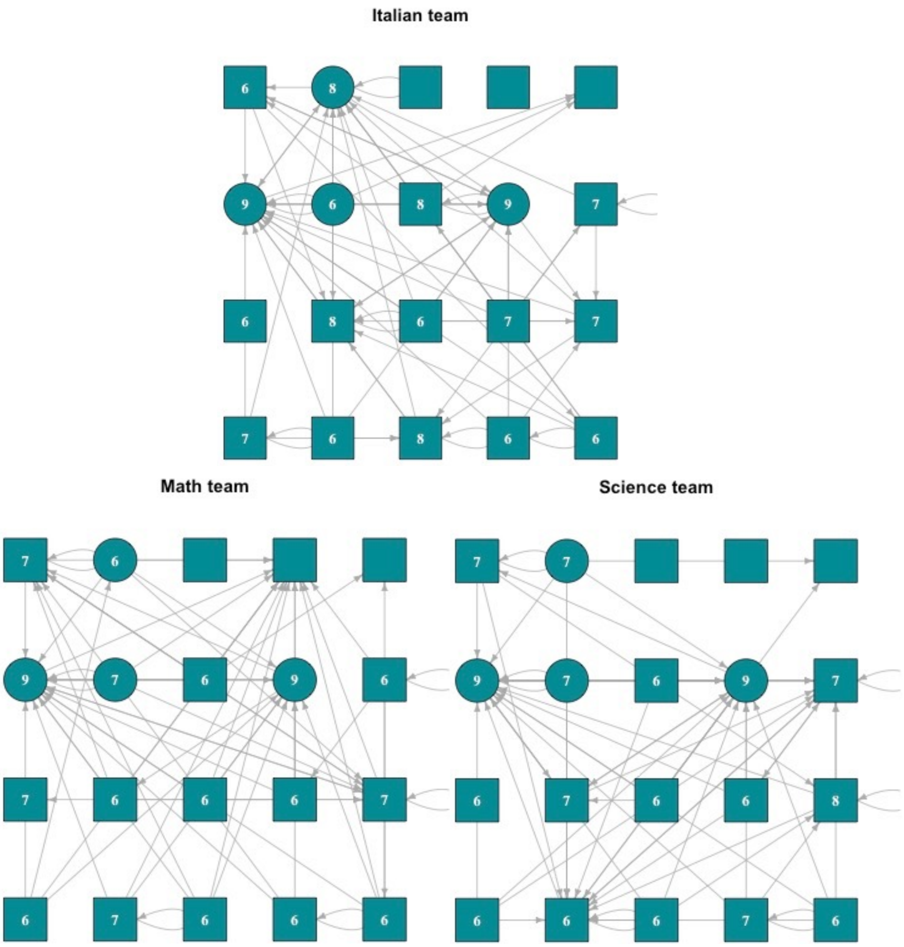


Italian team

Math team

Science team

Figure 3.4 *Network 3*

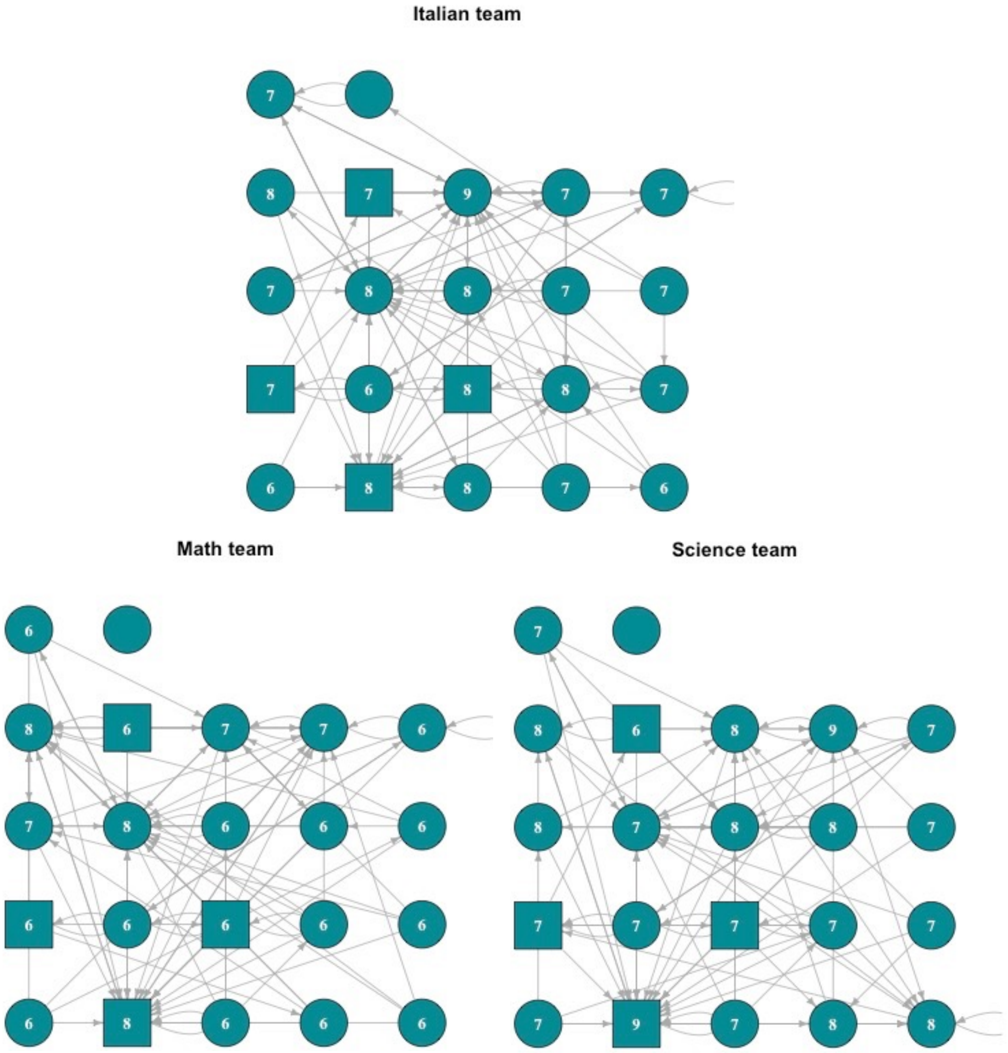


Italian team

Math team

Science team

Figure 3.5 *Network 4*

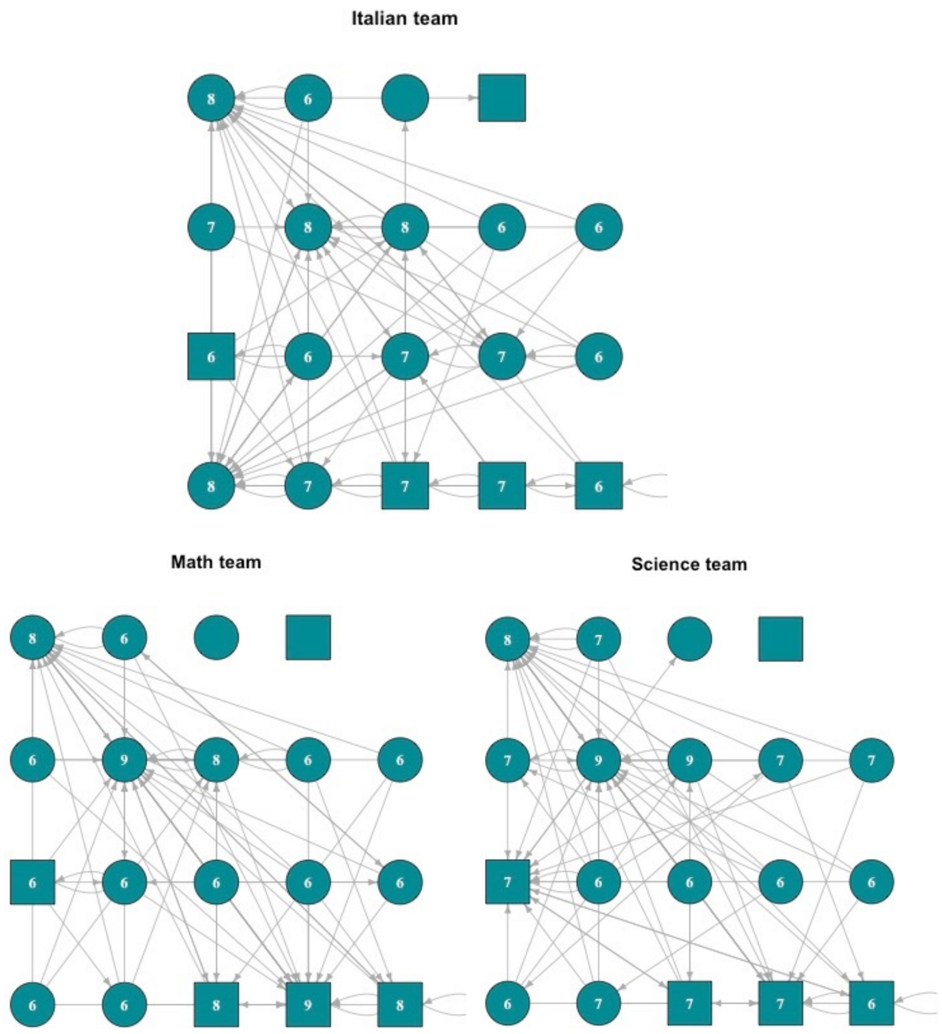


Italian team

Math team

Science team

Figure 3.6 *Network 5*


Italian team

Math team

Science team

Figure 3.7 *Network 6*

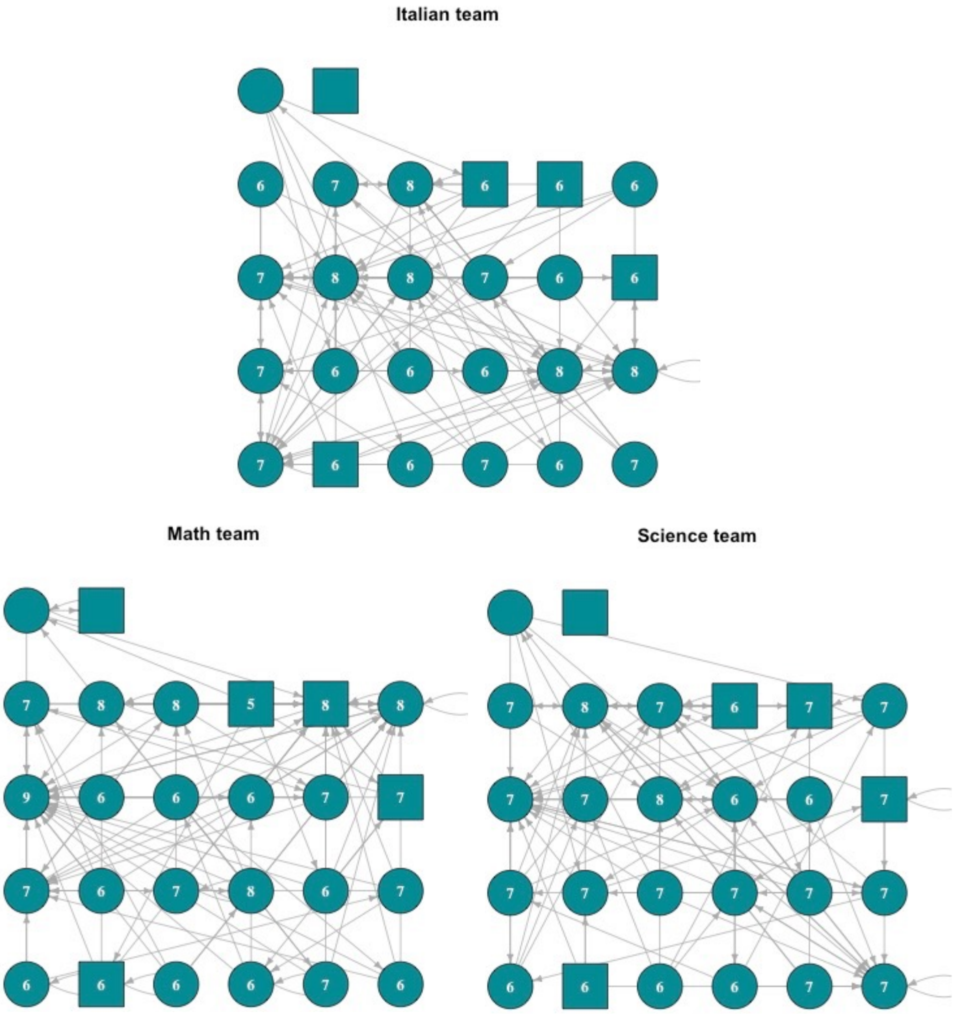
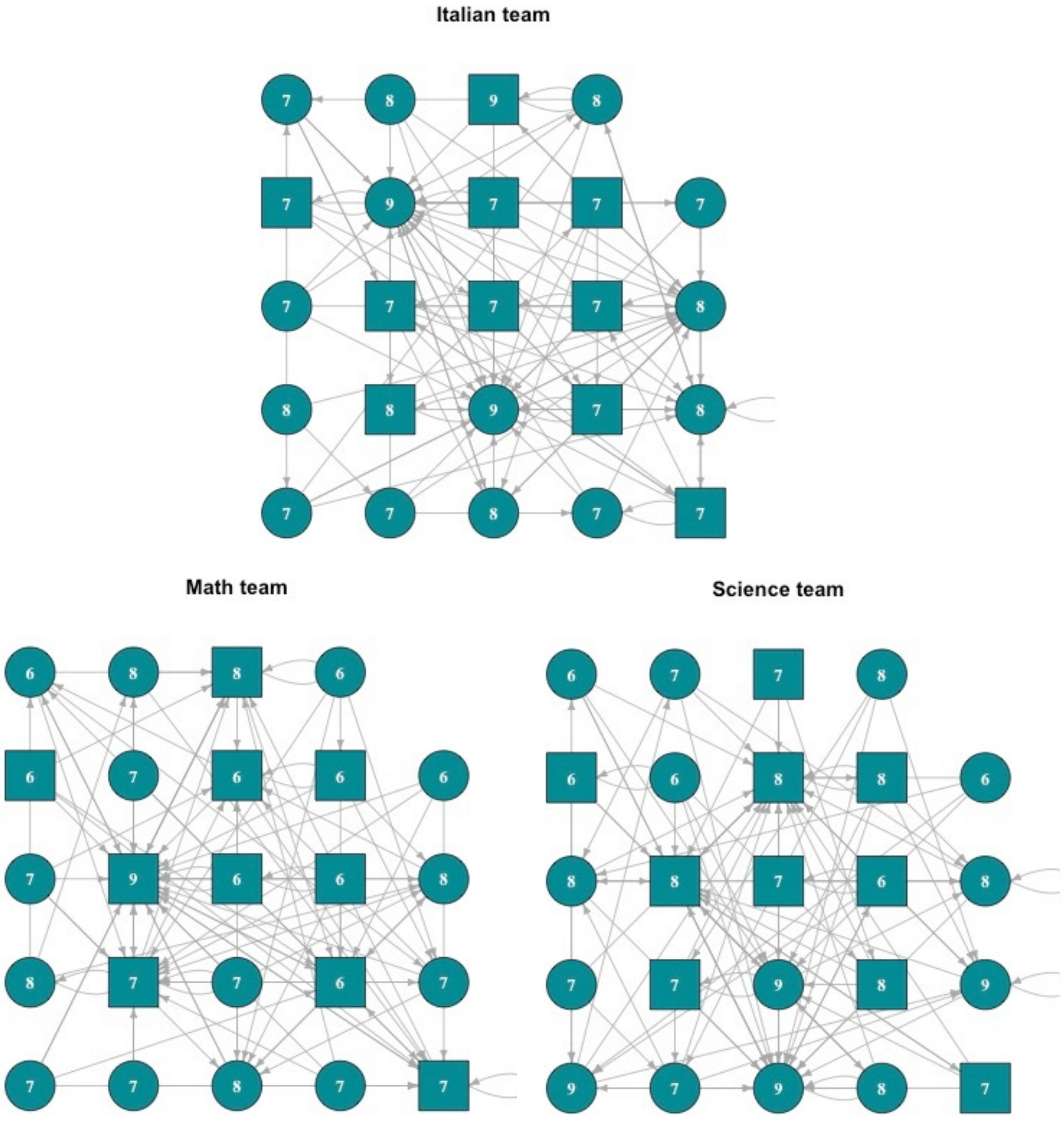


Figure 3.8 *Network 7*



Italian team

Math team

Science team

ERGM results

Models were estimated through the *ergm* package (v. 1.37.5; Handcock et al. 2020) implemented in the *R* programming language (v. 4.0.3; R Core Team 2020).

Table 3.4 summarizes the characteristics of networks.

Table 3.4 *Networks' characteristics*

Network	Number of students	Response rate	% of females	Type of school
Network 1	22	5%	50%	Science-oriented (liceo scientifico)
Network 2	19	5%	26%	Science-oriented (liceo scientifico)
Network 3	20	20%	15%	Science-oriented (liceo scientifico)
Network 4	22	5%	82%	Languages-oriented (liceo linguistico)
Network 5	19	11%	74%	Languages-oriented (liceo linguistico)
Network 6	26	8%	81%	Languages-oriented (liceo linguistico)
Network 7	24	0%	58%	Humanities-oriented (liceo classico)

There are two types of missing values in network data, i.e., missing on edges and missing on attributes. ERGM can deal with missingness in edges, but not in attributes. Given that in this case, data were missing randomly (from 4% to 20% in the 7 networks), replacements of missing attribute data were estimated via multivariate imputation by chained equations using the *mice* package in *R* (van Buuren and Groothuis-Oudshoorn 2011). This allowed to substitute missing values with plausible values derived from a distribution of the imputed variable. In particular, the following parameters were specified in the mice package.

- Number of imputed datasets: 5 (10 in network C because of greater number of missing values, 20%)
- Method: pmm
- Imputed variables: grade
- Predictors: gender, in-degree, type of school

Table 3.5, Table 3.6 and Table 3.7 report the outcomes of the ERGM for each network. These outcomes were used to perform the meta-analysis presented in Table 3.1 of Chapter 3. For each variable in the model, Table 3.1 shows the mean effect size, computed as a weighted least squares estimator, its standard error and the t-ratio testing that the mean effect is zero. Snijders and Baerveldt (2003:15-17) were followed for the formula to compute the effect size, the standard error and the t-test. These are the three inputs of the function: (1) a vector with the coefficients from the seven ERGM function estimates; (2) a vector with the corresponding standard errors; and (3) the number of networks included in the meta-analysis.

Models' goodness of fit can be retrieved at https://github.com/elenadegio/The-stubbornness-of-gender-stereotypes-in-education/tree/main/Chapter%203.

Table 3.5 *Regression results of reading competition*

	Science-oriented			Languages-oriented			Humanities -oriented
	Net 1	Net 2	Net 3	Net 4	Net 5	Net 6	Net 7
Receiver: gender [female]	.84*	.60	.69*	.11	.87*	.19	.71*
	(.39)	(.61)	(.27)	(.24)	(.45)	(.22)	(.31)
Receiver: school grade	1.12***	1.92***	.38**	1.09***	.95***	1.03***	.98***
	(.28)	(.44)	(.15)	(.31)	(.28)	(.05)	(.24)
Edges	-10.73***	-15.02***	-5.67***	-11.03***	-10.45***	-8.76***	-10.05***
	(2.30)	(3.17)	(1.22)	(2.42)	(2.37)	(2.62)	(2.06)
Reciprocity	-.61	-.38	-1.33	-.06	-.91	-1.42*	-.75
	(.66)	(.85)	(.89)	(.64)	(.74)	(.47)	(.68)
GWESP	.56	.45	1.64**	1.05***	1.25***	.37	.59**
	(.31)	(.41)	(.37)	(.27)	(.38)	(.06)	(.20)
GWIDEGREE	-1.08	-1.29	-1.34	.71	-.12	-1.43	.08
	(.87)	(1.00)	(.87)	(.85)	(.99)	(.59)	(.76)
Friendship	.72*	.66	1.09**	.84**	.96*	1.24***	.87**
	(.33)	(.46)	(.41)	(.31)	(.38)	(.10)	(.28)

Note. Standard errors in parenthesis

***p < 0.001; **p < 0.01; *p < 0.05

Table 3.6 *Regression results of math competition*

School type	Science-oriented			Languages-oriented			Humanities-oriented
	Net 1	Net 2	Net 3	Net 4	Net 5	Net 6	Net 7
Receiver: Gender [female]	-3.08***	-.01	-.13	-.18	1.47**	-.14	-1.25***
	(1.43)	(.19)	(.36)	(.41)	(.48)	(.31)	(.29)
Receiver: School grade	1.91***	.44**	.87**	1.41**	1.42***	.97***	.73***
	(.99)	(.15)	(.32)	(.49)	(.31)	(.25)	(.16)
Edges	-17.67***	-4.88***	-9.11***	-11.91***	-14.76***	-10.11***	-4.90***
	(8.56)	(1.23)	(2.46)	(3.29)	(2.74)	(1.79)	(1.18)
Reciprocity	-.98	-.88	-2.44*	-.02	.09	-.75	-2.25**
	(.87)	(.77)	(1.07)	(.79)	(.87)	(.62)	(.74)
GWESP	.08	1.14*	1.46***	1.08*	1.12*	1.33***	.31
	(.58)	(.48)	(.39)	(0.56)	(.55)	(.34)	(.28)
GWIDEGREE	.63	-3.72***	-1.26	.58	-.53	1.56	-3.57***
	(2.38)	(1.11)	(.94)	(.94)	(1.15)	(.82)	(.81)
Friendship	1.78***	1.98***	1.38**	.97*	1.94**	1.31***	1.25***
	(.57)	(.45)	(.43)	(.39)	(.51)	(.32)	(.32)

Note. Standard errors in parenthesis

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3.7 *Regression results of science competition*

School type	Science-oriented		Languages-oriented			Humanities-oriented
	Net 1	Net 2	Net 4	Net 5	Net 6	Net 7
Receiver: Gender [female]	$-.73^{*}$	-.27	-.28	$-.76^{*}$	.39	-.39
	(.32)	(.29)	(.22)	(.39)	(.29)	(.22)
Receiver: School grade	$.67^{**}$	.64	$.35^{*}$	$.58^{**}$	.24	$.59^{**}$
	(.25)	(.22)	(.17)	(.29)	(.14)	(.21)
Edges	-6.05^{***}	-7.75^{***}	-4.46^{***}	-5.97^{***}	-3.48^{**}	-6.29^{***}
	(1.77)	(1.75)	(1.29)	(1.81)	(1.06)	(1.64)
Reciprocity	-.87	-.29	-1.55	-1.10	-.28	-.71
	(.67)	(.66)	(.82)	(.74)	(.52)	(.68)
GWESP	$.97^{*}$	1.50^{**}	$.77^{*}$	1.22^{**}	$.72^{**}$	$.84^{***}$
	(.37)	(0.44)	(.31)	(.42)	(.23)	(.19)
GWIDEGREE	-2.19	-.12	-1.03	-.95	-1.60^{*}	-.83
	(.99)	(1.00)	(.79)	(1.02)	(.63)	(.77)
Friendship	.40	1.29^{***}	.48	.57	.28	$.79^{**}$
	(.32)	(.38)	(.33)	(.38)	(.28)	(.26)

Note. Standard errors in parenthesis

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

References

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., Krivitsky, P. N., & Morris, M. (2020). *ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks* (R package version 3.11.0) [Computer software]. The Statnet Project (\url{https://statnet.org}). https://CRAN.R-project.org/package=ergm

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Snijders, T. A. B., & Baerveldt, C. (2003). A multilevel network study of the effects of delinquent behavior on friendship evolution. *The Journal of Mathematical Sociology*, *27*(2–3), 123–151. https://doi.org/10.1080/00222500305892

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://www.jstatsoft.org/v45/i03/

Regression analysis without multiple imputation of missing data

In case of few clusters, clustered standard errors can be downwards biased. The best technique to deal with this issue is to use wild cluster bootstrap standard errors, a type of cluster bootstrap resampling method with asymptotic refinement (Cameron and Miller 2015; Roodman et al. 2019). However, existing commands used in common statistical software are not compatible with those used for multiple imputation of missing data. To solve this issue, the results reported in Chapter 3 use another bootstrap technique suggested in case of few clusters, yet less efficient compared to the wild cluster bootstrap (Cameron and Miller 2015), i.e., the percentile cluster bootstrap standard errors.

To check whether there are differences in the results when we use wild cluster bootstrap standard errors, Table 3.8 reports the results of the two logistic regression models estimated using the original dataset including missing values. Note that in case of a binary dependent variable, residuals are not well-defined, which prevents application of the wild bootstrap. I thus used the adaptation of the wild bootstrap developed by Kline and Santos (2010) and available in the command 'boottest' (Roodman 2015). Results are consistent with those estimated on the imputed dataset.

Table 3.8 *Regression results on self-nominations (non-imputed data)*

	Estimate	SE	Score bootstrap-t Wald test	
			z	p-value
Reading competition				
Gender [Female]	-1.38	.45	-2.70	.027
School grade	1.14	.40	3.49	.010
Identification with Italian	1.05	.41	2.49	.040
Math competition				
Gender [Female]	-1.68	.53	-3.05	.017
School grade	1.01	.29	4.04	.003
Identification with Italian	.34	.14	2.11	.091
Science competition				
Gender [Female]	-2.13	.40	-5.06	.000
School grade	.66	.46	2.11	.086
Identification with Italian	.74	.29	2.59	.058

Note. Total N = 257; SE = Standard error

[a] Reference category ‘Above high school’

[b] Reference category ‘Humanities-oriented’

References

Cameron, A. C., & Miller, D. L. (2015). A Practitioner’s Guide to Cluster-Robust Inference. *Journal of Human Resources*, *50*(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

Kline, P., & Santos, A. (2010). *A Score Based Approach to Wild Bootstrap Inference* (No. w16127; p. w16127). National Bureau of Economic Research. https://doi.org/10.3386/w16127

Roodman, D. (2015). *BOOTTEST: Stata module to provide fast execution of the wild bootstrap with null imposed* [Statistical Software Components S458121]. Boston College Department of Economics.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal: Promoting Communications on Statistics and Stata*, *19*(1), 4–60. https://doi.org/10.1177/1536867X19830877

Chapter 4 Supplementary material

Descriptive statistics

Figure 4.2 shows students' majoring intentions for nine disciplinary sectors, i.e., Arts and humanities, Linguistic, Socio-political, Statistical and economic, Law, Health, Veterinary and agrarian, Scientific, and Architecture and Engineering., by gender.

Figure 4.2 *Students' major intentions by gender*

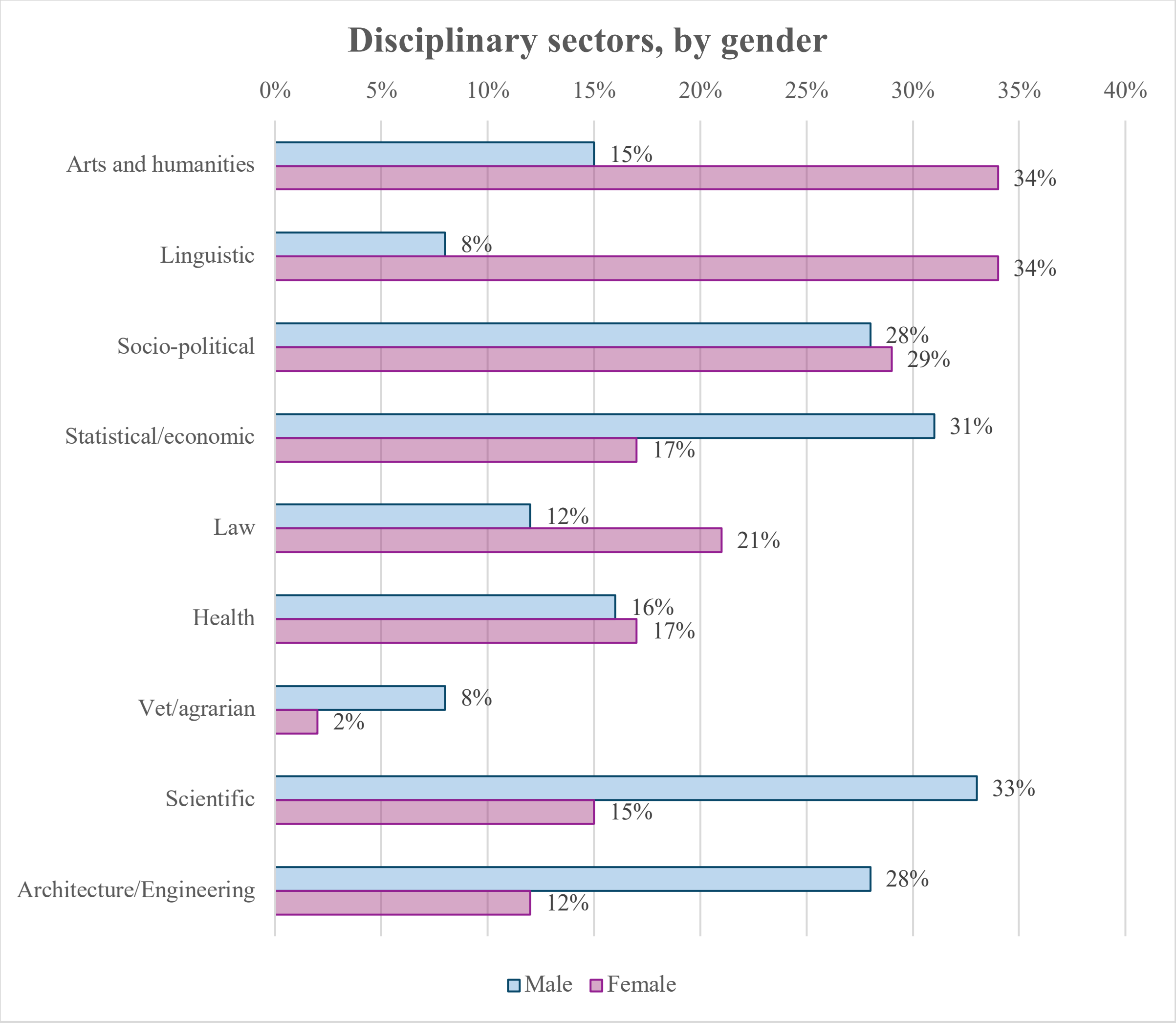


Table 4.5 shows the result of a two-sample t-test comparing male and female students.

Table 4.5 *t-test of gender differences*

	Male		Female			
	Mean	SD	Mean	SD	t[a]	p-value[b]
Intention of majoring in STEM	.50	.5	.23	.4	4.97	.000
Math grade	6.82	1.1	6.88	1.1	-.42	.336
Science grade	6.89	.9	7.23	1.0	-2.86	.002
Interest in STEM subjects	0.73	.5	.35	.5	6.89	.000
Identification with STEM	3.28	.8	2.92	.8	3.69	.000
Relevance of job’s salary on choice	.35	.5	.14	.3	4.53	.000
Intention of majoring in humanities	.15	.4	.34	.5	-3.78	.000
Italian grade	6.85	.7	7.39	1.1	-4.87	.000
Interest in humanistic subjects	.65	.5	.66	.5	-.23	.407
Identification with humanities	3.5	.7	3.7	.7	-3.15	.001
Relevance of job’s social utility on choice	.33	.5	.44	.5	-1.77	.039
IAT D score	.23	.4	.35	.4	-2.38	.009

Note. SD = Standard deviation; t = t statistic

[a] H_0: $(\text{mean}_{\text{male}} - \text{mean}_{\text{female}}) = 0$; DF = 302

[b] Pr(T<t) when t<0; Pr(T>t) when t>0

As expected, compared to their female peers, male students were more likely to express the intention of majoring in STEM, were more interested and had a higher identification with the STEM field. They were also more likely to consider future job’s salary as relevant in the choice of majors. On the contrary, female students were more likely to express the intention of majoring in humanities, had higher grades in Italian and a higher identification with the humanistic field. They were also more frequently interested in the expected social utility of their future job. In line with recent trends on the improved performance of female students in math and science at school (OECD, 2020), female students outperformed male students in science. Finally, females had stronger stereotypical associations of gender and STEM/humanities fields than their male peers and the difference was statistically significant.

However, the characteristics of the context, i.e., the fact that secondary school in Italy is structured so that students in high school have already chosen between humanities and STEM orientation, requires careful attention whenever examining these gender differences. When comparing male and female students within the same type of school, i.e., humanities-oriented or STEM-oriented, the differences were not statistically significant anymore for students in a humanities-oriented school, except for grades in science and Italian. On the contrary, for those enrolled in a STEM-oriented high school, gender differences were still found in intentions to choose a humanistic major, grade in Italian and identification with humanities. Finally, the difference in the strength of implicit gender stereotypes disappeared among students in STEM-oriented schools, while it remained statistically significant among students in humanities-oriented schools.

This suggests that controlling for the type of belonging school is necessary whenever assessing the association between moderating and independent variables on major choice.

Table 4.6 *t-test of gender differences by type of school*

	Male		Female		t[a]	p-value[b]
	Mean	SD	Mean	SD		
Humanities-oriented school						
Intention of majoring in STEM	.20	.4	.15	.4	.71	.239
Math grade	6.9	.9	6.8	.9	.86	.194
Science grade	6.9	.9	7.3	1.0	-1.66	.005
Interest in STEM subjects	.33	.9	.23	.9	1.16	.124
Identification with science	2.5	1.1	2.8	.9	-1.15	.125
Identification with math	3.1	1.0	2.8	.9	1.26	.105
Relevance of job's salary on choice	.2	.4	.13	.3	1.05	.148
Intention of majoring in humanities	.27	.5	.37	.5	-1.11	.135
Italian grade	6.8	.8	7.4	1.0	-2.97	.002
Interest in humanistic subjects	.73	.5	.65	.5	0.89	.186
Identification with humanities	3.6	.7	3.7	.8	-.65	.259
Relevance of job's social utility on choice	.37	.5	.44	.5	.32	.5

IAT D score	.17	.5	.37	.5	-2.17	.016
STEM-oriented school						
Intention of majoring in STEM	.6	.5	.5	.5	.82	.207
Math grade	6.8	1.1	7.2	1.2	-1.92	.028
Science grade	6.9	.9	7.1	1.3	-1.21	.115
Interest in STEM subjects	.86	.4	.74	.5	1.63	.052
Identification with Science	3.3	.8	3.1	.9	.9	.184
Identification with Math	3.6	.8	3.6	.7	.08	.468
Relevance of job's salary on choice	.4	.5	.17	.4	2.72	.004
Intention of majoring in humanities	.11	.3	.24	.4	-1.91	.029
Italian grade	6.9	.7	7.3	1.2	-2.58	.006
Interest in humanistic subjects	.62	.5	.71	.5	-1.03	.152
Identification with humanities	3.4	.7	3.7	.6	-2.24	.013
Relevance of job's social utility on choice	.23	.4	.25	.4	-.26	.399
IAT D score						

Note. SD = standard deviation; t = t statistic; p = p-value

[a] Humanities-oriented school H_0: $(\text{mean}_{\text{male}} - \text{mean}_{\text{female}}) = 0$; DF = 170; STEM-oriented school H_0: $(\text{mean}_{\text{male}} - \text{mean}_{\text{female}}) = 0$; DF = 130

[b] Pr(T<t) when t<0; Pr(T>t) when t>0

References

OECD. (2020). Girls' and boys' performance in PISA. In PISA 2018 Results (Volume II): Where All Students Can Succeed Retrieved November 19, 2021 (https://www.oecd-ilibrary.org/education/pisa-2018-results-volume-ii_f56f8c26-en).

Implicit association test: items and categories

The implicit association test (IAT) was designed by Greenwald, McGhee, and Schwartz (1998) to measure individual differences in implicit cognition. The test evaluates the association between two dimensions, namely the target and the attribute. Each dimension has two categories so that there is an association between one category of the target and one category of the attribute which is considered compatible and an association between the other category of the target and the other category of the attribute which is considered incompatible. Therefore, differences in implicit cognition are measured as the difference in the time needed to do an association between compatible constructs and the time needed to do an association with incompatible constructs.

In this case, I used the gender-science IAT (Greenwald et al., 2003; Nosek et al., 2009) in which the target are male and female names and the attribute STEM- or humanities-related majors. In this case, the compatible association is that of female names with humanities and male names with STEM, while the incompatible association is that of female names with STEM and male names with humanities. The English version of the test was translated into Italian. Table A1 reports the words used as stimuli (translated into English).

Table 4.7 *Stimuli*

Gender	Majors
Female names: Cecilia, Rachele, Eva, Azzurra, Caterina, Benedetta, Anita	Humanities: Arts, Literature, Philosophy, History, Music, Languages, Art History
Male names: Kevin, Raffaele, Domenico, Stefano, Nicola, Daniel, Luigi	STEM: Physics, Mathematics, Engineering, Astronomy, Geology, IT, Chemistry

The data collection was conducted on the SoSci Survey platform (Leiner, 2019) which includes a module to perform the IAT. The procedure is that suggested in Greenwald et al. (2003), with 40 trials instead of 20 in the fifth block. The number of trials for each block is summarized in Table A2. The sequence of the stimuli within the blocks was randomly varied. The “improved” D-score suggested in Greenwald et al. (2003) was used as a measure of implicit gender stereotypes.

Table 4.8 *Blocks and trials*

Block	Left	Right	Function	Trials
1	Male	Female	Exercise	20
2	STEM	Humanities	Exercise	20
3	Male+STEM	Female+Humanities	Exercise	20
4	Male+STEM	Female+Humanities	Test	40
5	Female	Male	Exercise	40
6	Female+STEM	Male+Humanities	Exercise	20
7	Female+STEM	Male+Humanities	Test	40

References

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Leiner, D. J. (2019). *SoSci Survey* (3.1.06) [Computer software]. https://www.soscisurvey.de

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., Bar-Anan, Y., Bergh, R., Cai, H., Gonsalkorale, K., Kesebir, S., Maliszewski, N., Neto, F., Olli, E., Park, J., Schnabel, K., Shiomura, K., Tulbure, B. T., Wiers, R. W., … Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593–10597. https://doi.org/10.1073/pnas.0809921106

Regression analysis without multiple imputation of missing data

In case of few clusters, clustered standard errors can be downwards biased. The best technique to deal with this issue is to use wild cluster bootstrap standard errors, a type of cluster bootstrap resampling method with asymptotic refinement (Cameron and Miller 2015; Roodman et al. 2019). However, existing commands used in common statistical software are not compatible with those used for multiple imputation of missing data. To solve this issue, the results reported in Chapter 4 use another bootstrap technique suggested in case of few clusters, yet less efficient compared to the wild cluster bootstrap (Cameron and Miller 2015), i.e., the percentile cluster bootstrap standard errors.

To check whether there are differences in the results when we use wild cluster bootstrap standard errors, Table 4.9 reports the results of the two logistic regression models estimated using the original dataset including missing values. Note that in case of a binary dependent variable, residuals are not well-defined, which prevents application of the wild bootstrap. I thus used the adaptation of the wild bootstrap developed by Kline and Santos (2010) and available in the command 'boottest' (Roodman 2015). Results are consistent with those estimated on the imputed dataset.

Table 4.9 *Results from logistic regression analysis (non-imputed data)*

	Estimate	SE	Score bootstrap-t Wald test	
			z	p-value
STEM				
Gender [Female]	.26	.47	.54	.613
IAT D score	.21	1.19	1.69	.122
Gender*IAT D	-1.67	.72	-2.53	.024
School grade in science	.20	.15	1.47	.168
Identification with STEM	.91	.41	3.19	.006
Interest in STEM subjects	1.11.	.56	1.99	.081
Relevance of job's salary	.80	.37	1.99	.061
Mother's education [a] [below high school]	-.14	.58	-.25	.819
Mother's education [a] [high school]	-.52	.54	-.99	.381

Type of school [b] [STEM-oriented]	.82	.43	1.79	.097
Humanities				
Gender [Female]	.59	.39	1.57	.150
IAT D score	-2.27	.78	-2.67	.019
Gender*IAT D	1.41	.38	2.76	. 020
School grade in Italian	-.10	.10	.37	.733
Identification with humanities	.59	.22	3.283	.007
Interest in humanities-related subjects	.11	.32	-1.06	.319
Relevance of job's utility	.19	.24	.87	.423
Mother's education [a] [below high school]	-.12	.45	-.27	.804
Mother's education [a] [high school]	-.64	.27	-2.13	.054
Type of school [b] [STEM-oriented]	-.88	.39	-2.55	.017

Note. Total N = 257; SE = Standard error

[a] Reference category 'Above high school'

[b] Reference category 'Humanities-oriented'

References

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, *50*(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

Kline, P., & Santos, A. (2010). *A Score Based Approach to Wild Bootstrap Inference* (No. w16127; p. w16127). National Bureau of Economic Research. https://doi.org/10.3386/w16127

Roodman, D. (2015). *BOOTTEST: Stata module to provide fast execution of the wild bootstrap with null imposed* [Statistical Software Components S458121]. Boston College Department of Economics.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal: Promoting Communications on Statistics and Stata*, *19*(1), 4–60. https://doi.org/10.1177/1536867X19830877

Chapter 5 Supplementary material

Search strategy

Table 5.4 *Search terms*

Database	Search strategy
PsycInfo	((‘role model’ OR ‘role modeling’ OR ‘role models’) AND (gender OR m#n OR wom#n OR masculin* OR feminin* OR male* OR female* OR sex OR girl* OR boy*) AND (STEM OR stem OR math* OR scien*)).mp
Web of Science	TS=((‘role model’ OR ‘role modeling’ OR ‘role models’) AND (gender OR m?n OR wom?n OR masculin* OR feminin* OR male* OR female* OR sex OR girl* OR boy*) AND (STEM OR stem OR math* OR scien*))
Scopus	(TITLE-ABS-KEY((‘role model’ OR ‘role modeling’ OR ‘role models’) AND (gender OR m#n OR wom#n OR masculin* OR feminin* OR male* OR female* OR sex OR girl* OR boy*) AND (STEM OR stem OR math* OR scien*)))
Proquest	AB,TI,SU((‘role model’ OR ‘role modeling’ OR ‘role models’) AND (gender OR m#n OR wom#n OR masculin* OR feminin* OR male* OR female* OR sex OR girl* OR boy*) AND (STEM OR stem OR math* OR scien*))

Sources included in the review by type of publication

Table 5.5 *List of studies included in the review*

Type of publication	Reference
Article (n = 54)	Bagès et al., 2016; Bagès & Martinot, 2011; Bamberger, 2014; Baylor et al., 2006; Betz & Sekaquaptewa, 2012; Bhatia & Amati, 2010; Breda et al., 2020; Buck et al., 2002; Cheryan et al., 2011, 2013; Clark et al., 2016; Davies et al., 2002; Dennehy & Dasgupta, 2017; Evans et al., 1995; González-Pérez et al., 2020; Good et al., 2010; Hernandez et al., 2017, 2018; Herrmann et al., 2016; Hoffman & Kurtz-Costes, 2018; Jethwani et al., 2017; LaCosse et al., 2016; Lamers & Mason, 2018; Lewis et al., 2019; Luong & Knobloch-Westerwick, 2017; Marx et al., 2005, 2013; Marx & Ko, 2012; Marx & Roman, 2002; Merritt et al., 2021; Morin-Messabel et al., 2017; O'Brien et al., 2017; Pietri, Drawbaugh, et al., 2021; Pietri et al., 2018; Pietri, Johnson, et al., 2021; Plant et al., 2009; Ramsey et al., 2013; Redmond & Gutke, 2019; Reid et al., 2017; Rule et al., 2019; Shapiro et al., 2013; Single et al., 2005; Stoeger et al., 2013, 2016; Stout et al., 2011, 2017; Van Camp et al., 2019; Van Loo & Rydell, 2014; Walton et al., 2015; Wei et al., 2018; Wyss et al., 2012; Yanowitz & Vanderpool, 2004; Ziegler & Stoeger, 2008; Zurn-Birkhimer & Serrano Anazco, 2018
Ph.D. dissertation (n = 9)	Bailer, 1998; Cherchiglia, 2019; Evans, 1992; Gilbert, 2015; Granville, 1985; Rudy, 1981; Scott, 2013; Thiem, 2016; Wessels, 1987
MA thesis (n = 3)	Howard, 2015; Lawner, 2014; Van Raden, 2011
Conference paper (n = 1)	Nickerson et al., 2017
Report (n = 1)	Bennett et al., 1998

References

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*(1), 55–64. https://doi.org/10.1177/0361684315608842

Bailer, J. (1998). The effects of 'Women Are Scientists, Too' program on middle school students' perceptions of scientists and their attitudes toward women in science [Ed.D., University of Houston]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/effects-women-are-scientists-too-program-on/docview/304436212/se-2?accountid=12459

Bamberger, Y. M. (2014). Encouraging Girls into Science and Technology with Feminine Role Model: Does This Work? *Journal of Science Education and Technology*, *23*(4), 549–561. https://doi.org/10.1007/s10956-014-9487-7

Baylor, A. L., Rosenberg-Kima, R. B., & Plant, E. A. (2006). *Interface agents as social models: The impact of appearance on females' attitude toward engineering*. 526. http://dl.acm.org/citation.cfm?doid=1125451.1125564

Bennett, D., Tsikalas, K., Hupert, N., Meade, T., & Honey, M. (1998). *The benefits of online mentoring for high school girls: Telementoring young women in science, engineering and computing project—Year 3 Evaluation*. Center for children and technology. http://cct.edc.org/sites/cct.edc.org/files/publications/telement_bomhsg98.pdf

Betz, D. E., & Sekaquaptewa, D. (2012). My fair physicist? Feminine math and science role models demotivate young girls. *Social Psychological and Personality Science*, *3*(6), 738–746. https://doi.org/10.1177/1948550612440735

Bhatia, S., & Amati, J. P. (2010). "If These Women Can Do It, I Can Do It, Too": Building Women Engineering Leaders through Graduate Peer Mentoring. *Leadership and Management in Engineering*, *10*(4), 174–184. https://doi.org/10.1061/(ASCE)LM.1943-5630.0000081

Breda, T., Grenet, J., Monnet, M., & Effenterre, C. V. (2020). *Do Female Role Models Reduce the Gender Gap in Science? Evidence from French High Schools* (IZA Discussion Papers No. 13163). Institute of Labor Economics (IZA). http://hdl.handle.net/10419/216475

Buck, G. A., Leslie-Pelecky, D., & Kirby, S. K. (2002). Bringing Female Scientists into the Elementary Classroom: Confronting the Strength of Elementary Students' Stereotypical Images of Scientists. *Journal of Elementary Science Education*, *14*(2), 1–9.

Cherchiglia, L. L. (2019). The Creation of STEM Role Models: An Exploratory Study on the Design of Mentoring Characters for a STEM Gaming Website [Ph.D., Michigan State University]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/creation-stem-role-models-exploratory-study-on/docview/2300630345/se-2?accountid=12459

Cheryan, S., Drury, B. J., & Vichayapai, M. (2013). *Enduring Influence of Stereotypical Computer Science Role Models on Women's Academic Aspirations*. https://doi.org/10.1177/0361684312459328

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*(6), 656–664. https://doi.org/10.1177/1948550611405218

Clark, E. K., Fuesting, M. A., & Diekman, A. B. (2016). Enhancing interest in science: Exemplars as cues to communal affordances of science. *Journal of Applied Social Psychology*, *46*(11), 641–654. https://doi.org/10.1111/jasp.12392

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming Images: How Television Commercials that Elicit Stereotype Threat Can Restrain Women Academically and Professionally. *Personality and Social Psychology Bulletin*, *28*(12), 1615–1628. https://doi.org/10.1177/014616702237644

Dennehy, T. C., & Dasgupta, N. (2017). Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences*, *114*(23), 5964–5969. https://doi.org/10.1073/pnas.1613117114

Evans, M. A. (1992). *Changing attitudes toward science and women in science: Assessing the impact of a role model intervention on ninth grade students* (p. 6371779) [Doctor of Philosophy, Iowa State University, Digital Repository]. https://doi.org/10.31274/rtd-180813-9471

Evans, M. A., Whigham, M., & Wang, M. C. (1995). The effect of a role model project upon the attitudes of ninth-grade science students. *Journal of Research in Science Teaching*, *32*(2), 195–204. https://doi.org/10.1002/tea.3660320208

Gilbert, P. N. (2015). *The role of role models: How does identification with STEM role models impact women's implicit STEM stereotypes and STEM outcomes?* [Ph.D. Dissertation, Tulane University]. https://digitallibrary.tulane.edu/islandora/object/tulane%3A27945

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, J. J., Woodzicka, J. A., & Wingfield, L. C. (2010). The Effects of Gender Stereotypic and Counter-Stereotypic Textbook Images on Science Performance. *The Journal of Social Psychology*, *150*(2), 132–147. https://doi.org/10.1080/00224540903366552

Granville, R. C. (1985). The influence of science role models on the attitudes of middle school students toward women in science [Educat.D., University of Miami]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/influence-science-role-models-on-attitudes-middle/docview/303419629/se-2?accountid=12459

Hernandez, P. R., Bloodhart, B., Adams, A. S., Barnes, R. T., Burt, M., Clinton, S. M., Du, W., Godfrey, E., Henderson, H., Pollack, I. B., & Fischer, E. V. (2018). Role modeling is a viable retention strategy for undergraduate women in the geosciences. *Geosphere*. https://doi.org/10.1130/GES01659.1

Hernandez, P. R., Bloodhart, B., Barnes, R. T., Adams, A. S., Clinton, S. M., Pollack, I., Godfrey, E., Burt, M., & Fischer, E. V. (2017). Promoting professional identity, motivation, and persistence: Benefits of an informal mentoring program for female undergraduate students. *PLOS ONE*, *12*(11), e0187531. https://doi.org/10.1371/journal.pone.0187531

Herrmann, S. D., Adelman, R. M., Bodford, J. E., Graudejus, O., Okun, M. A., & Kwan, V. S. Y. (2016). The Effects of a Female Role Model on Academic Performance and Persistence of Women in STEM Courses. *Basic and Applied Social Psychology*, *38*(5), 258–268. https://doi.org/10.1080/01973533.2016.1209757

Hoffman, A. J., & Kurtz-Costes, B. (2018). Promoting science motivation in American Indian middle school students: An intervention. *Educational Psychology*, *39*(4), 448–469.

Howard, C. (2015). Do STEM fields need a makeover?: The effect of role model femininity on men and women's interest in STEM [M.A., San Diego State University]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/do-stem-fields-need-makeover-effect-role-model/docview/1726897936/se-2?accountid=12459

Jethwani, M. M., Memon, N., Seo, W., & Richer, A. (2017). "I Can Actually Be a Super Sleuth": Promising Practices for Engaging Adolescent Girls in Cybersecurity Education. *Journal of Educational Computing Research*, *55*(1), 3–25. https://doi.org/10.1177/0735633116651971

LaCosse, J., Sekaquaptewa, D., & Bennett, J. (2016). STEM Stereotypic Attribution Bias Among Women in an Unwelcoming Science Setting. *Psychology of Women Quarterly*, *40*(3), 378–397. https://doi.org/10.1177/0361684316630965

Lamers, D., & Mason, R. (2018). Advertising CS/IT degrees to female students in Australia. *Proceedings of the 20th Australasian Computing Education Conference*, 1–8. https://doi.org/10.1145/3160489.3160497

Lawner, E. K. (2014). *Impact of Role Model Gender and Communality on College Women's Math Performance and Interest in STEM* [University of Connecticut]. https://opencommons.uconn.edu/gs_theses/688

Lewis, N. A., Sekaquaptewa, D., & Meadows, L. A. (2019). Modeling gender counter-stereotypic group behavior: A brief video intervention reduces participation gender gaps on STEM teams. *Social Psychology of Education*. https://doi.org/10.1007/s11218-019-09489-3

Luong, K. T., & Knobloch-Westerwick, S. (2017). Can the Media Help Women Be Better at Math? Stereotype Threat, Selective Exposure, Media Effects, and Women's Math

Performance: Media and Stereotype Threat. *Human Communication Research*, *43*(2), 193–213. https://doi.org/10.1111/hcre.12101

Marx, D. M., & Ko, S. J. (2012). Superstars "like" me: The effect of role model similarity on performance under threat: Superstars "like" me. *European Journal of Social Psychology*, *42*(7), 807–812. https://doi.org/10.1002/ejsp.1907

Marx, D. M., Monroe, A. H., Cole, C. E., & Gilbert, P. N. (2013). No Doubt About It: When Doubtful Role Models Undermine Men's and Women's Math Performance Under Threat. *The Journal of Social Psychology*, *153*(5), 542–559. https://doi.org/10.1080/00224545.2013.778811

Marx, D. M., & Roman, J. S. (2002). Female Role Models: Protecting Women's Math Test Performance. *Personality and Social Psychology Bulletin*, *28*(9), 1183–1193. https://doi.org/10.1177/01461672022812004

Marx, D. M., Stapel, D. A., & Muller, D. (2005). We Can Do It: The Interplay of Construal Orientation and Social Comparisons Under Threat. *Journal of Personality and Social Psychology*, *88*(3), 432–446. https://doi.org/10.1037/0022-3514.88.3.432

Merritt, S. K., Hitti, A., Van Camp, A. R., Shaffer, E., Sanchez, M. H., & O'Brien, L. T. (2021). Maximizing the impact of exposure to scientific role models: Testing an intervention to increase science identity among adolescent girls. *Journal of Applied Social Psychology*, jasp.12774. https://doi.org/10.1111/jasp.12774

Morin-Messabel, C., Ferrière, S., Martinez, F., Devif, J., & Reeb, L. (2017). Counter-stereotypes and images: An exploratory research and some questions. *Social Psychology of Education*, *20*(1), 1–13. https://doi.org/10.1007/s11218-017-9370-5

Nickerson, S., Bjorkman Sei, K., Ko, J., & Marx, D. M. (2017). Identification Matters: Effects of Female Peer Role Models Differ By Gender Between High and Low Mathematically Identified Students. *Proceedings of the 20th Annual Conference on Research in Undergraduate Mathematics Education*, 6. https://www.researchgate.net/publication/333489183_Identification_Matters_Effects_of_Female_Peer_Role_Models_Differ_By_Gender_Between_High_and_Low_Mathematically_Identified_Students

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*(3), 301–309. https://doi.org/10.1177/1948550616671997

Pietri, E. S., Drawbaugh, M. L., Johnson, I. R., & Colvin, V. E. (2021). Maybe He Is Relatable Too: Encouraging Women to Identify With Male Scientists by Highlighting Bias Against Fathers. *Psychology of Women Quarterly*, *45*(1), 79–105. https://doi.org/10.1177/0361684320972118

Pietri, E. S., Johnson, I. R., Majid, S., & Chu, C. (2021). Seeing What’s Possible: Videos are more Effective than Written Portrayals for Enhancing the Relatability of Scientists and Promoting Black Female Students’ Interest in STEM. *Sex Roles*, *84*(1–2), 14–33. https://doi.org/10.1007/s11199-020-01153-x

Pietri, E. S., Johnson, I. R., Ozgumus, E., & Young, A. I. (2018). Maybe She Is Relatable: Increasing Women’s Awareness of Gender Bias Encourages Their Identification With Women Scientists. *Psychology of Women Quarterly*, *42*(2), 192–219. https://doi.org/10.1177/0361684317752643

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*(2), 209–215. https://doi.org/10.1016/j.compedu.2009.01.013

Ramsey, L. R., Betz, D. E., & Sekaquaptewa, D. (2013). The effects of an academic environment intervention on science identification among women in STEM. *Social Psychology of Education*, *16*(3), 377–397. https://doi.org/10.1007/s11218-013-9218-6

Redmond, P., & Gutke, H. (2019). STEMming the Flow: Supporting Females in STEM. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-019-09963-6

Reid, J., Smith, E., Iamsuk, N., & Miller, J. (2017). Balancing the Equation: Mentoring First-Year Female STEM Students at a Regional University. *International Journal of*

Innovation in Science and Mathematics Education (Formerly CAL-Laborate International), *24*(4), 18–30. https://openjournals.library.sydney.edu.au/index.php/CAL/article/view/10707

Rudy, M. J. K. (1981). *Gender, Sex-Role Stereotypes, Interest In Science, And Responses To Science/Technology Role Models In The Television Science Series '3-2-1 Contact'* [Ph.D., New York University]. https://www.proquest.com/dissertations-theses/gender-sex-role-stereotypes-interest-science/docview/303161625/se-2?accountid=12459

Rule, A. C., Blaine, D. A., Edwards, C. M., & Gordon, M. (2019). Girls Explore Math Careers by Making 3D-Constructions of Diverse Women Mathematicians' Lives. *The Mathematics Enthusiast*, *16*(1). https://scholarworks.umt.edu/tme/vol16/iss1/27

Scott, J. R. (2013). *Einstein girls: Exploring STEM careers, interest, and identity in an online mentoring community* [Ph.D. Dissertation, University of Florida]. https://ui.adsabs.harvard.edu/abs/2013PhDT.......600S/abstract

Shapiro, J. R., Williams, A. M., & Hambarchyan, M. (2013). Are All Interventions Created Equal? A Multi-Threat Approach to Tailoring Stereotype Threat Interventions. *Journal of Personality and Social Psychology*, *104*(2), 277–288. https://doi.org/10.1037/a0030461

Single, P. B., Muller, C. B., Cunningham, C. M., Single, R. M., & Carlsen, W. S. (2005). Mentornet: E-mentoring For Women Students In Engineering And Science. *Journal of Women and Minorities in Science and Engineering*, *11*(3), 295–310. https://doi.org/10.1615/JWomenMinorScienEng.v11.i3.60

Stoeger, H., Duan, X., Schirner, S., Greindl, T., & Ziegler, A. (2013). The effectiveness of a one-year online mentoring program for girls in STEM. *Computers & Education*, *69*, 408–418. https://doi.org/10.1016/j.compedu.2013.07.032

Stoeger, H., Hopp, M., & Ziegler, A. (2017). Online Mentoring as an Extracurricular Measure to Encourage Talented Girls in STEM (Science, Technology, Engineering, and Mathematics): An Empirical Study of One-on-One Versus Group Mentoring. *Gifted Child Quarterly*, *61*(3), 239–249. https://doi.org/10.1177/0016986217702215

Stoeger, H., Schirner, S., Laemmle, L., Obergriesser, S., Heilemann, M., & Ziegler, A. (2016). A contextual perspective on talented female participants and their development in extracurricular STEM programs: Talented females in an extracurricular STEM program.

Annals of the New York Academy of Sciences, *1377*(1), 53–66. https://doi.org/10.1111/nyas.13116

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*(2), 255–270. https://doi.org/10.1037/a0021385

Stout, J. G., Tamer, B., Wright, H. M., Clarke, L. A., Dwarkadas, S., & Howard, A. M. (2017). The Grad Cohort Workshop: Evaluating an Intervention to Retain Women Graduate Students in Computing. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.02071

Thiem, K. (2016). Gender stereotypes and academic performance: The influence of salient role models on stereotype validation [Ph.D., The University of Iowa]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/gender-stereotypes-academic-performance-influence/docview/1878893949/se-2?accountid=12459

Van Camp, A. R., Gilbert, P. N., & O'Brien, L. T. (2019). Testing the effects of a role model intervention on women's STEM outcomes. *Social Psychology of Education*, *22*(3), 649–671. https://doi.org/10.1007/s11218-019-09498-2

Van Loo, K. J., & Rydell, R. J. (2014). Negative Exposure: Watching Another Woman Subjected to Dominant Male Behavior During a Math Interaction Can Induce Stereotype Threat. *Social Psychological and Personality Science*, *5*(5), 601–607. https://doi.org/10.1177/1948550613511501

Van Raden, S. J. (2011). The Effect of Role Models on the Attitudes and Career Choices of Female Students Enrolled in High School Science [M.S.T., Portland State University]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/effect-role-models-on-attitudes-career-choices/docview/915678364/se-2?accountid=12459

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology*, *107*(2), 468–485. https://doi.org/10.1037/a0037461

Wei, B., Strage, A., & Rhee, J. (2018). Silicon Valley Women in Engineering Conference—Creating Community and Nurturing Engineering Identity. *2018 IEEE Frontiers in Education Conference (FIE)*, 1–4. https://doi.org/10.1109/FIE.2018.8659187

Wessels, J. S. (1987). Effects Of Two Audiovisual Presentations On Motivating Ninth Grade Girls To Pursue Scientific, Technical Careers [Ph.D., University of Denver]. In *ProQuest Dissertations and Theses*. https://www.proquest.com/dissertations-theses/effects-two-audiovisual-presentations-on/docview/303473817/se-2?accountid=12459

Wyss, V. L., Heulskamp, D., & Siebert, C. J. (2012). Increasing Middle School Student Interest in STEM Careers with Videos of Scientists. *International Journal of Environmental and Science Education*, *7*(4), 501–522.

Yanowitz, K. L., & Vanderpool, S. S. (2004). Assessing Girls' Reactions to Science Workshops. *Journal of Science Education and Technology*, *13*(3), 353–359. JSTOR.

Ziegler, A., & Stoeger, H. (2008). Effects of Role Models from Films on Short-Term Ratings of Intent, Interest, and Self-Assessment of Ability by High School Youth: A Study of Gender-Stereotyped Academic Subjects. *Psychological Reports*, *102*(2), 509–531. https://doi.org/10.2466/pr0.102.2.509-531

Zurn-Birkhimer, S., & Serrano Anazco, M. I. (2018). *Impact of Non-technical Conferences in Female Engineering Students' Self- esteem and Engineering Self-efficacy*. ASEE Annual Conference & Exposition, Salt Lake City, UT.

Sources excluded after full-text review

Table 5.6 *Sources excluded and reason*

Reference	Reason for exclusion
Duyilemi, 2008; Johnson, 1989; Song et al., 2020; Woodcock, 2012	Impossible to retrieve/Unpublished
Abedi et al., 2010; Adams et al., 2006; Ashby & Wittmaier, 1978; Bailey & Nihlen, 1990; Behar-Horenstein et al., 2010; Blake-Beard et al., 2011; Bowles et al., 2010; Clarke-Midura et al., 2018; Crisp et al., 2018; Dahlberg et al., 2011; Downing et al., 2005; Farland-Smith, 2009; Finzel et al., 2018; Golan & You, 2021; Good et al., 2003; Hackett et al., 1989; Heilman et al., 2004; Hondred et al., 2012; Kalokerinos et al., 2017; Kao & Harrell, 2015; Kardash, 2000; Kier, 2013; Kunda & Oleson, 1995; Leblebicioglu et al., 2011; Liu et al., 2014; Lockwood, 2006; Luong et al., 2020; McIntyre et al., 2003, 2005, 2011; Mills & Katzman, 2015; Muller & Barsion, 2003; Murphy et al., 2007; Nathanson et al., 2002; Neve & Keith-Marsoun, 2017; Nhundu, 2007; Parks-Stamm et al., 2008; Perkins et al., 2018; Pietri et al., 2017; Quimby & DeSantis, 2006; Rosenthal et al., 2013; Savenye, 1990; Schinske et al., 2016; Scott & Feldman-Summers, 1979; Shin, 2016; Shin et al., 2016; Steinke et al., 2006; Syed et al., 2012; Van Buren et al., 1993; Weisgram & Bigler, 2007	No (or not only) gender-science stereotypes
Aish et al., 2017; Banchefsky et al., 2016; Bennett & Sekaquaptewa, 2014; Black et al., 2011; Bloor et al., 2007; Bonnot & Croizet, 2011; Boucher & Murphy, 2017; Bruchmann & Evans, 2019; Buck et al., 2008; Buck, 2008; Campbell & Skoog, 2004; Carnes et al., 2015; Cheryan et al., 2013; Chesler & Chesler, 2002; Coyle & Liben, 2016; Cozza, 2011; Dar-Nimrod & Heine, 2006; Dasgupta, 2011; Dasgupta et al., 2015; Dawson et al., 2015; Delaney, 2019; Diekman et al., 2011; Farland-Smith, 2015; Fogg-Rogers et al., 2017; Froschl & Sprung, 2013; Fuesting & Diekman, 2017; Ganley et al., 2013;	No effect of role models

Grande & Daniels, 2017; Gutierrez, 2018; Hartman & Hartman, 2008; He et al., 2016; Huguet & Régner, 2007; Kekelis, 2017; Kelleher, 2006; Kemmelmeier & Oyserman, 2001; Kerr & Robinson Kurpius, 2004; Kissane et al., 2015; Kitzinger et al., 2008; Ko et al., 2020; Koch & Gorges, 2016; Krämer et al., 2016; Liben & Coyle, 2014; Logel et al., 2009; Marx & Stapel, 2006; Master et al., 2017; Morgenroth et al., 2015; Nauta & Kokaly, 2001; Newbill, 2005; Park et al., 2018; Quinn & Spencer, 2001; Rhoton, 2011; Richman et al., 2011; Robinson, 2007; Schimke et al., 2007; Scholl & Fuhrmann, 2018; Schuster & Martiny, 2017; Settles et al., 2007; Shaffer et al., 2013; Shull & Weiner, 2002; J. L. Smith et al., 2007; Spieler et al., 2020; Steinke, 1999; Steinke et al., 2021; Szelényi et al., 2013; Szelényi & Inkelas, 2011; Weisgram & Bigler, 2006; Xu & Martin, 2011; Yoder & Schleicher, 1996; Zhao et al., 2018	
Benitez-Herrera et al., 2019; Bobbie et al., 2004; Bodzin & Gehringer, 2001; Canali et al., 2019; Chandler & Parsons, 1995; Chen et al., 2011; Chen et al., 2020; Chesler et al., 2003; Conner & Danielson, 2016; Demetry et al., 2009; Demetry & Sontgerath, 2017, 2013; Dewar, 2017; Egbue et al., 2015; Gaspard et al., 2015; Hochstein et al., 2009; Holmes et al., 2012; Kelly et al., 2013; Kitzinger et al., 2008; Lang et al., 2010; Laursen et al., 2007; Lofton, 2017; Machina & Gokhale, 2015; Marcu et al., 2010; Miel et al., 2018; Nama Namakshi, 2016; Pietri et al., 2019; Rattaro et al., 2020; Schmidt et al., 2003; Smith & Erb, 1986; Stoeger et al., 2017; Todd, 2015; Tyler-Wood et al., 2012; Veltri et al., 2010; Ventresca, 2020; Virnoche & Eschenbach, 2007; Watermeyer, 2012; Weber, 2011	Impossible to disentangle role models' effect
Fried & MacCleave, 2009; Phelan, 2010; Young et al., 2013	Known role models

References

Abedi, F. H., Noorbala, F., & Saeedi, Z. (2010). Stereotypical beliefs about sex role-typed occupations and the mediating effect of exposure to counterstereotypic examples in Iranian students. *Procedia - Social and Behavioral Sciences*, *5*, 1425–1428. https://doi.org/10.1016/j.sbspro.2010.07.300

Adams, G., Garcia, D. M., Purdie-Vaughns, V., & Steele, C. M. (2006). The detrimental effects of a suggestion of sexism in an instruction situation. *Journal of Experimental Social Psychology*, *42*(5), 602–615. https://doi.org/10.1016/j.jesp.2005.10.004

Aish, N., Asare, P., & Miskioğlu, E. E. (2017). People like me increasing likelihood of success for underrepresented minorities in STEM by providing realistic and relatable role models. *2017 IEEE Frontiers in Education Conference (FIE)*, 1–4. https://doi.org/10.1109/FIE.2017.8190454

Ashby, M. S., & Wittmaier, B. C. (1978). Attitude changes in children after exposure to stories about women in traditional or nontraditional occupations. *Journal of Educational Psychology*, *70*(6), 945–949. https://doi.org/10.1037/0022-0663.70.6.945

Bailey, B. A., & Nihlen, A. S. (1990). Effect of Experience with Nontraditional Workers on Psychological and Social Dimensions of Occupational Sex-Role Stereotyping by Elementary School Children. *Psychological Reports*, *66*(3_suppl), 1273–1282. https://doi.org/10.2466/pr0.1990.66.3c.1273

Banchefsky, S., Westfall, J., Park, B., & Judd, C. M. (2016). But You Don’t Look Like A Scientist!: Women Scientists with Feminine Appearance are Deemed Less Likely to be Scientists. *Sex Roles; New York*, *75*(3–4), 95–109. http://dx.doi.org.pros.lib.unimi.it/10.1007/s11199-016-0586-1

Behar-Horenstein, L. S., Roberts, K. W., & Dix, A. C. (2010). Mentoring Undergraduate Researchers: An Exploratory Study of Students’ and Professors’ Perceptions. *Mentoring & Tutoring: Partnership in Learning*, *18*(3), 269–291. https://doi.org/10.1080/13611267.2010.492945

Benitez-Herrera, S., Spinelli, P. F., Mano, S., & Paula Germano, A. (2019). Pursuing gender equality in Astronomy in basic education: The case of the project “Girls in the Museum of Astronomy and Related Sciences”. *EPJ Web of Conferences*, *200*, 02010. https://doi.org/10.1051/epjconf/201920002010

Bennett, J., & Sekaquaptewa, D. (2014). *Setting an Egalitarian Social Norm in the Classroom: Improving Attitudes towards Diversity among Male Engineering Students*. *17*(2), 343–355.

Black, J., Curzon, P., Myketiak, C., & McOwan, P. W. (2011). A study in engaging female students in computer science using role models. *Proceedings of the 16th Annual Joint Conference on Innovation and Technology in Computer Science Education - ITiCSE ’11*, 63. https://doi.org/10.1145/1999747.1999768

Blake-Beard, S., Bayne, M. L., Crosby, F. J., & Muller, C. B. (2011). Matching by Race and Gender in Mentoring Relationships: Keeping our Eyes on the Prize: Matching by Race and Gender in Mentoring Relationships. *Journal of Social Issues*, *67*(3), 622–643. https://doi.org/10.1111/j.1540-4560.2011.01717.x

Bloor, A., Krenitsky, L., & Wellenstein, M. J. (2007). Work in progress - women in technology: An initiative to reach females in rural Wisconsin. *2007 37th Annual Frontiers In Education Conference - Global Engineering: Knowledge Without Borders, Opportunities Without Passports*, S2H-14-S2H-15. https://doi.org/10.1109/FIE.2007.4418079

Bobbie, P. O., Uboh, J., & Davis, B. (2004). A project in embedded systems design and development: A partnership with area high school scholars. *34th Annual Frontiers in Education, 2004. FIE 2004.*, 841–844. https://doi.org/10.1109/FIE.2004.1408656

Bodzin, A., & Gehringer, M. (2001). Breaking Science Stereotypes. *Science and Children*, *38*(4), 36–41.

Bonnot, V., & Croizet, J.-C. (2011). Stereotype threat and stereotype endorsement: Their joint influence on women’s math performance. *Revue internationale de psychologie sociale*, *24*(2), 105–120. Cairn.info. https://www.cairn.info/revue-internationale-de-psychologie-sociale-2011-2-page-105.htm

Boucher, K. L., & Murphy, M. C. (2017). Why so few? The role of social identity and situational cues in understanding the underrepresentation of women in STEM fields. In *Self,* *social identity, and education* (K. I. Mavor, M. J. Platow, B. Bizumic, pp. 93–111). Routledge/Taylor & Francis Group.

Bowles, T., Grant, C., Martin, P., & Carpenter, E. (2010). “Advance Eng Girls To Women: An Innovative Engineering Faculty Student Mentoring Summit For Underrepresented Minority (Urm) Girls And Their Mothers”. *2010 Annual Conference & Exposition Proceedings*, 15.1389.1-15.1389.13. https://doi.org/10.18260/1-2--15966

Bruchmann, K., & Evans, M. C. (2019). Comparing to ingroup and outgroup members: Do we assimilate, contrast, or neither? *The Journal of Social Psychology*, *159*(3), 313–327. https://doi.org/10.1080/00224545.2018.1472546

Buck, G. (2008). Science role models for adolescent girls. *Science Scope; Washington*, *32*(4), 40–43. http://search.proquest.com/docview/225971011/citation/68F3D12D92524598PQ/1

Buck, G., Clark, V. L. P., Leslie-Pelecky, D., Lu, Y., & Cerda-Lizarraga, P. (2008). Examining the cognitive processes used by adolescent girls and women scientists in identifying science role models: A feminist approach. *Science Education*, *92*(4), 688–707. https://doi.org/10.1002/sce.20257

Campbell, A. B., & Skoog, G. (2004). Preparing Undergraduate Women for Science Careers: Facilitating Success in Professional Research. *The Journal of College Science Teaching*, *33*, 24–26.

Canali, C., Addabbo, T., & Moumtzi, V. (2019). A Best Practice for Attracting Female Students to Enrol in ICT Studies. *Proceedings of the 2nd International Conference on Gender Research, Roma Tre University, Rome, Italy, 11-12 April 2019*, 121–128.

Carnes, M., Devine, P. G., Baier Manwell, L., Byars-Winston, A., Fine, E., Ford, C. E., Forscher, P., Isaac, C., Kaatz, A., Magua, W., Palta, M., & Sheridan, J. (2015). The Effect of an Intervention to Break the Gender Bias Habit for Faculty at One Institution: A Cluster Randomized, Controlled Trial. *Academic Medicine*, *90*(2), 221–230. https://doi.org/10.1097/ACM.0000000000000552

Chandler, F. T., & Parsons, S. (1995, April 23). *A Case Study of a Mother/Daughter Science Club*. Annual Meeting of the National Association for Research in Science Teaching (San Francisco, CA, April 22-25, 1995). https://eric.ed.gov/?id=ED387331

Chen, C.-F. J., Jiang, A., Litkowski, E., Elia, A. R., Shuen, J. A., Xu, K., Bonhivert, A., Hsu-Kim, H., & Schwartz-Bloom, R. D. (2011). Females Excelling More In Math, Engineering, And Science (Femmes): An After-Schools STEM Program For Girls That Fosters Hands-On Learning And Female-To-Female Mentorship. *Journal of Women and Minorities in Science and Engineering*, *17*(4), 313–324. https://doi.org/10.1615/JWomenMinorScienEng.2011002293

Chen, Y., Chow, S. C. F., & So, W. W. M. (2020). School-STEM professional collaboration to diversify stereotypes and increase interest in STEM careers among primary school students. *Asia Pacific Journal of Education*, 1–18. https://doi.org/10.1080/02188791.2020.1841604

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

Chesler, N. C., & Chesler, M. A. (2002). Gender-informed mentoring strategies for women engineering scholars: On establishing a caring community. *Journal of Engineering Education; Washington*, *91*(1), 49.

Chesler, N. C., Single, P. B., & Mikic, B. (2003). On belay: Peer-mentoring and adventure education for women faculty in engineering. *Journal of Engineering Education; Washington*, *92*(3), 257. https://doi.org/10.1002/J.2168-9830.2003.TB00766.X

Clarke-Midura, J., Poole, F., Pantic, K., Hamilton, M., Sun, C., & Allan, V. (2018). How Near Peer Mentoring Affects Middle School Mentees. *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, 664–669. https://doi.org/10.1145/3159450.3159525

Conner, L. D. C., & Danielson, J. (2016). Scientist role models in the classroom: How important is gender matching? *International Journal of Science Education*, *38*(15), 2414–2430. https://doi.org/10.1080/09500693.2016.1246780

Coyle, E. F., & Liben, L. S. (2016). Affecting Girls' Activity and Job Interests Through Play: The Moderating Roles of Personal Gender Salience and Game Characteristics. *Child Development*, *87*(2), 414–428. https://doi.org/10.1111/cdev.12463

Cozza, M. (2011). Bridging Gender Gaps, Networking in Computer Science. *Gender, Technology and Development*, *15*(2), 319–337. https://doi.org/10.1177/097185241101500207

Crisp, G., Carales, G., Walls, C., & Cassill, A. (2018). A phenomenological study of community college students' experiences with mentor teachers. *Mentoring & Tutoring: Partnership in Learning*, *26*(2), 183–206.

Dahlberg, T., Barnes, T., Buch, K., & Rorrer, A. (2011). The STARS Alliance: Viable Strategies for Broadening Participation in Computing. *ACM Transactions on Computing Education*, *11*(3), 1–25. https://doi.org/10.1145/2037276.2037282

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to Scientific Theories Affects Women's Math Performance. *Science*, *314*(5798), 435–435. https://doi.org/10.1126/science.1131100

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

Dasgupta, N., Scircle, M. M., & Hunsinger, M. (2015). Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. *Proceedings of the National Academy of Sciences*, *112*(16), 4988–4993. https://doi.org/10.1073/pnas.1422822112

Dawson, A. E., Bernstein, B. L., & Bekki, J. M. (2015). Providing the Psychosocial Benefits of Mentoring to Women in STEM: *Career* WISE as an Online Solution: Providing the Psychosocial Benefits of Mentoring to Women in STEM. *New Directions for Higher Education*, *2015*(171), 53–62. https://doi.org/10.1002/he.20142

Delaney, A. (2019). *The influence of a female-centered figured world on the STEM identities of young girls.* [Iowa State University]. https://lib.dr.iastate.edu/etd/17438/

Demetry, C., Hubelbank, J., Blaisdell, S. L., Sontgerath, S., Nicholson, M. E., Rosenthal, E., & Quinn, P. (2009). SUPPORTING YOUNG WOMEN TO ENTER ENGINEERING:

LONG-TERM EFFECTS OF A MIDDLE SCHOOL ENGINEERING OUTREACH PROGRAM FOR GIRLS. *Journal of Women and Minorities in Science and Engineering*, *15*(2), 119–142. https://doi.org/10.1615/JWomenMinorScienEng.v15.i2.20

Demetry, C., & Sontgerath, S. (2013). Does a Middle School Intervention for Girls Have Long-Lasting Differential Effects on Their Perceptions of and Participation in Engineering? (Research to practice). *2013 ASEE Annual Conference & Exposition Proceedings*, 23.441.1-23.441.13. https://doi.org/10.18260/1-2--19455

Demetry, C., & Sontgerath, S. (2017). A middle school engineering outreach program for girls yields STEM undergraduates. *2017 ASEE Annual Conference & Exposition Proceedings*, 27481. https://doi.org/10.18260/1-2--27481

Dewar, J. M. (2017). Women and mathematics: A course and a scholarly investigation. *BSHM Bulletin: Journal of the British Society for the History of Mathematics*, *32*(3), 246–253. https://doi.org/10.1080/17498430.2017.1319160

Diekman, A. B., Clark, E. K., Johnston, A. M., Brown, E. R., & Steinberg, M. (2011). Malleability in communal goals and beliefs influences attraction to stem careers: Evidence for a goal congruity perspective. *Journal of Personality and Social Psychology*, *101*(5), 902–918. https://doi.org/10.1037/a0025199

Downing, R. A., Crosby, F. J., & Blake-Beard, S. (2005). The Perceived Importance of Developmental Relationships on Women Undergraduates' Pursuit of Science. *Psychology of Women Quarterly*, *29*(4), 419–426. https://doi.org/10.1111/j.1471-6402.2005.00242.x

Duyilemi, A. (2008). Role Modelling as a Means of Enhancing Performance of Nigerian Girls in Science,Technology and Mathematics Education. *International Journal of Learning*, *15*(3), 227–234.

Egbue, O., Long, S., & Ng, E.-H. (2015). Charge It! Translating Electric Vehicle Research Results to Engage 7th and 8th Grade Girls. *Journal of Science Education and Technology*, *24*(5), 663–670. https://doi.org/10.1007/s10956-015-9555-7

Farland-Smith, D. (2009). Exploring Middle School Girls' Science Identities: Examining Attitudes and Perceptions of Scientists when Working "Side-by-Side" with Scientists. *School*

Science and Mathematics, *109*(7), 415–427. https://doi.org/10.1111/j.1949-8594.2009.tb17872.x

Farland-Smith, D. (2015). Mentoring Girls in Science: Eight Case Studies of a Science Camp Experience. In *Innovative Professional Development Methods and Strategies for STEM Education* (1st edition, pp. 17–31). IGI Global.

Finzel, B., Deininger, H., & Schmid, U. (2018). From beliefs to intention: Mentoring as an approach to motivate female high school students to enrol in computer science studies. *Proceedings of the 4th Conference on Gender & IT - GenderIT '18*, 251–260. https://doi.org/10.1145/3196839.3196879

Fogg-Rogers, L., Sardo, M. S., & Boushel, C. N. D. (2017). *'Robots Vs Animals': Establishing a Culture of Public Engagement and Female Role Modeling in Engineering Higher Education*. https://doi.org/10.1177/1075547017696169

Fried, T., & MacCleave, A. (2009). Influence of Role Models and Mentors on Female Graduate Students' Choice of Science as a Career. *Alberta Journal of Educational Research*, *55*.

Froschl, M., & Sprung, B. (2013). The Role of Out-of-School Time in Encouraging Girls in STEM. In *Girls and Women in STEM. A Never Ending Story: Vol. Research on Women and Education* (Paperback, p. 258). Information Age Publishing.

Fuesting, M. A., & Diekman, A. B. (2017). Not By Success Alone: Role Models Provide Pathways to Communal Opportunities in STEM. *Personality and Social Psychology Bulletin*, *43*(2), 163–176. https://doi.org/10.1177/0146167216678857

Ganley, C. M., Mingle, L. A., Ryan, A. M., Ryan, K. M., Vasilyeva, M., & Perry, M. (2013). An examination of stereotype threat effects on girls' mathematics performance. *Developmental Psychology*, *49*(10), 1886–1897. https://doi.org/10.1037/a0031412

Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Nagengast, B., & Trautwein, U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology*, *51*(9), 1226–1240. https://doi.org/10.1037/dev0000028

Golan, J., & You, J. (2021). Raising Aspirations of Boys and Girls through Role Models: Evidence from a Field Experiment. *The Journal of Development Studies*, *57*(6), 949–979. https://doi.org/10.1080/00220388.2020.1826445

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, *24*(6), 645–662. https://doi.org/10.1016/j.appdev.2003.09.002

Grande, V., & Daniels, M. (2017). A diversity lens on the last decade of the FIE conference: Role models for the engineering community. *2017 IEEE Frontiers in Education Conference (FIE)*, 1–8. https://doi.org/10.1109/FIE.2017.8190453

Gutierrez, C. (2018). *Project Increasing Interest in STEM for Underrepresented Females Using Historical Vignettes* [M.A. Thesis, California State University]. https://www.proquest.com/docview/2081200047?pq-origsite=gscholar&fromopenview=true

Hackett, G., Esposito, D., & O'Halloran, M. S. (1989). The relationship of role model influences to the career salience and educational and career plans of college women. *Journal of Vocational Behavior*, *35*(2), 164–180. https://doi.org/10.1016/0001-8791(89)90038-9

Hartman, H., & Hartman, M. (2008). How Undergraduate Engineering Students Perceive Women's (and Men's) Problems in Science, Math and Engineering. *Sex Roles; New York*, *58*(3–4), 251–265. http://dx.doi.org.pros.lib.unimi.it/10.1007/s11199-007-9327-9

He, L., Murphy, L., & Luo, J. (2016). *Using Social Media to Promote STEM Education: Matching College Students with Role Models*. Joint European Conference on Machine Learning and Knowledge Discovery in Databases.

Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *The Journal of Applied Psychology*, *89*(3), 416–427. https://doi.org/10.1037/0021-9010.89.3.416

Hochstein, D., Moses, S., & Jones, D. (2009). *Expanding Your Horizons: A Stem Career Conference For 7 Th And 8 Th Grade Girls*. 14.602.1-14.602.13. https://peer.asee.org/expanding-your-horizons-a-stem-career-conference-for-7-th-and-8-th-grade-girls

Holmes, S., Redmond, A., Thomas, J., & High, K. (2012). Girls Helping Girls: Assessing the Influence of College Student Mentors in an Afterschool Engineering Program. *Mentoring & Tutoring: Partnership in Learning*, *20*(1), 137–150. https://doi.org/10.1080/13611267.2012.645604

Hondred, P. R., Leshem, A., & Kessler, M. R. (2012). Iowa State University Symbi GK12 program: A case study of the resident engineer's effects on 8 th graders attitudes toward science and engineering. *2012 ASEE Annual Conference & Exposition*, 25.863. 1-25.863. 9.

Huguet, P., & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology*, *99*(3), 545–560. https://doi.org/10.1037/0022-0663.99.3.545

Johnson, J. (1989). Effects of Successful Female Role Models on Young Women's Attitudes toward Traditionally Male Careers. *Proceedings of Selected Research Papers Presented at the Annual Meeting of the Association for Educational Communications and Technology*.

Kalokerinos, E. K., Kjelsaas, K., Bennetts, S., & Hippel, C. von. (2017). Men in pink collars: Stereotype threat and disengagement among male teachers and child protection workers. *European Journal of Social Psychology*, *47*(5), 553–565. https://doi.org/10.1002/ejsp.2246

Kao, D., & Harrell, D. F. (2015). Exploring the Impact of Role Model Avatars on Game Experience in Educational Games. *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '15*, 571–576. https://doi.org/10.1145/2793107.2810291

Kardash, C. M. (2000). Evaluation of undergraduate research experience: Perceptions of undergraduate interns and their faculty mentors. *Journal of Educational Psychology*, *92*(1), 191–201. https://doi.org/10.1037/0022-0663.92.1.191

Kekelis, L. S. (2017). TECHBRIDGE GIRLS Inspire a Girl to Change the World Through STEM. In *Out-Of-School-Time STEM Programs For Females: Implications For Research And Practice* (pp. 55–79).

Kelleher, C. (2006). *Motivating Programming: Using Storytelling to Make Computer Programming Attractive to Middle School Girls* [Ph.D. Dissertation, Carnegie Mellon University]. https://www.cs.cmu.edu/~caitlin/kelleherThesis_CSD.pdf

Kelly, K., Dampier, D. A., & Carr, K. (2013). WILLING, ABLE, AND UNWANTED: HIGH SCHOOL GIRLS' POTENTIAL SELVES IN COMPUTING. *Journal of Women and Minorities in Science and Engineering*, *19*(1), 67–85. https://doi.org/10.1615/JWomenMinorScienEng.2013004471

Kemmelmeier, M., & Oyserman, D. (2001). The ups and downs of thinking about a successful other: Self-construals and the consequences of social comparisons. *European Journal of Social Psychology*, *31*(3), 311–320. https://doi.org/10.1002/ejsp.47

Kerr, B., & Robinson Kurpius, S. E. (2004). Encouraging talented girls in math and science: Effects of a guidance intervention. *High Ability Studies*, *15*(1), 85–102. https://doi.org/10.1080/1359813042000225357

Kier, M. W. (2013). *Examining the Effects of a STEM Career Video Intervention on the Interests and STEM Professional Identities of Rural, Minority Middle School Students* [Ph.D. Dissertation]. http://search.proquest.com/docview/1459459780/abstract/9F4E9612BE7C4817PQ/1

Kissane, L. M., Fang, L., & Silver, R. (2015). Female Millennial Perceptions of Engineering's 'Brand'. *122nd ASEE Annual Conference & Exposition*. 122nd ASEE Annual Conference & Exposition, Seattle.

Kitzinger, J., Haran, J., Chimba, M., & Boyce, T. (2008). Role Models in the Media: An Exploration of the Views and Experiences of Women in Science, Engineering and Technology. *UK Resource Centre for Women in Science, Engineering and Technology (UKRC) and Cardiff University*. https://www.academia.edu/1624459/Role_Models_in_the_Media_An_Exploration_of_the_Views_and_Experiences_of_Women_in_Science_Engineering_and_Technology

Ko, S. J., Marx, D. M., Nickerson, S. D., & Bjorkman, K. (2020). Implementing a Peer Role Model Program in College Calculus Classes to Broaden Women's Participation in STEM. *PRIMUS*, *30*(3), 349–366. https://doi.org/10.1080/10511970.2019.1590749

Koch, M., & Gorges, T. (2016). Curricular Influences on Female Afterschool Facilitators' Computer Science Interests and Career Choices. *Journal of Science Education and Technology*, *25*(5), 782–794. https://doi.org/10.1007/s10956-016-9636-2

Krämer, N. C., Karacora, B., Lucas, G., Dehghani, M., Rüther, G., & Gratch, J. (2016). Closing the gender gap in STEM with friendly male instructors? On the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers & Education*, *99*, 1–13. https://doi.org/10.1016/j.compedu.2016.04.002

Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, *68*(4), 565–579. https://doi.org/10.1037//0022-3514.68.4.565

Lang, C., Craig, A., Fisher, J., & Forgasz, H. (2010). Creating digital divas: Scaffolding perception change through secondary school and university alliances. *Proceedings of the Fifteenth Annual Conference on Innovation and Technology in Computer Science Education - ITiCSE '10*, 38–42. https://doi.org/10.1145/1822090.1822103

Laursen, S., Liston, C., Thiry, H., & Graf, J. (2007). What Good Is a Scientist in the Classroom? Participant Outcomes and Program Design Features for a Short-Duration Science Outreach Intervention in K–12 Classrooms. *CBE— Life Sciences Education*, *6*(1), 49–64. https://doi.org/10.1187/cbe.06-05-0165

Leblebicioglu, G., Metin, D., Yardimci, E., & Cetin, P. S. (2011). The effect of informal and formal interaction between scientists and children at a science camp on their images of scientists. *Science Education International*, *22*(3), 158–174.

Liben, L. S., & Coyle, E. F. (2014). Developmental Interventions to Address the STEM Gender Gap: Exploring Intended and Unintended Consequences. In *Advances in Child Development and Behavior* (Vol. 47, pp. 77–115). Elsevier. https://doi.org/10.1016/bs.acdb.2014.06.001

Liu, Y., Lou, S., & Shih, R. (2014). The investigation of STEM Self-Efficacy and Professional Commitment to Engineering among female high school students. *South African Journal of Education*, *34*(2), 1–15. https://doi.org/10.15700/201412071216

Lockwood, P. (2006). “Someone Like Me can be Successful”: Do College Students Need Same-Gender Role Models? *Psychology of Women Quarterly*, *30*(1), 36–46. https://doi.org/10.1111/j.1471-6402.2006.00260.x

Lofton, J. J. (2017). Encouraging Young Women to Pursue Engineering: 25 Years of Summer Camp Successes and Challenges. *202017 ASEE Annual Conference & Exposition*, 20. https://doi.org/10.18260/1-2--28222

Logel, C., Walton, G. M., Spencer, S. J., Iserman, E. C., von Hippel, W., & Bell, A. E. (2009). Interacting with sexist men triggers social identity threat among female engineers. *Journal of Personality and Social Psychology*, *96*(6), 1089–1103. https://doi.org/10.1037/a0015703

Luong, K. T., Knobloch-Westerwick, S., & Niewiesk, S. (2020). Superstars within reach: The role of perceived attainability and role congruity in media role models on women’s social comparisons. *Communication Monographs*, *87*(1), 4–24. https://doi.org/10.1080/03637751.2019.1622143

Machina, K., & Gokhale, A. (2015). Interventions for increasing male and female undergraduate interest in information technology. *Computers & Education*, *87*, 277–284. https://doi.org/10.1016/j.compedu.2015.06.008

Marcu, G., Kaufman, S. J., Lee, J. K., Black, R. W., Dourish, P., Hayes, G. R., & Richardson, D. J. (2010). Design and evaluation of a computer science and engineering course for middle school girls. *Proceedings of the 41st ACM Technical Symposium on Computer Science Education - SIGCSE ’10*, 234. https://doi.org/10.1145/1734263.1734344

Marx, D. M., & Stapel, D. A. (2006). Distinguishing stereotype threat from priming effects: On the role of the social self and threat-based concerns. *Journal of Personality and Social Psychology*, *91*(2), 243–254. https://doi.org/10.1037/0022-3514.91.2.243

Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. N. (2017). Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology*, *160*, 92–106. https://doi.org/10.1016/j.jecp.2017.03.013

McIntyre, R. B., Lord, C. G., Gresky, D. M., Eyck, L. L. T., Frye, G. D. J., & Bond Jr., C. F. (2005). A social impact trend in the effects of role models on alleviating women's mathematics stereotype threat. *Current Research in Social Psychology*, *10*.

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*(1), 83–90. https://doi.org/10.1016/S0022-1031(02)00513-9

McIntyre, R. B., Paulson, R. M., Taylor, C. A., Morin, A. L., & Lord, C. G. (2011). Effects of role model deservingness on overcoming performance deficits induced by stereotype threat. *European Journal of Social Psychology*, *41*(3), 301–311. https://doi.org/10.1002/ejsp.774

Miel, K., Portsmore, M. D., Maltese, A. V., & Paul, K. (2018). Board 126: Examining the Interactions Related to Role Modeling in an Elementary Outreach Program (Work in Progress). *2018 ASEE Annual Conference & Exposition*, 8. https://doi.org/10.18260/1-2--29910

Mills, L., & Katzman, W. (2015). Examining the Effects of Field Trips on Science Identity. *12th International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA)*, 7.

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*(4), 465–483. https://doi.org/10.1037/gpr0000059

Muller, C. B., & Barsion, S. J. (2003). Assessment of a Large-Scale E-Mentoring Network for Women in Engineering and Science: Just How Good is MentorNet? *Proceedings of the 2003 WEPAN National Conference*, 6. https://journals.psu.edu/wepan/article/view/58327

Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling Threat: How Situational Cues Affect Women in Math, Science, and Engineering Settings. *Psychological Science*, *18*(10), 879–885. https://doi.org/10.1111/j.1467-9280.2007.01995.x

Namakshi, N. (2016). *Creating A Pathway To STEM: Role Of An Informal Mathematics Program* [Ph.D. Dissertation, Texas State University]. https://digital.library.txstate.edu/handle/10877/6856

Nathanson, A. I., Wilson, B. J., McGee, J., & Sebastian, M. (2002). Counteracting the Effects of Female Stereotypes on Television via Active Mediation. *Journal of Communication*, *52*(4), 922–937. https://doi.org/10.1111/j.1460-2466.2002.tb02581.x

Nauta, M. M., & Kokaly, M. L. (2001). Assessing Role Model Influences on Students' Academic and Vocational Decisions. *Journal of Career Assessment*, *9*(1), 81–99. https://doi.org/10.1177/106907270100900106

Neve, N., & Keith-Marsoun, S. K. K. (2017). The Invention Bootcamp, a Four-Week Summer Course For High School Underrepresented Students in a University Setting. *ASEE Annual Conference & Exposition*, 14. https://doi.org/10.18260/1-2--28983

Newbill, P. L. (2005). *Instructional strategies to improve women's attitudes toward science* [Ph.D. Dissertation, Virginia Polytechnic Institute and State University]. https://ui.adsabs.harvard.edu/abs/2005PhDT........22N/abstract

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*(9–10), 639–649. https://doi.org/10.1007/s11199-007-9204-6

Park, L. E., Kondrak, C. L., Ward, D. E., & Streamer, L. (2018). Positive Feedback From Male Authority Figures Boosts Women's Math Outcomes. *Personality and Social Psychology Bulletin*, *44*(3), 359–383. https://doi.org/10.1177/0146167217741312

Parks-Stamm, E. J., Heilman, M. E., & Hearns, K. A. (2008). Motivated to Penalize: Women's Strategic Rejection of Successful Women. *Personality and Social Psychology Bulletin*, *34*(2), 237–247. https://doi.org/10.1177/0146167207310027

Perkins, H., Wyer, M., & Schinske, J. N. (2018). Evaluation of Effects of an Intervention Aimed at Broadening Participation in STEM while Conveying Science Content. In T. L. S. Kishbaugh & S. G. Cessna (Eds.), *ACS Symposium Series* (Vol. 1301, pp. 59–82). American Chemical Society. https://doi.org/10.1021/bk-2018-1301.ch004

Phelan, J. E. (2010). *Increasing women's aspirations and achievement in science: The effect of role models on implicit cognitions*. https://doi.org/10.7282/T36H4HG4

Pietri, E. S., Hennes, E. P., Dovidio, J. F., Brescoll, V. L., Bailey, A. H., Moss-Racusin, C. A., & Handelsman, J. (2019). Addressing Unintended Consequences of Gender Diversity Interventions on Women's Sense of Belonging in STEM. *Sex Roles*, *80*(9–10), 527–547. https://doi.org/10.1007/s11199-018-0952-2

Pietri, E. S., Moss-Racusin, C. A., Dovidio, J. F., Guha, D., Roussos, G., Brescoll, V. L., & Handelsman, J. (2017). Using Video to Increase Gender Bias Literacy Toward Women in Science. *Psychology of Women Quarterly*, *41*(2), 175–196. https://doi.org/10.1177/0361684316674721

Quimby, J. L., & DeSantis, A. M. (2006). The Influence of Role Models on Women's Career Choices. *The Career Development Quarterly; Alexandria*, *54*(4), 297–306.

Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, *57*(1), 55–71. https://doi.org/10.1111/0022-4537.00201

Rattaro, C., Briozzo, I., Siniscalchi, M., Blasina, F., & Castillo, M. del. (2020). Encouraging Girls in STEM: workshops on analog electronics, sensors and robotics. *2020 XIV Technologies Applied to Electronics Teaching Conference (TAEE)*, 1–5. https://doi.org/10.1109/TAEE46915.2020.9163703

Rhoton, L. A. (2011). Distancing as a Gendered Barrier: Understanding Women Scientists' Gender Practices. *Gender & Society*, *25*(6), 696–716. https://doi.org/10.1177/0891243211422717

Richman, L. S., vanDellen, M., & Wood, W. (2011). How Women Cope: Being a Numerical Minority in a Male-Dominated Profession: Coping with Cues to Social Identity Threat. *Journal of Social Issues*, *67*(3), 492–509. https://doi.org/10.1111/j.1540-4560.2011.01711.x

Robinson, J. (2007). *Closing the race and gender gaps in computer science education* [Ph.D. Dissertation, Graduate School of Rowan University]. https://www.proquest.com/docview/304703095

Rosenthal, L., Levy, S. R., London, B., Lobel, M., & Bazile, C. (2013). In Pursuit of the MD: The Impact of Role Models, Identity Compatibility, and Belonging Among

Undergraduate Women. *Sex Roles*, *68*(7–8), 464–473. https://doi.org/10.1007/s11199-012-0257-9

Savenye, W. C. (1990). Role Models and Student Attitudes toward Nontraditional Careers. *Educational Technology Research and Development*, *38*(3), 5–13. JSTOR. https://www.jstor.org/stable/30218313

Schimke, D., Stoeger, H., & Ziegler, A. (2007). The Relationship Between Social Presence and Group Identification Within Online Communities and Its Impact on the Success of Online Communities. In D. Schuler (Ed.), *Online Communities and Social Computing* (Vol. 4564, pp. 160–168). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-73257-0_18

Schinske, J. N., Perkins, H., Snyder, A., & Wyer, M. (2016). Scientist Spotlight Homework Assignments Shift Students' Stereotypes of Scientists and Enhance Science Identity in a Diverse Introductory Science Class. *CBE Life Sciences Education*, *15*(3). https://doi.org/10.1187/cbe.16-01-0002

Schmidt, J. A., Smith, P. E., Vogt, K. E., & Schmidt, L. C. (2003). Innovative educational opportunities for women in stem: Research internships in science and engineering (rise). *33rd Annual Frontiers in Education, 2003. FIE 2003.*, *2*, F1D_18-F1D_23. https://doi.org/10.1109/FIE.2003.1264676

Scholl, M. C., & Fuhrmann, F. (2018). How Can We Succeed in Attracting (Young) Women to the Promising Careers of Information Security and Computing? – A Project Testing an Experience-Driven Approach. *The 12th International Multi-Conference on) Society, Cybernetics and Informatics*, 116–119. https://www.researchgate.net/publication/327282887_How_Can_We_Succeed_in_Attracting_Young_Women_to_the_Promising_Careers_of_Information_Security_and_Computing_-_A_Project_Testing_an_Experience-Driven_Approach

Schuster, C., & Martiny, S. E. (2017). Not Feeling Good in STEM: Effects of Stereotype Activation and Anticipated Affect on Women's Career Aspirations. *Sex Roles*, *76*(1–2), 40–55. https://doi.org/10.1007/s11199-016-0665-3

Scott, K. P., & Feldman-Summers, S. (1979). Children’s reactions to textbook stories in which females are portrayed in traditionally male roles. *Journal of Educational Psychology*, *71*(3), 396–402. https://doi.org/10.1037/0022-0663.71.3.396

Settles, I. H., Cortina, L. M., Stewart, A. J., & Malley, J. (2007). Voice Matters: Buffering the Impact of a Negative Climate for Women in Science. *Psychology of Women Quarterly*, *31*(3), 270–281. https://doi.org/10.1111/j.1471-6402.2007.00370.x

Shaffer, E. S., Marx, D. M., & Prislin, R. (2013). Mind the Gap: Framing of Women’s Success and Representation in STEM Affects Women’s Math Performance under Threat. *Sex Roles*, *68*(7–8), 454–463. https://doi.org/10.1007/s11199-012-0252-1

Shin, J. E. L. (2016). *Increasing college students’ interest and engagement in STEM: A comparison of strategies for challenging STEM stereotypes* [Ph.D. Dissertation, State University of New York at Stony Brook]. https://www.proquest.com/docview/1882305509

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*(7), 410–427. https://doi.org/10.1111/jasp.12371

Shull, P. J., & Weiner, M. (2002). Thinking Inside the Box: Self-Efficacy of Women in Engineering. *International Journal of Engineering Education*, *18*(4), 438:446.

Smith, J. L., Sansone, C., & White, P. H. (2007). The stereotyped task engagement process: The role of interest and achievement motivation. *Journal of Educational Psychology*, *99*(1), 99–114. https://doi.org/10.1037/0022-0663.99.1.99

Smith, W. S., & Erb, T. O. (1986). Effect of women science career role models on early adolescents’ attitudes toward scientists and women in science. *Journal of Research in Science Teaching*, *23*(8), 667–676. https://doi.org/10.1002/tea.3660230802

Song, J. J., Dow, D. E., Ma, G. G., Li, W., & Ma, L. (2020). STEM program for female high school students. *2020 ASEE Annual Conference and Exposition, Conference Proceedings*, *2020-June*, 9. https://doi.org/10.18260/1-2--35205

Spieler, B., Mikats, J., Valentin, S., Oates-Indruchová, L., & Slany, W. (2020). “RemoteMentor” Evaluation of Interactions Between Teenage Girls, Remote Tutors, and

Coding Activities in School Lessons. In P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies. Designing, Developing and Deploying Learning Experiences* (Vol. 12205, pp. 547–567). Springer International Publishing. https://doi.org/10.1007/978-3-030-50513-4_40

Steinke, J. (1999). Women Scientist Role Models on Screen: A Case Study of Contact. *Science Communication*, *21*(2), 111–136. https://doi.org/10.1177/1075547099021002002

Steinke, J., Applegate, B., Penny, J. R., & Merlino, S. (2021). Effects of Diverse STEM Role Model Videos in Promoting Adolescents' Identification. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-021-10168-z

Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-thomas, A., Williams, Y., Evergreen, S. H., & Kuchibhotla, S. (2006). Assessing media influences on middle school-aged children's perceptions of women in science and engineering using the Draw-AScientist Test (DAST. *In Proceedings of 2006 WEPAN Conference, Edited By*.

Stoeger, H., Hopp, M., & Ziegler, A. (2017). Online Mentoring as an Extracurricular Measure to Encourage Talented Girls in STEM (Science, Technology, Engineering, and Mathematics): An Empirical Study of One-on-One Versus Group Mentoring. *Gifted Child Quarterly*, *61*(3), 239–249. https://doi.org/10.1177/0016986217702215

Syed, M., Goza, B. K., Chemers, M. M., & Zurbriggen, E. L. (2012). Individual Differences in Preferences for Matched-Ethnic Mentors Among High-Achieving Ethnically Diverse Adolescents in STEM. *Child Development*, *83*(3), 896–910. https://doi.org/10.1111/j.1467-8624.2012.01744.x

Szelényi, K., Denson, N., & Inkelas, K. K. (2013). Women in STEM Majors and Professional Outcome Expectations: The Role of Living-Learning Programs and Other College Environments. *Research in Higher Education*, *54*(8), 851–873. https://doi.org/10.1007/s11162-013-9299-2

Szelényi, K., & Inkelas, K. K. (2011). The Role of Living–Learning Programs in Women's Plans to Attend Graduate School in STEM Fields. *Research in Higher Education*, *52*(4), 349–369. https://doi.org/10.1007/s11162-010-9197-9

Todd, B. (2015). *Little Scientists: Identity, Self-efficacy, And Attitudes Toward Science In A Girls' Science Camp* [Ph.D. Dissertation, University of Oregon]. https://ui.adsabs.harvard.edu/abs/2015PhDT.......111T/abstract

Tyler-Wood, T., Ellison, A., Lim, O., & Periathiruvadi, S. (2012). Bringing Up Girls in Science (BUGS): The Effectiveness of an Afterschool Environmental Science Program for Increasing Female Students' Interest in Science Careers. *Journal of Science Education and Technology*, *21*(1), 46–55. https://doi.org/10.1007/s10956-011-9279-2

Van Buren, J. B., Kelly, K. R., & Hall, A. S. (1993). Modeling Nontraditional Career Choices: Effects of Gender and School Location on Response to a Brief Videotape. *Journal of Counseling & Development*, *72*(1), 101–104. https://doi.org/10.1002/j.1556-6676.1993.tb02285.x

Veltri, N., Webb, H., & Papp, R. (2010). GETSMART: An Academic-Industry Partnership to Encourage Female Participation in Science, Technology, Engineering and Math Careers. *Women in Engineering, Science and Technology: Education and Career Challenges*, 56–72. https://doi.org/10.4018/978-1-61520-657-5.ch003

Ventresca, N. (2020). *Encouraging STEM participation among upper elementary females.* [Ph.D. Dissertation, Wilmington University]. https://www.proquest.com/docview/2388065033?pq-origsite=gscholar&fromopenview=true

Virnoche, M., & Eschenbach, E. (2007). *Expanding Girls' Horizons In Math And Science: A Longitudinal Evaluation Of Eyh Conference Outcomes*. 12.707.1-12.707.24. https://peer.asee.org/expanding-girls-horizons-in-math-and-science-a-longitudinal-evaluation-of-eyh-conference-outcomes

Watermeyer, R. (2012). Confirming the legitimacy of female participation in science, technology, engineering and mathematics (STEM): Evaluation of a UK STEM initiative for girls. *British Journal of Sociology of Education*, *33*(5), 679–700. https://doi.org/10.1080/01425692.2012.678751

Weber, K. (2011). Role Models and Informal STEM-Related Activities Positively Impact Female Interest in STEM. *Technology and Engineering Teacher; Reston*, *71*(3), 18–21. http://search.proquest.com/docview/902759039/abstract/66600B2781AF4695PQ/1

Weisgram, E. S., & Bigler, R. S. (2006). Girls and science careers: The role of altruistic values and attitudes about scientific tasks. *Journal of Applied Developmental Psychology*, *27*(4), 326–348. https://doi.org/10.1016/j.appdev.2006.04.004

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*(3), 262–269. https://doi.org/10.1111/j.1471-6402.2007.00369.x

Woodcock, A. (2012). *Implicit ingroup biases and performance under threat: Do successful counterstereotypic exemplars help or harm?* [Ph.D. Dissertation, Purdue University]. https://docs.lib.purdue.edu/dissertations/AAI3544586/

Xu, Y. J., & Martin, C. L. (2011). Gender Differences in STEM Disciplines: From the Aspects of Informal Professional Networking and Faculty Career Development. *Gender Issues*, *28*(3), 134–154. https://doi.org/10.1007/s12147-011-9104-5

Yoder, J. D., & Schleicher, T. L. (1996). Undergraduates regard deviation from occupational gender stereotypes as costly for women. *Sex Roles*, *34*(3), 171–188. https://doi.org/10.1007/BF01544294

Young, D. M., Rudman, L. A., Buettner, H. M., & McLean, M. C. (2013). The Influence of Female Role Models on Women's Implicit Science Cognitions. *Psychology of Women Quarterly*, *37*(3), 283–292. https://doi.org/10.1177/0361684313482109

Zhao, F., Zhang, Y., Alterman, V., Zhang, B., & Yu, G. (2018). Can Math-Gender Stereotypes Be Reduced? A Theory-Based Intervention Program with Adolescent Girls. *Current Psychology*, *37*(3), 612–624. https://doi.org/10.1007/s12144-016-9543-y

Data extraction instrument

Table 5.7 *Data extraction instrument*

Information on source
Citation details Author/s Date Title Journal Volume Issue Pages Type of publication
Information on research type
Research approach Research design
Information on target
Participants’ age Participants’ occupation Participants’ gender Participants’ nationality Sample size
Information on examples
Type (i.e., role model or mentor) Gender Number Profession
Information on intervention
Type Location

Duration

Information on dependent variables

Variable(s)

Instrument(s)

Information on effects

Aim

Effect

Inclusion and exclusion criteria

Table 5.8 *Inclusion and exclusion criteria*

	Inclusion criteria	Exclusion criteria
Participants	All age groups and gender/sex	
Concept	Exposure to typical and/or atypical role models/mentors, where typical and atypical refer to the adherence to the traditional gender stereotype (atypical= female in STEM, male in non-STEM; typical=female in non-STEM, male in STEM)	Role models and mentors should be unknown and thus not part of the family or educators in school (e.g., teachers, professors)
Context	Addressing gender stereotypes on STEM	Not addressing stereotypes or non-STEM context
Types of studies	All (include also grey literature)	

Chapter 6 Supplementary material

Descriptive statistics

On average, students were aware of the disproportion between women and men in the STEM and humanities sectors, with differences between male and female participants (STEM: Chi-squared(3) 19.6422, Pr 0.000; Humanities: Chi-squared(2) 3.0196, Pr 0.221). Female students were more likely to believe that the STEM sector was a prerogative of men. Similarly, they were more likely to believe that the humanities sector was a prerogative of women. In general, the gender gap in the humanities sector was perceived more prominent than the one in the STEM sector.

Table 6.5 *Descriptive statistics - Perceived gender gap by gender*

	Female	Male
STEM sector		
Gender balance	13.09%	28.83%
More men than women	80.73%	69.37%
Almost all men	6.18%	0.90%
Humanities sector		
Gender balance	5.82%	9.91%
More men than women	78.91%	79.28%
Almost all men	15.27%	10.81%

Table 6.6 shows the mean and standard deviation of implicit and explicit gender stereotypes.

Table 6.6 *Descriptive statistics – gender stereotypes*

	Female		Male		t	p-value
	Mean	SD	Mean	SD		
Implicit stereotypical beliefs	.43	.47	.39	.43	1.03	.306
Explicit stereotypical beliefs on scientific abilities	1.91	.89	1.99	1.08	-.77	.442
Explicit stereotypical beliefs on language-related abilities	2.13	.96	2.19	1.12	-.45	.654

Note. SD = Standard deviation; t = t-statistic

On average, participants endorsed weak implicit and explicit stereotypical beliefs. No gender differences emerged. Compared to explicit beliefs on scientific abilities, those on language-related abilities were slightly stronger for both male and female participants.

Finally, I found a different pattern in the attribution of gender differences in STEM compared to the humanities sector, with some differences between male and female participants (STEM: Chi-squared(2) 29.1653, Pr 0.000; Humanities: Chi-squared(2) 4.2389, Pr 0.120). As regards the STEM sector, female students were more likely to agree with the social explanation of gender differences than with the explanation related to gendered individual characteristics. On the contrary, male students' opinion was more heterogeneous. As regards the other sector, for both male and female students, gender-related characteristics were invoked more frequently than discrimination and social pressure to explain gender differences in humanistic studies.

Table 6.7 *Descriptive statistics - Causes attributed to gender differences*

	Female	Male
STEM sector		
Social pressure more relevant than gendered characteristics	54.91%	31.53%
No preference	30.55%	30.63%
Gendered characteristics more relevant than social pressure	14.55%	37.84%
Humanities sector		
Social pressure more relevant than gendered characteristics	16.00%	8.11%
No preference	39.27%	41.44%
Gendered characteristics more relevant than social pressure	44.73%	50.45%

Material

Role models were selected among young (under 30 years old) professionals to increase similarity with participants. The interviews avoided to emphasize extraordinary results of the interviewees and simply provided information on the daily routine of an ordinal working day. More specifically, I asked them:

1. What do you do for a living?
2. What is your job about?
3. What do you like about your job?
4. Which advice would you give to those wishing to undertake this path?

Figure 6.4 shows a frame of Treatment A's video, Figure 6.5 a frame of Treatment B's video and Figure 6.6 a frame of the control group's video.

Figure 6.4 *Frame of the video (Treatment A)*

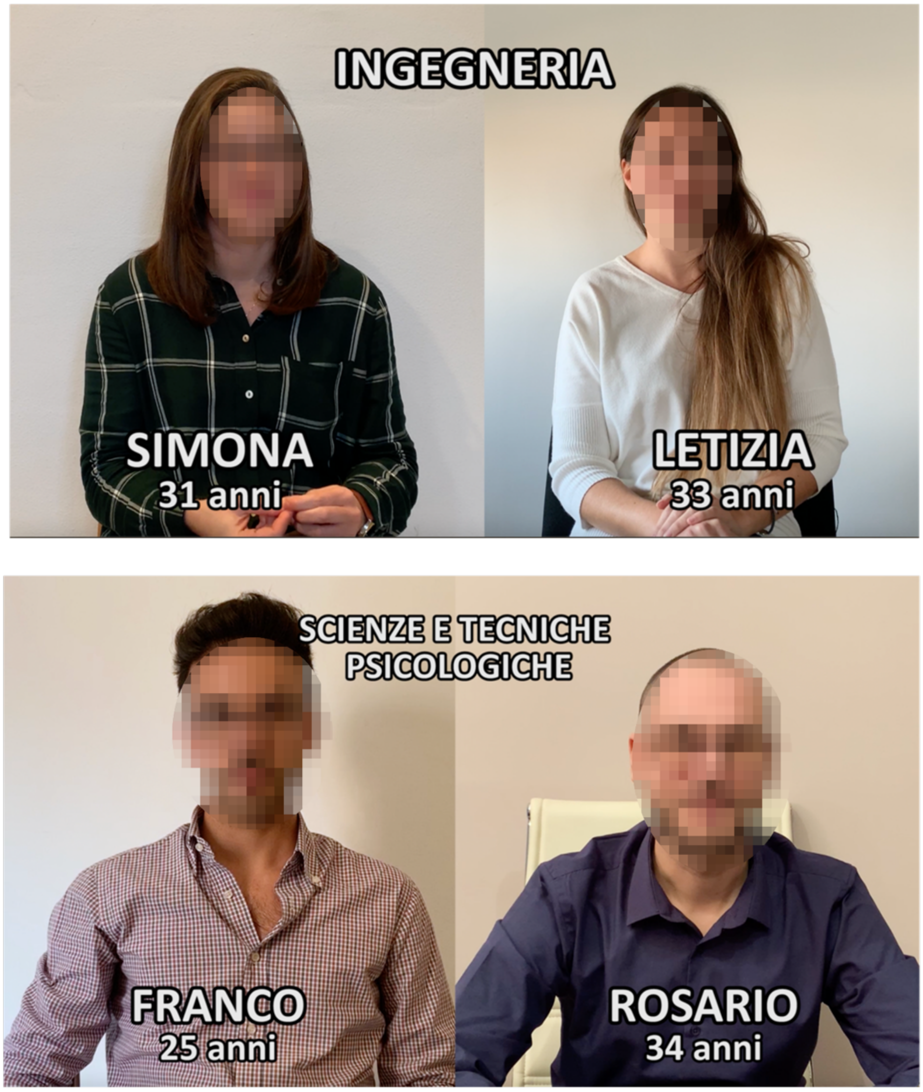


Figure 6.5 *Frame of the video (Treatment B)*

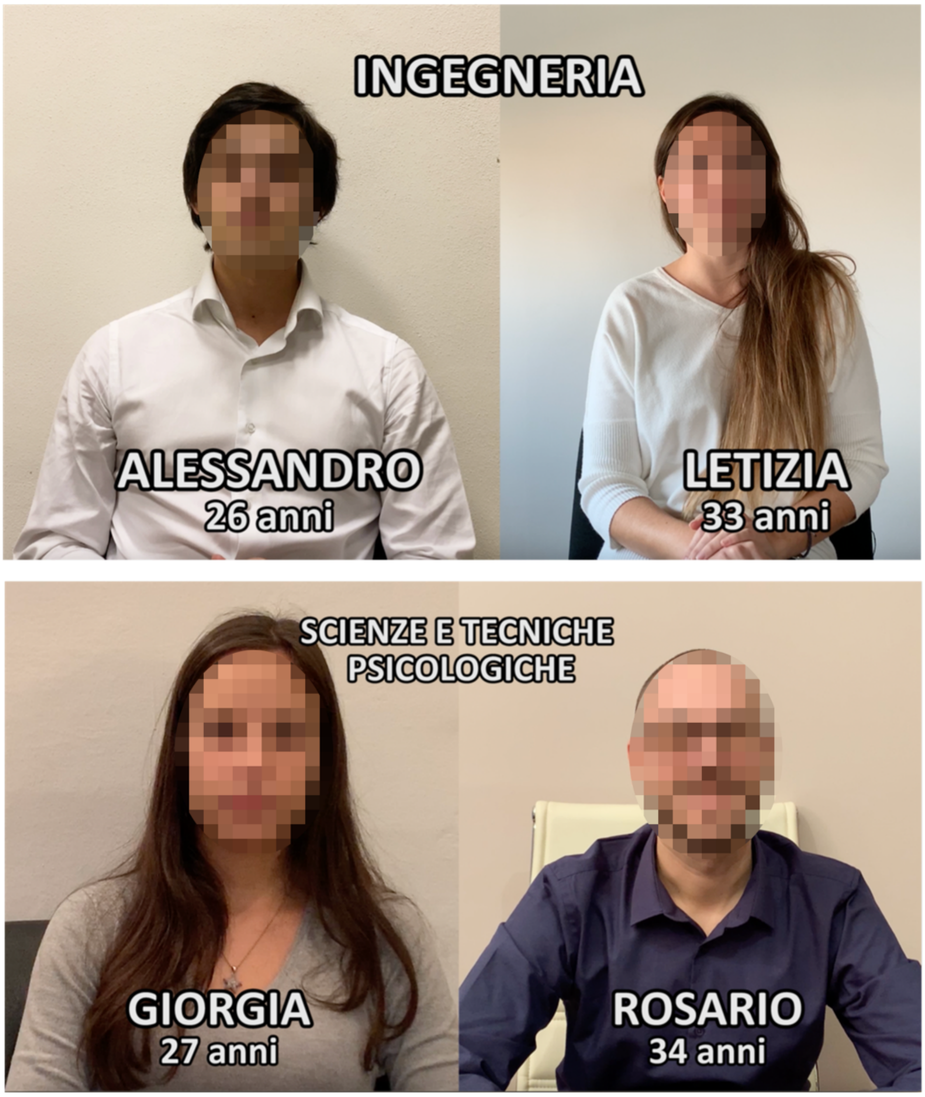
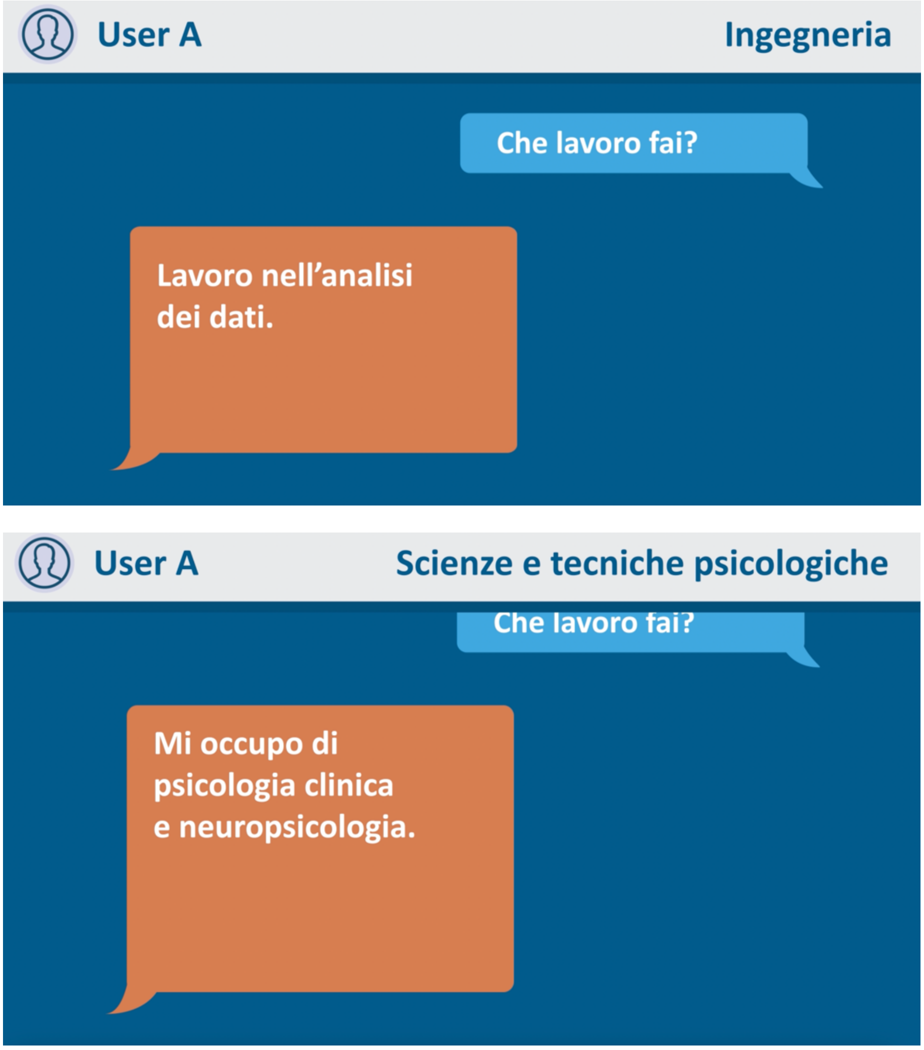


Figure 6.6 *Frame of the video (Control)*



Che lavoro fai?

Lavoro nell’analisi dei dati.

Che lavoro fai?

Mi occupo di psicologia clinica e neuropsicologia.

Instruments

Table 6.8 *Factor analysis of stereotypes' attribution*

PCAT item	Factor loading	
	1	2
STEM		
Factor 1: Attribution to gendered characteristics		
Women are generally better at humanities than men because they are naturally more sensitive	.09	**.81**
If there are more women in the humanities majors it is because they have less difficulty than men in studying these subjects	.05	**.74**
If there are more women in the humanities majors it is because they are more interested in these subjects than men	-.11	**.77**
Factor 2: Attribution to social pressure		
If there are more women in humanistic majors it is because men are hindered and discriminated against in this area	**.85**	-.07
Women are encouraged more than men to pursue studies in the humanities	**.71**	.06
Regardless of skills, women who work in science are disadvantaged compared to men just because they are women	**.76**	.00
Humanities		
Factor 1: Attribution to gendered characteristics		
Men are generally better than women in science subjects because they have a natural predisposition to logical and abstract reasoning	.03	**.45**
If there are more men in science majors it is because they have less difficulty than women in studying these subjects	-.00	**.43**
If there are more men in scientific majors it is because they are more interested in these subjects than women	-.02	**.35**
Factor 2: Attribution to social pressure		
If there are more men in scientific majors it is because women are hindered and discriminated against in this area	**.39**	-.08
Men are encouraged more than women to pursue scientific studies	**.38**	.05
Regardless of skills, men who work in the humanities are disadvantaged compared to women just because they are men	**.39**	.03

Note. N = 325.

The extraction method was principal axis factoring with an oblique (Oblimin with Kaiser Normalization) rotation. Factor loadings above .30 are in bold.

Complete results

Table 6.9 *Chi-square test*

Null hypothesis	Female		Male	
	Chi2	p-value	Chi2	p-value
Implicit gender-science stereotypes				
Treatment A = Control	2.02	.157	.08	.779
Treatment A = Treatment B	.04	.841	1.64	.204
Treatment B = Control	1.35	.246	2.45	.121
Humanities				
Explicit gender-science stereotypes				
Treatment A = Control	.47	.493	.42	.518
Treatment A = Treatment B	.40	.527	1.32	.251
Treatment B = Control	1.45	.229	.33	.564
Difference in stereotypes' attribution				
Treatment A = Control	.25	.621	.84	.359
Treatment A = Treatment B	.97	.324	2.66	.103
Treatment B = Control	.26	.614	.58	.447
STEM				
Explicit gender-science stereotypes				
Treatment A = Control	2.60	.107	.58	.446
Treatment A = Treatment B	.01	.904	.74	.388
Treatment B = Control	2.13	.144	.03	.867

Note. $N_{female} = 236$, $N_{male} = 87$, df = 1

Holm-adjusted p-values for multiple comparison

Table 6.10 *Direct effect of perceived gender imbalance (males)*

Predictor	Estimate	SE	t	p-value	95% CI	
					LB	UB
Implicit gender-science stereotypes						
Perceived gender imbalance^1 (hum)						
More women than men	.18	.24	.77	.441	-.29	.65
Almost all women	.18	.28	.62	.534	-.39	.74
Perceived gender imbalance^1 (STEM)						
Almost all men	-.52	.10	-5.29	.000	-.72	-.33
More men than women	.02	.11	.18	.857	-.20	.24
Explicit gender stereotypes (humanities)						
Perceived gender imbalance^1 (hum)						
More women than men	.97	.82	1.19	.235	-.63	2.58
Almost all women	1.66	1.00	1.66	.098	-.30	3.63
Explicit gender stereotypes (STEM)						
Perceived gender imbalance^1 (STEM)						
Almost all men	.28	.40	.70	.481	-.50	1.06
More men than women	.01	.46	.02	.985	-.89	.90

Note. b = coefficient; SE = Robust standard errors; CI = Confidence interval; LB = Lower bound; UB

[1] Reference category: “Equal number of women and men”

Table 6.11 *Indirect effects of perceived gender imbalance (females)*

Indirect effect	Estimate	SE	Percentile bootstrap 95% CI	
			LB	UB
Perceived imbalance (Humanities)				
Implicit gender stereotypes [Treat. A]	.03	.02	.002	.062
Implicit gender stereotypes [Treat. B]	-.00	.01	-.024	.015
Explicit gender stereotypes [Treat. A]	.02	.15	-.002	.056
Explicit gender stereotypes [Treat. B]	-.00	.01	-.024	.016
Perceived imbalance (STEM)				
Implicit gender stereotypes [Treat. A]	.00	.01	.004	.018
Implicit gender stereotypes [Treat. B]	-.00	.01	-.018	.003
Explicit gender stereotypes [Treat. A]	.01	.01	-.018	.039
Explicit gender stereotypes [Treat. B]	-.01	.01	-.046	.009

Note. SE = Bootstrap standard errors; CI = Confidence interval; LB = Lower bound; UB = Upper bound

Table 6.12 *Indirect effects of perceived gender imbalance (males)*

Indirect effect	Estimate	SE	Percentile bootstrap 95% CI	
			LB	UB
Perceived imbalance (Humanities)				
Implicit gender stereotypes [Treat. A]	.01	.02	-.016	.055
Implicit gender stereotypes [Treat. B]	-.01	.03	-.100	.025
Explicit gender stereotypes [Treat. A]	.01	.02	-.022	.070
Explicit gender stereotypes [Treat. B]	-.04	.04	-.123	.003
Perceived imbalance (STEM)				
Implicit gender stereotypes [Treat. A]	-.00	.02	-.029	.042
Implicit gender stereotypes [Treat. B]	.01	.02	-.040	.055
Explicit gender stereotypes [Treat. A]	.00	.02	-.036	.044
Explicit gender stereotypes [Treat. B]	-.05	.04	-.123	.003

Note. SE = Bootstrap standard errors; CI = Confidence interval; LB = Lower bound; UB = Upper bound