Acta Psychologica 194 (2019) 63–68

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

The role of implicit gender spatial stereotyping in mental rotation performance

Francesca Guizzo[a,*], Angelica Moè[b], Mara Cadinu[a], Chiara Bertolli[a]

[a] Department of Developmental and Socialization Psychology, University of Padova, Italy
[b] Department of General Psychology, University of Padova, Italy

ARTICLE INFO

Keywords:
Mental rotation
Gender stereotyping
Implicit measures
Spatial cognition

ABSTRACT

Men outscore women in mental rotation. Among the possible explanations for this result are gender stereotypes. Research has shown that instructions confirming or disconfirming the gender stereotype that men are more talented than women may affect performance in some spatial tasks, such as mental rotation, but research so far has shown inconsistent or null results. However, no research to date has assessed whether participants' implicit associations linking men to spatial abilities may modulate these effects. Thus, the goal of this study was to assess the moderating role of the implicit gender spatial stereotyping, that is the automatic associations between men vs. women and space, in male and female participants receiving either stereotypical (stating that men outscore women) or stereotype-nullifying (stating that there is no gender difference) explicit instructions. Results confirmed that men performed better than women in mental rotation, but also showed that in the stereotype-nullifying condition, the higher the automatic associations between space and men the lower men's performance. The discussion focuses on the importance of considering implicit gender spatial stereotyping as a factor that can modulate mental rotation performance.

1. Introduction

Mental rotation is the ability to mentally rotate 2D or 3D objects in space (Voyer, Voyer, & Bryden, 1995). It is very important for successful performance in some everyday tasks, such as orienting (Pazzaglia & Moè, 2013) or parking (Wolf et al., 2010), and for achievement in science, technology, engineering, and mathematics (STEM; Wai, Lubinski, & Benbow, 2009). Moreover, it is helpful in tasks that are not spatial in nature, such as understanding graphs, relationships, and metaphors (Frick, Möhring, & Newcombe, 2014). Importantly, one of the largest gender differences ever observed in cognitive abilities has been found for spatial cognition (Hyde, 2014), with men showing scores 1 standard deviation higher than women (Geiser, Lehman, & Eid, 2008). This gender gap can be explained by many factors ranging from biological (e.g., Grimshaw, Sitarenios, & Finegan, 1995) to experiential (e.g., Casey, 1996; Moè, 2016a), motivational or social factors (e.g., Dearing et al., 2012; Levine, Foley, Lourenco, Ehrlich, & Ratliff, 2016; Moè, 2016b). Among the social factors are gender stereotypes: men are believed to be more spatial talented than women (Halpern, Straight, & Stephenson, 2011) and this may affect performance (Campbell & Collaer, 2009).

Previous research has focused on explicit stereotypes, that is on conscious representations assessed through self-reports (e.g. Halpern et al., 2011; Moè, Meneghetti, & Cadinu, 2009), which have some limitations mainly due to social-desirability. For example, Halpern et al. found that individuals possess the stereotype that men outscore women in mental rotation tasks, even though they underestimate this gender difference compared to actual performance. One possible explanation for this finging is that the actual gender spatial stereotype is not very strong. However, another possibility is that this underestimation is due to social desirability concerns. To prevent social desirability concerns, the goal of this study is to capture for the first time spontaneous and automatic mental associations between spatial abilities and men, that is Implicit Gender Spatial Stereotyping (IGSS), and to assess their effects on performance in a Mental Rotation Test (MRT).

1.1. Spatial gender stereotypes and mental rotation performance

Consistent with the Stereotype Threat (ST) model proposed by Steele and collaborators (Steele & Aronson, 1995), activating a relevant negative stereotype about a minority group would decrease performance in tasks connected to the stereotype, an effect initially observed with African-Americans in verbal intelligence and female participants in math tasks (Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995).

* Corresponding author at: Dipartimento di Psicologia dello Sviluppo e della Socializzazione – DPSS, Via Venezia 8, 35131 Padova, Italy.
E-mail address: francesca.guizzo@unipd.it (F. Guizzo).

https://doi.org/10.1016/j.actpsy.2019.01.013
Received 18 September 2018; Received in revised form 20 December 2018; Accepted 29 January 2019

According to the ST model, the salience of the negative stereotype diverts threatened participants' attention from the task and increases their levels of anxiety, thus decreasing performance (in math see Cadinu, Maass, Rosabianca, & Kiesner, 2005; see also Schmader, Johns, & Forbes, 2008, for a review). However, ST manipulations do not always cause the expected performance decrement. Being reminded of belonging to a negatively stereotyped group may also cause stereotype reactance, that is increased effort and willingness to demonstrate that the stereotype is not true, thus favoring performance (Kray, Thompson, & Galinsky, 2001). This has also been demonstrated in mental rotation tests, for example when women are told that gender differences favoring males are not due to genetic factors (Moè, 2012). Moreover, research has shown that activating a positive stereotype about the ingroup may not favor performance due to the fear of not confirming the high expectations held by other people, causing a 'choking under pressure' effect, linked with anxiety and intrusive thoughts, which depletes working memory resources (in math see Beilock & Carr, 2005; with MRT see Moè, 2018).

In sum, both positive and negative stereotype instructions can either favor or harm performance. This depends on the way the stereotype-inducing message is appraised, either as a challenge or a threat (e.g., in math Shih, Ambady, Richeson, Fujita, & Gray, 2002). When appraised as a challenge an improvement in performance is expected, because achievement motivation increases as well as positive affect and success expectations (Walton & Cohen, 2003). On the contrary, when appraised as a threat, anxiety, intrusive thoughts, effort withdrawal, and a sense of uncertainty may negatively affect performance (in math: Beilock, Rydell, & McConnell, 2007; Cadinu et al., 2005; Schmader et al., 2008). Many psychological (e.g., in math Brown & Pinel, 2003) and contextual factors (for a meta-analysis in the math domain see Picho, Rodriquez, & Finnie, 2013; on mental rotation see Hirnstein, Andrews, & Hausmann, 2014; Moè, 2018) could affect the way the message is appraised. As a case in point, a challenge appraisal (i.e., reactance to a message suggesting worse performance) is more likely to occur when people are exposed to overt stereotypical messages than when the stereotype is activated in a subtle way (for a meta-analysis on race and intellectual performance as well as gender and mathematics, see Nguyen & Ryan, 2008).

The ST model has been strongly supported in the area of mathematics whereas results are more complex in the area of spatial abilities. A few studies (e.g., Heil, Jansen, Quaiser-Pohl, & Neuburger, 2012; McGlone & Aronson, 2006; Moè, 2009) have shown that women decrease MRT performance when the gender stereotype (i.e., men are better than women) is made salient. However, Doyle and Voyer (2016) have conducted a meta-analysis showing variable or null results based on ST manipulations regarding spatial tasks (as opposed to math). Therefore, the authors concluded that ST affects performance only in the case of female and mathematics. To overcome this state of impasse, it may be useful to look for individual differences that may moderate the unclear effects of stereotype threat manipulations. For example, Massa and colleagues (Massa, Mayer, & Bohon, 2005) found no direct effect of a stereotype threat manipulation on a spatial ability task, but found a moderation of gender role beliefs. Therefore, the assessment of individual differences may prove fruitful in the area of spatial abilities.

Implicit social cognition methodologies may contribute to assess the role of individual differences. Indeed, implicit measures such as the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) have been developed and successfully employed to measure a range of individual differences regarding attitudes and stereotypes (e.g., Gawronski & Bodenhausen, 2006). The success of implicit methods relies on the possibility to capture mostly automatic associations that prevent self-presentation biases, a problem that is often encountered when employing self-report explicit measures (Greenwald et al., 1998). Thus, we argue that such techniques may be helpful to assess the impact of gender stereotypes in the spatial domain. Indeed, the way in which an explicit stereotype provided with a specific manipulation is actually appraised may depend on previous implicit stereotypes held by participants (in math see Kiefer & Sekaquaptewa, 2007). Therefore, following the large use of implicit measures in the domain of science and math in the last two decades (e.g., Nosek, Banaji, & Greenwald, 2002; Nosek & Smyth, 2011), we propose to explore the role of such measures in the spatial domain as well.

1.2. Implicit gender spatial stereotyping

Implicit stereotypes are automatic associations between specific social categories (e.g., men) and given attributes (e.g. math) and are modeled on past experience (Greenwald & Banaji, 1995). During the last decades several studies have assessed stereotypes with implicit measures (Gawronski & Payne, 2010), among which the most frequently used is the Implicit Association Test (IAT; Greenwald et al., 1998), which is based on the assumption that the faster people react, the higher the activation of automatic associations. The IAT is a computer-based speeded categorization task based on the principle that it is easier to pair categories (e.g., men) with attributes (e.g., math) that are strongly associated in respondents' cognitive representation, as compared to categories with attributes that are less or not associated. The strength of automatic associations is measured by reaction times when categorizing together stimuli (by using the same response key) belonging to congruent pairings (e.g., men and math, in the presence of the contrasting pair women and verbal) as compared with the incongruent pairing (e.g., men and verbal vs. women and math).

The gender stereotype mostly studied with the IAT, even in very young children (Galdi, Cadinu, & Tomasetto, 2014), refers to abilities in the academic domain of mathematics and science as opposed to arts or verbal abilities: men and boys are associated with math more than women and girls. This implicit association has been found to moderate the effect of instructions on women math performance (Kiefer & Sekaquaptewa, 2007): the lower their implicit gender-math stereotyping the higher women's math performance in a reduced threat condition, i.e. when participants are told that the test is non-diagnostic of their math ability. Considering spatial abilities and mental rotation, to the best of our knowledge, no study till now has measured Implicit Gender *Spatial* Stereotyping (IGSS; i.e., the automatic associations between space and gender) and its moderating effects in the relation between explicit stereotype threat instructions and mental rotation performance.

1.3. Design and hypotheses

This study aims to measure for the first time the Implicit Gender Spatial Stereotyping (IGSS) and its moderating effects on mental rotation performance after participants receive one of two types of explicit instructions: one reminding the stereotype that men are more skilled than women in the spatial task at hand (i.e., *stereotypical* condition), the other stating there is no gender difference (i.e., *stereotype-nullifying* condition). It should be noticed that the latter condition enables researchers to control for the effects of personal beliefs by inducing a specific belief (i.e., no gender difference) instead of letting participants' personal beliefs affect performance in unpredictable or opposite ways, including for example the risk of inadvertently inducing stereotype threat, as suggested by Flore and Wicherts (2015). Importantly, the present design also includes a pre-manipulation MRT performance, which enables researchers to have a baseline to be compared with post-manipulation MRT performance.

Given the overall pattern of results in the spatial domain outlined by Doyle and Voyer (2016) we did not make explicit predictions on the effects of condition, but explored the possibility that the *stereotypical* message would generally be appraised as favorable by men and threatening by women, which may lead to increased and decreased performance respectively, whereas the *stereotype-nullifying* would generate the opposed pattern. Most importantly, we explored the possibility that the

IGSS would moderate the relationship between condition and MRT performance, leading to differential effects depending on the extent to which participants implicitly associate spatial abilities with men rather than women (IGSS). The direction of IGSS moderation could take several directions. To give an example, one possibility is that, in line with Kiefer and Sekaquaptewa (2007), for women the lower the IGSS (i.e. automatic association between men and spatial abilities) the higher their performance especially in the stereotype-nullifying condition, which is more congruent with their personal views. Another example is that for men the higher their IGSS the higher their performance in the stereotypical condition, which is more congruent with their personal views. Given the exploratory nature of the study we did not predict any specific direction of the moderation effects. Finally, in line with previous research (e.g., Voyer et al., 1995), we expect men to outscore women in the MRT.

2. Method

2.1. Participants

One hundred and forty-nine university students (91%) and workers (9%) participated. They were recruited either at University libraries, in quiet study rooms or among acquaintances unaware of the aims of the study by one of five experimenters (3 females, 2 males). Six participants (4 in the stereotypical condition and 2 in the nullifying condition) failed to understand the instructions (see Manipulation check below) and were excluded. The final sample included 143 participants (72 men, 71 women), $M_{age} = 22.34$, $SD_{age} = 2.49$. Please notice that the minimum sample size is 130, determined based on a-priori power analysis requiring an intermediate effect size ($f^2 = 0.15$) and $1 - \beta = 0.80$ (with $\alpha = 0.05$).

2.2. Measures

2.2.1. Mental rotation

Participants performed the 20-item MRT developed by Vandenberg and Kuse (1978). Each item consists of a target figure (three-dimensional representation of an object) and four stimuli figures, two representing a rotated version of the target object, and two distractors. Following previous research (Moè & Pazzaglia, 2006), 10 items were administered before and 10 after the experimental manipulation. Participants had to select the two correct rotated versions of the target object within 4 min. One point was assigned when both correct rotations were identified. The maximum score was 10 for each half of the two parts of the MRT.

2.2.2. Implicit gender spatial stereotyping (IGSS)

To assess IGSS ($\alpha = 0.72$) as the strength of automatic cognitive associations between spatial abilities and men vs. women, we employed the Single Attribute Implicit Association Test procedure (SA-IAT, Penke, Eichstaedt, & Asendorpf, 2006). Compared to the classic IAT, the SA-IAT procedure requires comparing a single attribute category (e.g., "Space") to two target categories (e.g., "Women" and "Men"). The Single Attribute IAT is particularly suitable for this study because there is no plausible counter-category for "Space" (see Bluemke & Friese, 2008, for other advantages of using a Single Target/Attribute Task).

In a pre-test, twenty-one university students (11 men) rated the 'spatial nature' of 25 words (e.g., square, line, map). For the category "Space" five words (orientation, rotation, coordinates, compass, map) were found to significantly *represent the concept of spatial ability* (i.e., scoring significantly higher than the midpoint 4 on a scale ranging from 1 = not at all to 7 = very much).

As target stimuli, we used five words representing the target "Women" (mother, female, she, wife, daughter) and five words representing the target "Men" (father, male, he, husband, son; see Cadinu & Galdi, 2012; Nosek et al., 2002). The task was completed on a laptop using Inquisit 4 (Millisecond ltd.). In the first training block participants categorized 20 randomized target words as quickly as possible pressing the *D* key when the word referred to "Women", or K for "Men". Two critical combined blocks composed of 40 randomized trials followed. In the first combined block, participants pressed K to categorize both "Men" and "Space" stimuli, whereas *D* was used to categorize "Women" stimuli. The reverse combination was presented in the second and last combined block. Stimuli words appeared one by one at the center of the screen, and a red cross followed incorrect responses. To avoid response bias in the two combined blocks, the stimuli were presented respectively with a ratio of 20:10:10 and 10:10:20, so that in both combined blocks the number of required right-hand and left-hand responses was equal. The order of presentation of the two combined blocks was counterbalanced across participants.

2.2.3. Manipulation check

After the IGSS, participants were asked to answer to the following written question 'In the instruction you were told that a) men score higher than women, b) women score higher than men, c) there is no gender difference'. Those participants assigned to the stereotypical condition who did not cross answer a) and those of the nullifying condition who did not cross option c) were excluded from the analyses (see Participants) because they did not properly understand the instructions.

2.2.4. Explicit gender spatial stereotype (EGSS)

To measure EGSS, at the end of the experiment participants were asked to express their personal belief on the gender spatial stereotype. Specifically, they were asked to think of the test they had just performed and to choose one of the following statements that better corresponds to their belief: *In your opinion, in this kind of task 1 = women score higher than men; 2 = men score higher than women; 3 = men and women score equally*.

2.3. Procedure

After signing informed consent, participants were presented with three MRT practice items. After the "start" signal, participants had to complete the first 10 items of MRT in 4 min. After the "stop" signal, they read a brief description of the aim of MRT. Participants then randomly received either stereotypical or [stereotype-nullifying] information: Research conducted on spatial abilities using this test has shown that men perform definitively better than women [there is no gender difference in performance in this test]. In other words, men obtain higher scores than women [women obtain scores as high as men] in this test. Instructions enclosed a graph showing higher scores for men than women in the stereotypical condition and similar scores in the stereotype-nullifying condition. Then, participants performed the post-manipulation MRT (10 items), performed the IGSS, completed the manipulation check and responded to the EGSS. Other measures were collected. At the end, participants were debriefed and thanked for their participation.

3. Results

3.1. Preliminary analyses

Following the improved *D*-algorithm procedure (Greenwald, Nosek, & Banaji, 2003) we calculated participants' IGSS scores so that positive higher values indicate stronger association between *men* and *spatial abilities* whereas negative lower values indicate stronger association between *women* and *spatial abilities*. The average IGSS score was close to zero ($M = 0.02$, $SD = 0.41$), but importantly the IGSS score distribution ranged from -1.07 to 1.50, thus demonstrating a good span of this individual difference. No effects of Gender ($M_{males} = 0.01$, $SD_{males} = 0.38$; $M_{females} = 0.03$, $SD_{females} = 0.44$; $F = 0.02$, $p = .88$), Condition or Gender × Condition interaction were found ($Fs < 0.2.31$, $ps > .13$).

Regarding the EGSS, 72% of the participants believed that there is no gender difference, 22% that men score higher than women, and 6% that women score higher than men. Logistic analysis showed no effects involving Gender or Condition (ORs < 2.70, $p > .08$). Notice that the EGSS response "women score higher" was excluded from analyses because it was chosen by only 9 participants. Moreover, the IGSS and EGSS scores did not correlate either with each other ($r = 0.01$, $p = .94$) or with MRT performance ($rs < 0.13$, $ps > .14$).

A 2 (Condition: stereotypical vs. stereotype-nullifying) × 2 (Gender) × 2 (Time: pre vs. post manipulation) repeated measures ANOVA on MRT scores, with the last variable within-person, showed a main effect of Gender, $F(1,139) = 11.43$, $p = .001$, Cohen's $d = 0.59$, with men obtaining higher scores ($M = 5.54$, $SD = 2.23$) than women ($M = 4.24$, $SD = 2.35$). A significant interaction also emerged between Time and Gender, $F(1,139) = 4.91$, $p = .03$, Cohen's $d = 0.13$. Simple effect analysis revealed an effect of Time for men $F(1,139) = 5.22$, $p = .02$, but not for women ($p = .40$): men performed better post-manipulation ($M = 5.81$, $SD = 2.41$) than pre-manipulation ($M = 5.28$, $SD = 2.50$), whereas no difference emerged for women ($Mpre = 4.34$, $SD = 2.54$; $Mpost = 4.14$, $SD = 2.53$).

3.2. IGSS moderation analysis

Since preliminary analyses showed that IGSS and EGSS were affected neither by type of Condition nor by the interaction Condition × Gender they were included as moderators in the analysis. A multiple regression tested whether IGSS moderates the effects of Condition and Gender on post-manipulation MRT scores by also including EGSS as a potential moderator. In the first step, Condition (stereotypical = 0, stereotype-nullifying = 1), Gender (women = 0, men = 1), IGSS score (continuous, centered) and EGSS (men equal to women = 0, men score higher than women = 1) were entered as predictors. In the second and third step, we entered the two-way and the three-way interactions among IGSS, Condition, and Gender as well as among EGSS, Condition, and Gender always including pre-manipulation MRT as a covariate. As shown in Table 1, the two-way interaction Gender × IGSS ($t = -3.37$, $p = .001$) and the three-way interaction Condition × Gender × IGSS ($t = 2.64$, $p = .009$) were significant. Importantly, including the three-way interaction led to a significant increase in the amount of explained variance $\Delta R^2 = 0.03$, $p = .03$, $R^2 = 0.58$, $F(12,125) = 12.93$, $p < .001$. On the contrary, EGSS did not lead to any significant effect. This demonstrates that IGSS, as opposed to EGSS, moderates the relation between Condition, Gender and post-manipulation MRT scores controlling for baseline performance. As shown in Fig. 1, for men receiving stereotype-nullifying information, the higher their level of IGSS the lower their performance in the MRT. In other words, the stronger men's automatic associations between spatial abilities and men, the lower their MRT performance when inconsistent information was delivered (stereotype-nullifying condition). Contrary to men, female MRT performance did not vary as a function of Condition and IGSS. Please notice that results were the same when EGSS was not included as a moderator in the model. Also, statistical results were unchanged excluding pre-manipulation MRT as a covariate except for the amount of variance explained.

Table 1
Post-manipulation MRT scores (controlling for pre-manipulation MRT scores) predicted by Condition (0 = stereotype-nullifying, 1 = stereotypical), Gender (0 = women, 1 = men), IGSS (continuous, centered) and EGSS (0 = men equal to women, 1 = men better than women).

	b	$SE\ b$	R^2	ΔR^2	ΔF (df)
Step 1 (simple predictors)			0.51		25.26***(5120)
Intercept	4.11***	0.29			
Pre-MRT	0.63***	0.07			
Condition	0.29	0.33			
Gender	1.19***	0.33			
IGSS	−0.64	0.39			
EGSS	0.47	0.40			
Step 2 (two-way interactions)			0.55	0.04	2.05 (5115)
Intercept	4.37***	0.34			
Pre-MRT	0.64***	0.07			
Condition	0.03	0.49			
Gender	0.67	0.47			
IGSS	−0.53	0.64			
EGSS	−0.88	0.79			
Condition × Gender	0.35	0.66			
Condition × IGSS	0.88	0.76			
Gender × IGSS	−1.60*	0.77			
Condition × EGSS	1.06	0.86			
Gender × EGSS	1.29	0.79			
Step 3 (three-way interactions)			0.58	0.03	3.48* (2113)
Intercept	4.33***	0.34			
Pre-MRT	0.63***	0.07			
Condition	0.06	0.51			
Gender	0.53	0.48			
IGSS	0.38	0.72			
EGSS	−0.62	0.94			
Condition × Gender	0.40	0.71			
Condition × IGSS	−0.78	0.98			
Gender × IGSS	−3.82**	1.13			
Condition × EGSS	0.88	1.15			
Gender × EGSS	1.12	1.40			
Condition × Gender × IGSS	4.01**	1.52			
Condition × Gender × EGSS	0.13	1.68			

* $p < .05$.
** $p < .01$.
*** $p < .001$.

4. Discussion

In the present study we assessed gender spatial stereotypes for the first time with an implicit measure (the IGSS), which has the advantage of preventing self-presentation biases, a problem that is often encountered when employing only self-report explicit measures (Greenwald et al., 1998). Results from the present study showed that men outscored women in MRT, confirming a large amount of previous research (e.g., Voyer et al., 1995) and, most importantly, that IGSS modulated men's performance as opposed to the explicit measure, which did not affect results. More specifically, when men received stereotype-nullifying instructions the higher their implicit gender spatial stereotyping the lower their MRT performance. One possible interpretation of this result is that men performance depends on the *incongruence* between their own automatic associations between gender and space and the explicit instructions provided by the experimenter. In other words, the higher the stereotype *incongruence* between the explicit instructions and men's implicit gender spatial stereotyping the lower their mental rotation performance.

This result confirms the novel role of implicit gender spatial stereotyping as opposed to explicit stereotyping. No effect of condition on performance was found without taking into account the IGSS, a result in line with Doyle and Voyer's (2016) meta-analysis showing null results on the role of ST in the spatial domain. Therefore, our findings capturing automatic gender spatial stereotypes suggest that the assessment of individual differences, especially through automatic association measures (Gawronski & Bodenhausen, 2006), is a promising research direction in this area. Nevertheless, given the exploratory nature of this study, future research should replicate the present findings.

In contrast with male results, no moderation effects were found for women. Although the reason for this gender difference is not clear, in future studies it may be useful to add a counter-stereotypical condition stating that women score higher than men (see Wraga, Duncan, Jacobs, Helt, & Church, 2006 for an example). This counter-stereotypical message may be appraised more favorably by women than the present no-difference condition, and improve their performance especially in

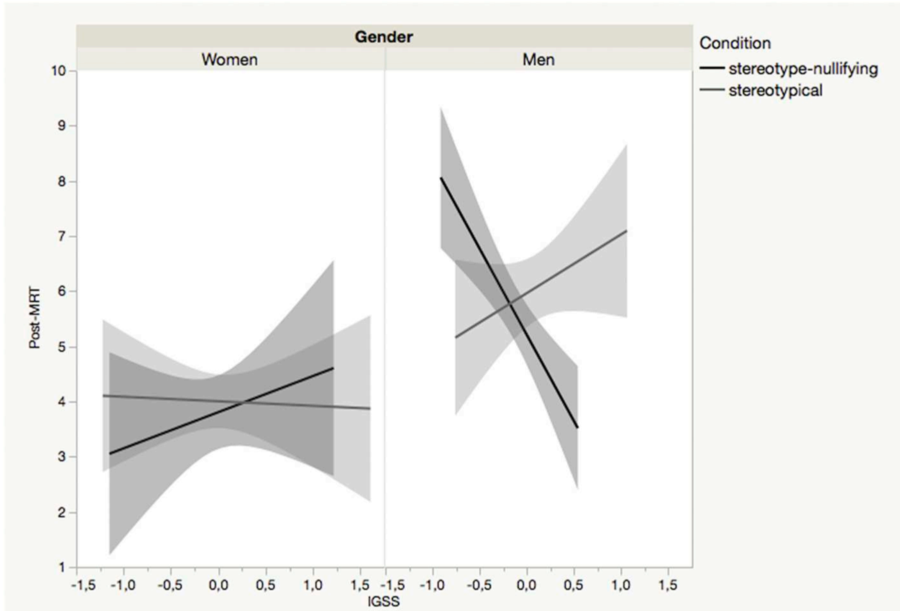

Fig. 1. Post-manipulation MRT scores as a function of Condition (stereotype-nullifying vs. stereotypical), Gender and Implicit Gender Spatial Stereotyping (IGSS) scores.

combination with lower implicit gender spatial stereotyping.

One may argue that a possible limitation of the present study is that it did not include a control condition in which participants are not provided with any information. However, such condition was purposely not included because participants already have their beliefs, which could differ within the same condition (some participants might believe that men are more skilled, others that there is no gender difference or even that women are better); this lack of information would cause a spurious condition, in which participants may put in place unpredictable and unknown beliefs that could affect performance even in opposite ways. Moreover, this condition could inadvertently cause stereotype threat, as suggested by Flore and Wicherts (2015). Therefore, the present study includes a more effective control condition, i.e., a stereotype-nullifying condition in which participants are told that there is no gender difference in mental rotation performance. In addition, it should be noticed that the present results were obtained controlling for the pre-manipulation MRT baseline performance, which represents an additional control for the post-manipulation MRT.

A limitation of the present study is that explicit stereotypes were measured with only one categorical item, which usually presents lower correlation with implicit measures such as IATs compared to continuous measures (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005). In fact, in the present study implicit and explicit gender stereotypes did not correlate with each other despite current meta-analysis research showing low but significant positive correlations (Hofmann et al., 2005). Moreover, the lack of correlation between the implicit and the explicit measures of gender spatial stereotypes may be due the fact that the IGSS words (i.e., orientation, rotation, coordinates, compass, map) mainly refer to navigation whereas the EGSS specifically refers to the mental rotation test that participants performed. Therefore, future studies should measure the explicit endorsement of the spatial stereotype with a continuous multi-item scale to obtain a measure more comparable to the implicit stereotype task, which is a continuous measure. At the same time, the IAT could also include words more relevant to mental rotation (e.g., axis, turn).

Another limitation of the present study is that the implicit and explicit moderators were measured after the manipulation. Although they were not affected by the manipulation, it would be useful for future research to measure them, for example, a week before the experimental procedure.

Future studies should also address potential mechanisms that may explain the present moderation results. For example, one possibility is that men under stereotype incongruence did not react to disconfirm the absence of gender differences, so they acted through effort withdrawal because the experimenter expected similar performance between genders. Another possibility is that men increased their level of anxiety and/or task intrusive thoughts when they were presented with information incongruent with their stereotypic automatic associations, thus affecting their performance, as found in previous research in the math domain (e.g., Cadinu et al., 2005; Osborne, 2007).

5. Conclusions

In the present study, we measured implicit gender spatial stereotyping for the first-time and found support for its role as a moderator: the stronger the implicit stereotype associating men to spatial abilities the lower men's mental rotation performance when stereotype-nullifying information was delivered. According to our interpretation, this lower performance depends on the incongruence between men's chronic implicit gender spatial stereotyping and the explicit instructions. Therefore, we propose that implicit gender spatial stereotypes are measured both in future research and when planning training and teaching interventions to improve mental rotation performance.

References

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science, 16*, 101–105. https://doi.org/10.1111/j.0956-7976.2005.00789.x.

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

Bluemke, M., & Friese, M. (2008). Reliability and validity of the Single-Target IAT (ST-IAT): Assessing automatic affect towards multiple attitude objects. *European Journal of Social Psychology, 38*, 977–997. https://doi.org/10.1002/ejsp.487.

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology, 39*, 626–633. https://doi.org/10.1016/S0022-1031(03)00039-8.

Cadinu, M., & Galdi, S. (2012). Gender differences in implicit gender self-categorization lead to stronger gender self-stereotyping by women than by men. *European Journal of Social Psychology, 42*, 546–551. https://doi.org/10.1002/ejsp.1881.

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? *Psychological Science, 16*, 572–578. https://doi.org/10.1111/j.0956-7976.2005.01577.x.

Campbell, S. M., & Collaer, M. L. (2009). Stereotype threat and gender differences in performance on a novel visuospatial task. *Psychology of Women Quarterly, 33*, 437–444. https://doi.org/10.1111/j.1471-6402.2009.01521.x.

Casey, M. B. (1996). Understanding individual differences in spatial ability within females: A nature/nurture interactionist framework. *Developmental Review, 16*, 241–260. https://doi.org/10.1006/drev.1996.0009.

Dearing, E., Casey, B. M., Ganley, C. M., Tillinger, M., Laski, E., & Montecillo, C. (2012). Young girls' arithmetic and spatial skills: The distal and proximal roles of family socioeconomics and home learning experiences. *Early Childhood Research Quarterly, 27*, 458–470. https://doi.org/10.1016/j.ecresq.2012.01.002.

Doyle, R. A., & Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences, 47*, 103–116. https://doi.org/10.1016/j.lindif.2015.12.018.

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology, 53*, 25–44. https://doi.org/10.1016/j.jsp.2014.10.002.

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Development of mental transformation abilities. *Trends in Cognitive Sciences, 18*, 536–542. https://doi.org/10.1016/j.tics.2014.05.011.

Galdi, S., Cadinu, M., & Tomasetto, C. (2014). The roots of stereotype threat: When automatic associations disrupt girls' math performance. *Child Development, 85*, 250–263. https://doi.org/10.1111/cdev.12128.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692–731. https://doi.org/10.1037/0033-2909.132.5.692.

Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. New York: Guilford Press.

Geiser, C., Lehman, W., & Eid, M. (2008). A note on sex differences in mental rotation in different age groups. *Intelligence, 36*, 556–563. https://doi.org/10.1016/j.intell.2007.12.003.

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

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

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

Grimshaw, G. M., Sitarenios, G., & Finegan, J. A. K. (1995). Mental rotation at 7 years: Relations with prenatal testosterone levels and spatial play experiences. *Brain and Cognition, 29*, 85–100. https://doi.org/10.1006/brcg.1995.1269.

Halpern, D. F., Straight, C. A., & Stephenson, C. L. (2011). Beliefs about cognitive gender differences: Accurate for direction, underestimated for size. *Sex Roles, 64*, 336–347. https://doi.org/10.1007/s11199-010-9891-2.

Heil, M., Jansen, P., Quaiser-Pohl, C., & Neuburger, S. (2012). Gender-specific effects of artificially induced gender beliefs in mental rotation. *Learning and Individual Differences, 22*, 350–353. https://doi.org/10.1016/j.lindif.2012.01.004.

Hirnstein, M., Andrews, L. C., & Hausmann, M. (2014). Gender-stereotyping and cognitive sex differences in mixed and same-sex groups. *Archives of Sexual Behavior, 43*, 1663–1673. https://doi.org/10.1007/s10508-014-0311-5.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*, 1369–1385. https://doi.org/10.1177/0146167205275613.

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology, 65*, 373–398. https://doi.org/10.1146/annurev-psych-010213-115057.

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*, 825–832. https://doi.org/10.1016/j.jesp.2006.08.004.

Kray, L. J., Thompson, L., & Galinsky, A. D. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology, 80*, 942–958.

Levine, S. C., Foley, A., Lourenco, S., Ehrlich, S., & Ratliff, K. (2016). Sex differences in spatial cognition: Advancing the conversation. *Wiley Interdisciplinary Reviews: Cognitive Science, 7*, 127–155. https://doi.org/10.1002/wcs.1380.

Massa, L. J., Mayer, R. E., & Bohon, L. M. (2005). Individual differences in gender role beliefs influence spatial ability test performance. *Learning and Individual Differences, 15*, 99–111. https://doi.org/10.1016/j.lindif.2004.11.002.

McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology, 27*, 486–493. https://doi.org/10.1016/j.appdev.2006.06.003.

Moè, A. (2009). Are males always better than females in mental rotation? Exploring a gender belief explanation. *Learning and Individual Differences, 19*, 21–27. https://doi.org/10.1016/j.lindif.2009.08.002.

Moè, A. (2012). Gender difference does not mean genetic difference: Externalizing improves performance in mental rotation. *Learning and Individual Differences, 22*, 20–24. https://doi.org/10.1016/j.lindif.2011.11.001.

Moè, A. (2016a). Does experience with spatial school subjects favour girls' mental rotation performance? *Learning and Individual Differences, 47*, 11–16. https://doi.org/10.1016/j.lindif.2015.12.007.

Moè, A. (2016b). Teaching motivation and strategies to improve mental rotation abilities. *Intelligence, 59*, 16–23. https://doi.org/10.1016/j.intell.2016.10.004.

Moè, A. (2018). Effects of group gender composition on Mental Rotation Test performance in women. *Archives of Sexual Behavior, 47*, 2299–2305. https://doi.org/10.1007/s10508-018-1245-0.

Moè, A., Meneghetti, C., & Cadinu, M. (2009). Women and mental rotation: Incremental theory and spatial strategy use enhance performance. *Personality and Individual Differences, 46*, 187–191. https://doi.org/10.1016/j.paid.2008.09.030.

Moè, A., & Pazzaglia, F. (2006). Following the instructions! Effects of gender beliefs in mental rotation. *Learning and Individual Differences, 16*, 369–377. https://doi.org/10.1016/j.lindif.2007.01.002.

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

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology, 83*, 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*, 1125–1156. https://doi.org/10.3102/0002831211410683.

Osborne, J. W. (2007). Linking stereotype threat and anxiety. *Educational Psychology, 27*, 135–154. https://doi.org/10.1080/01443410601069929.

Pazzaglia, F., & Moè, A. (2013). Cognitive styles and mental rotation ability in map learning. *Cognitive Processing, 14*, 391–399. https://doi.org/10.1007/s10339-013-0572-2.

Penke, L., Eichstaedt, J., & Asendorpf, J. B. (2006). Single-Attribute Implicit Association Tests (SA-IAT) for the assessment of unipolar constructs. *Experimental Psychology, 53*, 283–291. https://doi.org/10.1027/1618-3169.53.4.283.

Picho, K., Rodriquez, A., & Finnie, L. (2013). Exploring the moderating role of context on the mathematics performance of females under stereotype threat: A meta-analysis. *Journal of Social Psychology, 153*, 299–333. https://doi.org/10.1080/00224545.2012.737380.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*, 336–356. https://doi.org/10.1037/0033-295X.115.2.336.

Shih, M., Ambady, N., Richeson, J. A., Fujita, K., & Gray, H. M. (2002). Stereotype performance boosts: The impact of self-relevance and the manner of stereotype activation. *Journal of Personality and Social Psychology, 38*, 638–647. https://doi.org/10.1037/0022-3514.83.3.638.

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

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797–811.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of 3-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599–604.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250–270.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*, 817–835. https://doi.org/10.1037/a0016127.

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology, 39*, 456–467. https://doi.org/10.1016/S0022-1031(03)00019-2.

Wolf, C. C., Ocklenburg, S., Ören, B., Becker, C., Hofstätter, A., Bös, C., ... Güntürkün, O. (2010). Sex differences in parking are affected by biological and social factors. *Psychological Research, 74*, 429–435. https://doi.org/10.1016/j.lindif.2008.02.002.

Wraga, M., Duncan, L., Jacobs, E. H., Helt, M., & Church, J. (2006). Stereotype susceptibility narrows the gender gap in imagined self-rotation performance. *Psychonomic Bulletin & Review, 13*, 813–819. https://doi.org/10.3758/BF03194002.