Learning and Individual Differences 47 (2016) 103–116

Contents lists available at ScienceDirect

Learning and Individual Differences

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

Stereotype manipulation effects on math and spatial test performance: A meta-analysis

Randi A. Doyle *, Daniel Voyer

Psychology Department, University of New Brunswick, Canada

ARTICLE INFO

Article history:
Received 25 April 2015
Received in revised form 31 October 2015
Accepted 22 December 2015

Keywords:
Mathematical cognition
Spatial cognition
Stereotype threat
Gender differences

ABSTRACT

A meta-analysis of 224 effect sizes (*d*) drawn from 86 studies examined the relationship between gender, stereotype manipulations, and math and spatial performance. Stereotype manipulations were analyzed separately as a function of gender (threat to males, threat to females, lift for males, lift for females). Only the threat to females grouping ($d = 0.29$) showed a mean effect size that was significantly different from zero, indicating significant deleterious effects of stereotype threat instructions. Analyses for the threat to females and lift for females categories in an attempt to account for significant variability in these groupings showed that task, sex of experimenter, and control group type accounted for significant variance in effect sizes. Essentially, the effects of stereotype threat on women can be interpreted as relatively small but significant in math performance, but non-significant in spatial performance. Implications for interpretations of gender differences in math and spatial performance are discussed.

1. Introduction

Throughout the history of psychology, researchers have been interested in examining and understanding cognitive gender differences (Halpern, 2012). The extensive literature review conducted by Maccoby and Jacklin (1974) was a landmark in this pursuit as it identified three specific areas where gender differences in cognition exist: math and spatial abilities in favor of males, and verbal abilities in favor of females.

Since Maccoby and Jacklin's (1974) seminal work, much of the research has moved away from the study of gender differences in verbal abilities as the meta-analysis by Hyde and Linn (1988) showed that gender differences in verbal ability were variable both in magnitude and direction of advantage. In contrast, the male advantage in spatial abilities has been confirmed in at least two meta-analyses (Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995). Various explanations have been proposed for gender differences in spatial abilities. They span biological factors such as prenatal sex hormones (Berenbaum, Korman, & Leveroni, 1995), circulating sex hormones (Hausmann, Slabbekoorn, Van Goozen, Cohen-Kettenis, & Güntürkün, 2000), and gender differences in brain activation reflecting differential strategy choices (Jordan, Wüstenberg, Heinze, Peters, & Jäncke, 2002). Various environmental and procedural factors have also been proposed as contributors to gender differences in spatial ability, including the role of childhood spatial activities (Doyle, Voyer, & Cherney, 2012; Voyer, Nolan, & Voyer, 2000), gender-role identity (Signorella & Jamison, 1978; McGlone & Aronson, 2006), scoring procedures (Goldstein, Haldane, & Mitchell, 1990), and timing procedures (Voyer, 2011; Voyer, Rodgers, & McCormick, 2004).

Gender differences in math abilities provide a puzzling situation. On the one hand, the most recent research findings provide little support for the view that men are better at math than are women as the magnitude of gender differences in math performance reflect such small effects that they have been interpreted as non-existent (Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Lindberg, Hyde, Petersen, & Linn, 2010), or as even favoring females when considering school grades (Voyer & Voyer, 2014). On the other hand, there seems to be a commonly held belief in Western culture that men outperform women in mathematics (Matlin, 2011; Nosek, Banaji, & Greenwald, 2002; Tartre & Fennema, 1995). Therefore, although research suggests that gender differences in math performance are negligible, the belief that males are better at math than are females persists (Else-Quest, Hyde, & Linn, 2010; Lindberg et al., 2010; Nosek et al., 2002). This might explain why there has been a great deal of research on the concept of stereotype threat in the context of gender differences in math.

Stereotype threat (Steele & Aronson, 1995; Steele, 1997) refers to stereotypes about societal groups that can influence the intellectual functioning of group members. Steele and Aronson (1995) suggested that individuals experience stereotype threat when they are at risk of confirming a negative stereotype about their group. In a typical study illustrating the influence of stereotype threat on women's math performance, researchers instruct participants that a difficult math test yielded gender differences in the past, without necessarily specifying the direction of the effect. Under these instructions, females typically perform worse on a math test than a control group that did not receive

* Corresponding author at: Department of Psychology, University of New Brunswick, Fredericton, NB E3B 5A3, Canada.
E-mail address: randi.doyle@unb.ca (R.A. Doyle).

http://dx.doi.org/10.1016/j.lindif.2015.12.018
1041-6080/

such instructions (e.g., Ambady, Shih, Kim, & Pittinsky, 2001; Keller, 2002; Rydell, McConnell, & Beilock, 2009). Considering that such findings have been replicated numerous times, it is not surprising that a few meta-analyses have been conducted to examine stereotype threat as a factor relevant to gender differences in math performance. We will now examine more closely the relevant meta-analyses that have been published to date.

1.1. Previous meta-analyses

The first meta-analysis of relevance that we could retrieve, conducted by Walton and Cohen (2003), examined the role of stereotype lift on intellectual performance, an effect closely related to stereotype threat. Stereotype lift occurs when non-stereotyped group members actually experience a boost in performance in testing situations developed to produce a stereotype threat for a stereotyped group (Marx & Stapel, 2006; Shih, Pittinsky, & Ambady, 1999).

Walton and Cohen's (2003) meta-analysis examined the effect of both stereotype threat and stereotype lift on participants' intellectual test performance when they were *not* the target of a negative stereotype (e.g., men, Caucasians). Their meta-analytic findings revealed a relatively small but significant overall effect of stereotype lift among non-stereotyped participants ($d = 0.24$). In particular, these participants performed better in stereotype lift conditions (i.e., men who were told that men are better at math than women before a math test) than in stereotype irrelevant conditions. Walton and Cohen's meta-analysis also showed that the effect of stereotype lift on non-stereotyped participants' performance was similar across studies that implemented an implied stereotype threat (i.e., stated that the test was a math test, but did not mention gender differences) and studies that implemented an explicit stereotype threat (i.e., explicitly described gender differences in favor of the non-stereotyped group). Aside from these finding, Walton and Cohen did not examine at all the effect of stereotype manipulation on participants experiencing the threat (the experimental group or minority group).

Extending the work of Walton and Cohen (2003), Nguyen and Ryan's (2008) meta-analysis focused on the effect of stereotype threat on minorities' and women's cognitive test performance, including quantitative, verbal, analytic, and nonverbal intelligence tests. Nguyen and Ryan found that women who identify little with math (i.e., felt that math held little importance in their life) showed only a small effect of stereotype threat ($d = 0.11$) compared to women who report moderate ($d = 0.52$) or high ($d = 0.29$) levels of math identification. Threat salience also moderated the effect of stereotype threat on women's math performance, as subtle threat-activating cues (e.g., stressing the evaluative nature of the test) seemed to be more detrimental to women's math performance than moderately explicit (e.g., informing participants that a test has revealed gender differences in performance) and blatant threat-activating cues (e.g., stating that men tend to score higher on the test than women). Women also appeared to benefit more from explicit threat removal strategies, such as an explicit statement that the test is free of gender biases, than subtle threat removal strategies, such as stating that test performance will not be assessed. Finally, test difficulty moderated the effect of stereotype threat on women's math performance, with more difficult tests producing larger effect sizes. Although Nguyen and Ryan's findings are informative in regards to the effect of stereotype threat on women's math performance, their meta-analysis did not include males, therefore examining gender differences and the effect of stereotype threat on men's performance was impossible.

In another meta-analysis, Stoet and Geary (2012) focused only on research replicating the Spencer, Steele, and Quinn's (1999) study on the effect of stereotype threat on women's math performance. Stoet and Geary reported that the evidence for stereotype threat effects on women's math performance is generally weak, and that these effects appear to vary in magnitude when the use of adjusted scores is considered. According to Stoet and Geary, adjusted performance scores (*i.e., adjusted for a previous mathematics test score by means of analysis of covariance*) produce significant stereotype threat effects in women whereas unadjusted performance scores produce non-significant effects. Based on their findings, Stoet and Geary concluded that the effect of stereotype threat on women's math performance is not as robust as many assume. However, it is important to note that Stoet and Geary did not examine gender differences in the effect of math performance and they focused only on the small number of studies that used the exact same paradigm as tested by Spencer et al. (1999). Therefore, the results from Stoet and Geary's meta-analysis should be interpreted with caution.

The most recent meta-analysis of stereotype threat and mathematics performance examined the moderating role of context on females' performance (Picho, Rodriguez, & Finnie, 2013). With a sample of 103 effect sizes, Picho et al. found a small deleterious effect of stereotype threat on females' math performance ($d = 0.24$). Stereotype threat priming (explicit, implicit) and testing group composition (single sex, mixed sex) did not significantly moderate the effect of stereotype threat on females' math performance; Picho et al. suggested that this may be due to their small sample size. Important research questions remain to be answered as, like Nguyen and Ryan's (2008) and Stoet and Geary (2012), Picho et al.'s meta-analysis only included female samples.

1.2. Stereotype threat and spatial abilities

Up to this point, our focus has been on math performance and how it might be affected by stereotype threat, especially among females. Considering that, contrary to math, spatial abilities still produce large gender differences (favoring males), it is surprising that it has yet to receive much attention in the context of a meta-analysis examining stereotype threat. In reality, past research has suggested that stereotype threat might be plausible as an explanation for gender differences in spatial abilities. For example, in what appears to be the first work investigating a manipulation akin to stereotype threat on spatial abilities, Sharps, Welton, and Price (1993) investigated the effect of various instruction types on performance. In this context, Sharps et al. had participants complete the Vandenberg and Kuse (1978) mental rotations test in one of two instructional conditions: one diminishing the spatial aspects of the task, and one emphasizing the use of spatial ability in a stereotypically masculine visuospatial task. Although Sharps et al. did not describe their paradigm as manipulating "stereotype threat", the design of their experiment was essentially identical to those that test the effect of implicit stereotype threat on performance (e.g., Davies, Spencer, Quinn, & Gerhardstein, 2002). As expected, in the context of stereotype threat, Sharps et al. found significant gender differences in mental rotation performance in the spatial instruction condition, but no gender difference in performance in the non-spatial instruction condition. Since this first study by Sharps et al., these results have been replicated with other spatial tasks (e.g. Delgado & Prieto, 2008; Massa, Mayer, & Bohon, 2005; McGlone & Aronson, 2006; Wraga, Helt, Jacobs, & Sullivan, 2007). Still, from the literature considered so far, previous meta-analyses investigating the effects of stereotype threat and stereotype lift manipulations have not considered performance on spatial tasks. Accordingly, the present meta-analysis reflects a novel contribution to the literature as it examined the effect of stereotype manipulations on both math and spatial performance in the same analysis. In addition, previous meta-analyses did not examine the differential effects of stereotype manipulations on males and females performance. Therefore, the present analysis also considered gender of the participants as a function of the type of manipulation (threat, lift), thereby contributing further to the existing literature.

1.3. Research questions

The current meta-analysis aimed to shed light on the effects of stereotype manipulations on math and spatial performance, and potential