Acta Psychologica 188 (2018) 65–73

Contents lists available at ScienceDirect

Acta Psychologica

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

Can gender priming eliminate the effects of stereotype threat? The case of simple dynamic systems

Vivien Lungwitz*, Peter Sedlmeier, Marcus Schwarz

Department of Research Methods and Evaluation in Psychology, Technische Universität Chemnitz, Germany

ARTICLE INFO

Keywords:
Dynamic systems
Gender
Priming
Stereotype threat
Stock–flow systems

ABSTRACT

Mathematics and mental rotation are classic fields where it has been shown that priming women with their gender identity impedes performance. Whereas past research focused mainly on stereotype threat effects in women in a narrowly defined context, this study broadened the research focus: We primed 264 women and men equally with a male, a neutral, or a female prime before they had to solve a simple dynamic system task. As expected, female-primed women subsequently performed worst of all six groups. Solution rates were almost 14% higher for the women in the male-primed condition. Men performed better than women in all three priming conditions. However, this difference was reduced in the male-primed condition as women's performance had increased as anticipated. Unexpected was a decline in the male performance in the same condition. The study showed that gender priming had a significant effect on women in tasks involving simple dynamic systems. However, mathematical knowledge and area of occupation clearly were stronger predictors for both men and women. Priming alone cannot eliminate the effects of stereotype threat.

1. Introduction

Mathematics, physics, and mental rotation are generally believed to be domains in which men perform better than women, and in fact past research has found this to be the case for mathematical tasks (Schmader, 2002; Spencer, Steele, & Quinn, 1999) and tasks involving spatial perception and mental rotation (for two meta-analyses see Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995). One possible explanation for this difference can be found in the concept of stereotype threat, which holds that fear of confirming a negative stereotype about the group to which they belong prevents people from reaching their full potential (Brannon, 2004; Shih, Pittinsky, & Ambady, 1999; Spencer et al., 1999; Steele & Aronson, 1995). A common stereotype is that women perform worse in math-related tasks than men. A gender-related stereotype threat effect was shown not only when comparing male and female participants but also for samples of women whose gender identity was made particularly salient compared to a group of women for whom their gender identity was not stressed (Cadinu, Maass, Rosabianca, & Kiesner, 2005; Krendl, Richeson, Kelley, & Heatherton, 2008; McGlone & Aronson, 2006; Schmader, 2002; Shih et al., 1999; Steele & Ambady, 2006). Inzlicht and Ben-Zeev (2000) showed that even the composition of a small group of people could trigger stereotype threat reactions, with women performing best in a women-only group. It seemed as if making their own sex salient to the women facilitated access to concepts such as “women can't do math” and caused deteriorated performance in mathematical tasks.

In any discussion about gender stereotype threat, one must remember that people differ in the degree to which they identify with their given sex and adopt a gender-conforming role. There are different ways to measure gender identity, including well-established self-report measures such as the Bem Sex Role Inventory (*BSRI*; Bem, 1974) and newer approaches that aim to assess perceived gender identity with graphical instruments (Schubert & Otten, 2002) or through implicit measures (Nosek, Banaji, & Greenwald, 2002), which can also be used to predict achievement in science and mathematics (Steffens, Jelenec, & Noack, 2010).

The above studies revealed that women performed worse on tasks involving math, mental rotation, and spatial orientation when they were made aware of their female gender. Yet these studies often lacked three things: First, many of them focused too much on women and neglected possible gender-priming effects on men. Second, only gender-congruent priming (women primed female but not primed male) was used. Ortner and Sieverding (2008) addressed these two drawbacks. They assessed the effects of male versus female primes on men and women for a mental rotation task. Men and women differed significantly in their performance only in the female-primed condition. In the male-primed condition, the gender effect vanished. A neutral control condition was not included, however. Nevertheless, gender-

* Corresponding author at: Technische Universität Chemnitz, Department of Psychology, Wilhelm-Raabe-Str. 43, 09107 Chemnitz, Germany.
E-mail address: vivien.lungwitz@psychologie.tu-chemnitz.de (V. Lungwitz).

https://doi.org/10.1016/j.actpsy.2018.05.010
Received 30 January 2017; Received in revised form 25 May 2018; Accepted 25 May 2018

congruent and gender-incongruent priming were shown to influence the results. A third drawback in past research is that the cognitive tasks examined seem to have been quite similar. Do the results also hold true for a broader context?

Instead of focusing narrowly on these same domains, we decided to systematically examine another: dynamic systems. Besides adding a new field to the test performance variables, we chose this domain for two other reasons: First, people deal with dynamic systems, such as money in a bank account, weight of a body, or water level in a bathtub, every day when they make personal decisions. Decisions involving dynamic systems can also affect others. For instance, if politicians or industrial tycoons are making decisions about, say, the carbon dioxide concentration in the atmosphere, millions of people can be affected. The simplest form of a dynamic system is a stock–flow system that includes one inflow, one outflow, and one stock. Understanding how flows change a stock over time in such a simple system is the basic key to understanding complex dynamic systems that also include aspects of feedback loops, nonlinearity, and time delays and are adaptive and self-evolving (Saleh, 2000; Sterman, 2002).

Our second reason for choosing simple dynamic systems is that stock–flow tasks have consistently shown gender effects, with men outperforming women (Booth Sweeney & Sterman, 2000; Brockhaus, Arnold, Schwarz, & Sedlmeier, 2013; Jensen & Brehmer, 2003; Kainz & Ossimitz, 2002; Kapmeier, 2004; Kasperidus, Langfelder, & Biber, 2006; Schwarz, Epperlein, Brockhaus, & Sedlmeier, 2013; Veldhuis & Korzilius, 2012). This gender effect seems to be independent of other task-specific or participant-specific factors (Schwarz, 2016).

Considering the above-mentioned three drawbacks of previous research —focusing on women, focusing on priming congruent to the given biological sex, and focusing on narrowly defined dependent variables — led us to our main research question: Is it possible to minimize or even eliminate performance differences between men and women on simple dynamic system tasks through priming? Taking past findings into consideration, we expected neutral-primed men to perform better than neutral-primed women. The more interesting question was how male and female priming would affect the task solution rates. In accordance with previous studies, poorer performance was expected for both men and women in the female-primed than in the male-primed condition, and we expected this to be more pronounced for the women because they were expected to experience a strong stereotype threat when confronted with their sex (Ortner & Sieverding, 2008). Furthermore, we anticipated that female-primed men would still perform somewhat better than male-primed women, as a few minutes of priming would be unlikely to overrule a stereotype that had developed over the course of 20 to 30 years.

2. Method

2.1. Participants

Gender effects in stock–flow research have been medium to large (Brockhaus et al., 2013; Kapmeier, 2004; Schwarz et al., 2013; Veldhuis & Korzilius, 2012). Taking these previous effects (using a mean of $d = 0.65$) and a 90% chance of finding an existing effect as the basis of an a priori power analysis for a one-tailed test in G*Power (version 3.1.2; Faul, Erdfelder, Lang, & Buchner, 2007), we determined we needed at least 42 participants per group for conducting *t*-tests between groups with a directed hypothesis. Although a *t*-test between two of the six groups was not necessary for every research question, we aimed at keeping group sizes comparable. Therefore, we collected data from 264 participants — 44 per group.

Participants were recruited via bulletin boards and e-mails or were approached personally on the university campus. Participants received 5 euros or got course credit for participating. The vast majority (90.9%) of participants were students (overall $M_{age} = 23.1$ years, $SD = 3.3$). Fourteen people had participated in an earlier stock–flow study. These participants' data patterns did not differ from the rest so they were left in the sample and not treated separately.

2.2. Materials

2.2.1. Priming texts

The study used three priming texts: one describing the day of a stereotypical male person, one describing the day of a stereotypical female person, and one neutral text describing a day in the mountains. The gender texts were taken from the study of Ortner and Sieverding (2008) and described either a caring, agreeable, and insightful woman or a determined, tough, and self-confident man, both “generally happy” with their life. Participants were asked to put themselves in the position of the described person and to write down adjectives to characterize themselves. For the neutral text they were asked to imagine the described landscape and to feel the atmosphere the text evoked. Then they should list adjectives characterizing the landscape. An independent pilot study where 20 volunteers (9 men, 11 women) rated the texts on a 15-point scale (from strongly male to strongly female — adapted from the workload scale of Schießl, 2007) revealed that the three texts were perceived as stereotypically male ($M = 11.75$, $SD = 1.92$), stereotypically female ($M = 3.90$, $SD = 2.17$), or neutral ($M = 7.90$, $SD = 2.57$) as intended. In the main study the priming task was labeled as an independent pilot study to measure the ability to understand someone else's position. After completing the whole questionnaire, participants were thoroughly debriefed and told the real purpose of the text.

2.2.2. Stock–flow tasks

To test participants' understanding of simple dynamic systems, we used three tasks of varying difficulty that have commonly been employed in stock–flow research (e.g., Booth Sweeney & Sterman, 2000; Cronin, Gonzalez, & Sterman, 2009). Flow information was presented in a line graph depicting the rates of inflow and outflow per minute with two distinguishable (e.g., dotted and solid) lines. There were no precise numbers given so that we could test participants' general understanding of the dynamic system rather than their calculation abilities. The three tasks and their solutions are shown in Fig. 1.

Following the procedure of Booth Sweeney and Sterman (2000), we gave participants a short introduction explaining the general concept of simple dynamic systems in a familiar context, in our case a bathtub with water in the tub, flowing into the tub, and draining out of the tub. The order of the three stock–flow tasks was completely counterbalanced across participants, each task followed by an empty diagram with an identical time axis in which participants were supposed to draw the change in stock over a 16-minutes time period.

2.2.3. BSRI-R

To compare participants' self-perceived gender identity in the different priming conditions, we used the revised German version of Bem's *BSRI* (Bem, 1974), the *BSRI*-R (Troche & Rammsayer, 2011). The inventory considers femininity and masculinity as two independent dimensions. The *BSRI*-R uses 15 items each to measure participants' self-concept regarding their own perceived masculinity and femininity. The items, which mainly consist of single adjectives such as “sensitive” or “logical,” need to be rated on a 7-point scale that ranges from *never applies* to *always applies*. The mean of all female and male item responses is calculated and then used as an indicator for the femininity and masculinity scale. The *BSRI*-R served as a manipulation check for the male and female priming conditions and was used to evaluate the influence of self-perceived gender identity on stock–flow performance in the neutral priming condition.

2.2.4. Other measures

After each stock–flow task, participants were asked to rate how difficult they perceived the task to be and how confident they were