Journal of Experimental Social Psychology 65 (2016) 26–30

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

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

A burden for the boys: Evidence of stereotype threat in boys' reading performance

Pascal Pansu [a,*], Isabelle Régner [b,**], Sylvain Max [c], Pascale Colé [b], John B. Nezlek [d,e], Pascal Huguet [f]

[a] Univ. Grenoble Alpes, LSE, Grenoble, France
[b] Aix Marseille Université, CNRS, LPC UMR 7290, 13331 Marseille, France
[c] Groupe ESC Dijon Bourgogne, 29 rue Sambin, BP 50608, 21006 Dijon Cedex, France
[d] Department of Psychology, P.O. Box 8795, College of William & Mary, Williamsburg, VA 23187-8795, USA
[e] SWPS University of Social Sciences and Humanities, Poznań, ul. gen. Tadeusza Kutrzeby 10, 60-995 Poznań, Poland
[f] Université Blaise Pascal, CNRS, LAPSCO UMR 6024, 63034 Clermont-Ferrand, France

HIGHLIGHTS

- We provide evidence for stereotype threat targeting boys in reading.
- Gender differences in reading can be reversed when the stereotype is made irrelevant.
- Boys underperformed girls when the task was presented as a reading test.
- Simply reframing the task as a game led boys to outperform girls.
- These findings matter given the importance of reading literacy in our societies.

ARTICLE INFO

Article history:
Received 21 July 2015
Revised 12 February 2016
Accepted 26 February 2016
Available online 3 March 2016

Keywords:
Stereotype threat
Children
Reading performance
Gender differences

ABSTRACT

There is ample evidence that Stereotype Threat (ST) contributes to gender differences favoring males on standardized math tests; however, whether ST also contributes to gender differences favoring females in reading remains unanswered. This is surprising as the gender gap in reading is three times bigger than the gender gap in math (OECD, 2014). In this study, we examined whether ST may explain gender differences favoring schoolgirls in reading, assuming that boys are negatively stereotyped in this domain. Eighty students (3rd grade) took a reading test while being assigned to either a threat or a reduced-threat condition (test presented as diagnostic of reading abilities versus as a game, respectively). Boys underperformed girls in the threat condition, whereas they outperformed girls in the reduced-threat condition. Consistent with ST theory, this pattern was obtained only among highly-identified students. These findings offer another explanation for the well-known gender gap favoring girls in reading.

Data collected by the Organisation for Economic Co-operation and Development (OECD)'s (2014) Program for International Student Assessment—PISA[1] (children from 34 OECD countries plus 41 partner countries), show that boys outperformed girls in math by an average of 11 points. Girls, however, outperformed boys in reading in every participating country and by an average of 38 points, the equivalent of an average school year's progress (See also Retelsdorf, Schwartz, and Asbrock, 2015; Stoet and Geary, 2013). Girls also read more frequently than boys, they have more positive attitudes about reading, higher reading motivation, and greater self-assurance about their reading skills than boys have (Hedges and Nowell, 1995; Logan and Johnston, 2009; McGeown, Goodwin, Henderson, and Wright, 2012; Reilly, 2012; Stoet and Geary, 2013). Although one may assume there to be intrinsic gender differences favoring females in reading ability and motivation, such intuitive accounts can be challenged by the role of the negative stereotype targeting boys in reading (Frome and Eccles, 1998; Hyde and Kling, 2001; Martinot, Bagès, and Désert, 2012; Retelsdorf et al., 2015).

Considerable research indicates that the threat of confirming a negative stereotype about one's group interferes with cognitive

* Correspondence to: P. Pansu, Université Grenoble Alpes, LSE, BP 47, 38040 Grenoble, France.
** Correspondence to: I. Régner, Aix Marseille Universite, CNRS, LPC UMR 7290, 13331 Marseille, France.
E-mail addresses: pascal.pansu@univ-grenoble-alpes.fr (P. Pansu), isabelle.regner@univ-amu.fr (I. Régner).

[1] Every 3 years since 2000, the Programme for International Student Assessment (PISA) has tested fifteen-year-old students from randomly selected schools worldwide in reading, mathematics, and science, with a focus on one subject each year. PISA is unique as it provides internationally comparable measures of student achievement.

http://dx.doi.org/10.1016/j.jesp.2016.02.008
0022-1031/

processes and leads to underperformance and stereotype confirmation (Régner et al., 2010; Schmader, Johns, and Forbes, 2008; Steele, 1997; Steele and Aronson, 1995). Nevertheless, most stereotype threat (ST) studies have addressed the gender gap in math favoring males, leaving the gender gap in reading favoring females unexplored. This is surprising because the gender gap in *reading* is three times the size of the gender gap in math. Some studies have examined the impact of the verbal-gender or academic achievement stereotypes, which both favor females, using explicit, if not directional, activation of those stereotypes. For example, participants were told that "gender differences are expected" or "girls are expected to do better than boys" (Hartley and Sutton, 2013; Keller, 2007; Seibt and Förster, 2004). No study has examined the specific impact of the reading-gender stereotype using subtle/implicit and nondirectional activation.

Here, we examine this impact in children, a population that has not received sufficient attention in ST research (with exceptions in the math domain, see Ambady, Shih, Kim, and Pittinsky, 2001; Huguet and Régner, 2007, 2009; for a review see Régner, Steele, Ambady, Thinus-Blanc, and Huguet, 2014), and we took into account their level of identification to reading. In adults, ST effects are typically stronger among individuals who highly identify with a subject matter, those who have much to lose in the event of poor performance (Steele, 1997; Walton and Cohen, 2003). Whether this also applies to children remains an unanswered question. Assuming that children may identify more or less with basic academic subjects such as reading, we expected 1) boys to underperform relative to girls under ST, while performing equally well in a reduced-threat condition; and 2) highly-identified students to be the most susceptible to ST effects.

1. Method

1.1. Participants

Participants were 80 third graders (9 years old, 48 boys) from four classes across three public elementary schools. All were French native speakers and were normal-readers, as assessed by the Alouette standard reading test (Lefavrais, 1967), with normal nonverbal intelligence, as assessed by Raven's Progressive Matrices (Raven, Raven, and Court, 1998). All children, parents, and teachers were given the opportunity not to participate (none refused). Consent and permission from all appropriate authorities were obtained.

1.2. Procedure

Children first took a standard reading test ("la pipe et le rat"; Lefavrais, 1986) designed to measure the recognition and comprehension of written words in children aged 6 to 20 years. This test is one of the few French tests assessing silent reading that can be administered in classroom settings (Colé, Duncan, and Blaye, 2014). Children took the test in their ordinary classroom setting. The experimental conditions were implemented at the classroom level, so that all the students in a specific class were randomly assigned to either the ST or reduced-threat condition. There were two classes in each condition. In the stereotype-threat condition, children were told by the experimenter (supposedly a reading teacher) that the task was a reading test designed to evaluate their "ability in reading". In the reduced-threat condition, children were told by the experimenter (supposedly a game designer) that the task was a new game called "the animal-fishing game" designed for a fun magazine. In both conditions, the test consisted of a silent reading task in which children had to underline, under time pressure (3 min), as many animal names as possible in a list of 486 words, half of which were animal names. Following the standard scoring system for this silent reading test (Lefavrais, 1986, p.18), children's performance was the number of animal names correctly underlined within 3-min, minus the total number of errors (words wrongly underlined and animals' names not underlined).[2]

After the reading test, children answered a series of questions about the test, including how important reading was to them (self-identification). All responses were made using 5-point scales with endpoints labeled strongly disagree (1) and strongly agree (5). Children indicated how interesting and difficult the test was ("I found this activity very interesting", "I found this activity difficult") and they evaluated their performance ("I did well in this activity"). To measure children's identification with reading, they answered two questions "Doing well on reading tasks is very important to me" and "I think it's very important to be good at reading" ($r(75) = .46$, $p < .001$). The mean response to these two items was computed, and higher scores indicated higher identification to reading. Descriptive statistics for all variables are presented in Table 1.

1.3. Analytical strategy

Although conceptually, the present data constituted a multilevel data structure (students nested within classrooms), we used Ordinary Least Squares regression analyses in which students were the unit of analysis. The primary reason for this was that four classes were not enough to estimate the parameters of the multilevel models that would be needed to test our hypotheses of interest using multilevel modeling (Nezlek, 2011, pp. 64–65). That being said, the means for the prime individual difference measures did not vary meaningfully across the classes (see text S1).

In our analyses, children's performance was regressed on gender (male = 1, female = −1), condition (reduced-threat = 1, threat = −1), identification to reading (mean-centered), and their interactions. This model was tested while controlling for children's level in reading (mean-centered) as assessed by the Alouette test, and its interaction with condition (Yzerbyt, Muller, and Judd, 2004). Our hypotheses led us to expect a Gender by Condition interaction, itself moderated by students' level of identification to reading. Testing this 3-way interaction implied that identification to reading (the moderator) was unaffected by Gender, Condition, and their interaction, which was indeed the case (Fs < 1; ps > .25).

Three participants were removed from this analysis, two with missing data (on either the covariates or the moderator) and one outlier (standardized residual exceeding +2.5), resulting in a sample of 77 participants. This sample size was adequate for the present analysis. It was determined a priori, as recommended by Tabachnick and Fidell (2007), on the basis of the desired power (.80), alpha level (.05), number of predictors (9 in the main analysis using covariates, 7 without), and anticipated size of the ST effect. Since the size of the ST effect was unknown in the reading domain, we used that found in math by Walton and Cohen's meta-analysis (2003) where f^2 is around .25 in highly identified participants. Soper's sample size calculator (Soper, 2013) indicated that the minimum required sample size for our multiple regression analysis was 72 (our initial sample size was slightly higher).

2. Results

2.1. Task performance

As expected, the Gender by Condition interaction was significant, $b = 4.12$, $t(67) = 2.81$, $p = .006$ (95% CI = [1.20, 7.05]), which was due to a ST effect unfavorable to boys. Boys underperformed relative to girls under threat condition ("reading test"), $b = -4.87$, $t(67) = -2.63$, $p = .011$, whereas they performed as well as girls in reduced-threat condition ("game"), $b = 3.37$, $t(67) = 1.48$, $p = .143$. It is noteworthy that girls' performance was not significantly higher in

[2] Students made very few errors ($M = .87$, $SD = 1.04$). Analyses of the number of items correct provided the same results as those we report.