European Journal of Social Psychology
Eur. J. Soc. Psychol. **37**, 586–597 (2007)
Published online 7 September 2006 in Wiley InterScience
(www.interscience.wiley.com) **DOI**: 10.1002/ejsp.379

Improving performance expectancies in stereotypic domains: Task relevance and the reduction of stereotype threat

HARRIET E. S. ROSENTHAL*,
RICHARD J. CRISP AND MEIN-WOEI SUEN
University of Birmingham, UK

Abstract

The experiments presented here extend previous research on reducing stereotype threat, along with examining the mediating role of performance expectancies. Women who generated shared academic characteristics between men and women predicted higher scores for themselves on a math test compared to the baseline and those who generated shared non-academic characteristics or shared physical characteristics. No effects were found for male participants' performance expectancies on an English test. Extending the relevance of these findings for stereotype threat research, women completing a math test, who first completed the shared academic characteristics task, both expressed higher performance expectancies and greater accuracy in math performance than participants in all other conditions. A partially mediating role of performance expectancies in relation to task and math performance was also found. Copyright © 2006 John Wiley & Sons, Ltd.

> *Ought not education to bring out and fortify the differences [between the sexes] rather than the similarities?*
>
> Virginia Woolf, *A Room of One's Own*

Education is a place where differences persist between the genders, with males and females often compelled to take gender-stereotypical subjects, so that more males than females choose to take math and physics, while more females opt to study French language and English (Department for Education and Skills, 2006). A growing body of work has addressed the conformity to such stereotypes, in the form of the phenomenon *stereotype threat*. Research has outlined the context and processes of stereotype threat, and more recently, means to reduce its negative impact. In this paper it is argued that to counter perceived differences between the genders in terms of academic abilities the focus must be on promoting similarity and not, as Virginia Woolf (an ardent feminist) ponders above, fortify the differences.

*Correspondence to: Harriet E. S. Rosenthal, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. E-mail: h.e.rosenthal@bham.ac.uk

Received 31 August 2005
Accepted 31 May 2006

STEREOTYPE THREAT

Although stereotype threat only came to light as an important psychological effect a decade ago (Steele & Aronson, 1995), research into the area is extensive. Stereotype threat is the predicament individuals face in certain situations where they risk being judged in terms of a negative stereotype associated with their group membership (Steele, 1997). For instance, it may be the case that fewer young women take math and physics subjects at high school, and fewer young men take English because of the stereotype that their gender does not excel at these subjects.

Stereotype threat has been studied in a number of settings. Steele and Aronson (1995) examined the threat of inferior intellectual ability attributed to African Americans. Informing White and Black participants that a test they were about to take was indicative (diagnostic) of intellectual ability resulted in Black participants underperforming on the test, compared to Black participants who were told that the test was not indicative of intelligence (non-diagnostic), and White participants in both conditions.

Besides contexts involving African Americans and the stereotype of intelligence (Aronson, Fried, & Good, 2002; McKay, Doverspike, Bowen-Hilton, & Martin, 2002), other researchers have found similar effects in additional stereotyped domains, such as low socio-economic status and intelligence (Croizet & Claire, 1998; Good, Aronson, & Inzlicht, 2003), homosexual men and childcare (Bosson, Haymovitz, & Pinel, 2004), men and affective tasks (Leyens, Désert, Croizet, & Darcis, 2000), women and career choice (Davies, Spencer, Quinn, & Gerhardstein, 2002; Rosenthal & Crisp, 2006) and women and mathematics (Ambady, Paik, Steele, Smith, & Mitchell, 2004; Brown & Pinel, 2003; Gonzales, Blanton, & Williams, 2002; Inzlicht & Ben-Zeev, 2000; Josephs, Newman, Brown, & Beer, 2003; Keller & Dauenheimer, 2003; Marx & Roman, 2002; McIntyre, Paulson, & Lord, 2003; Pronin, Steele, & Ross, 2004; Rosenthal, & Crisp, 2006). The latter is the most highly researched area with regard to stereotype threat and is one of two domains that is examined here. Typically, when women are informed that the genders will be compared on a mathematics test they under-perform. The experiments reported here aimed to investigate if the tendency for women to experience stereotype threat in the math domain could be reduced by focusing on a presumed part of the process leading to stereotype threat: performance expectancies.

PERFORMANCE EXPECTANCIES AND STEREOTYPE THREAT

A number of underlying processes have been investigated as potentially accounting for stereotype threat effects. Some of these potential mediators include anxiety (Bosson, Haymovitz, & Pinel, 2004), working memory capacity (Schmader & Johns, 2003), mental load (Croizet et al., 2004), intrusive thoughts (Cadinu, Maass, Rosabianca, & Kiesner, 2005) and performance expectancies (Cadinu, Maass, Frigerio, Impagliazzo, & Latinotti, 2003). It is the last proposed mediator that we concentrate on here. If performance expectancies are a component of why stereotype threat occurs, it follows that if negative performance expectancies can be attenuated, so too may stereotype threat.

The role of performance expectancies in contributing to stereotype threat has received mixed support, which indicates the need to examine its mediating role further. For instance, Stone, Lynch, Sjomeling, and Darley (1999) found that participants under stereotype threat displayed lower performance expectancies, reporting that they would require a larger number of attempts to complete a golf course than non-threatened participants. However, due to a possible confound between *previous* performance and expectancies, they concluded that the results were inconclusive. Similarly, Cadinu et al. (2003) found a partially mediating role of expectancy in relation to stereotype threat and

DOI: 10.1002/ejsp

performance for female students on a mathematics test. To date this study is the most supportive of performance expectancies as having an explanatory role in stereotype threat.

In sum, it is apparent from previous research that performance expectations may play a key role in stereotype threat, but that there is currently limited support for this role. In an attempt to shed some further light on this issue, in the experiments reported here, expectancies concerning math and english test performance were focused on, in the context of developing an intervention designed to *reduce* stereotype threat.

REDUCING STEREOTYPE THREAT

The research presented here predicted that placing participants under conditions conductive of stereotype threat would result in these participants predicting lower performance expectancies. However, it was also predicted that such an effect could be reduced by using a task successfully tested by Rosenthal and Crisp (2006). Specifically, asking participants to state characteristics shared between women and men reduced stereotype threat in terms of both stereotypic career choice and math performance. Female participants who completed the shared characteristics task chose less stereotypical careers (Experiment 1) and completed more math questions correctly than participants in the baseline condition (Experiments 2 and 3).

Drawing on intergroup discrimination research, there are good theoretical reasons why thinking about shared characteristics between ingroups and outgroups could reduce stereotype threat effects. Doise's (1978) *category differentiation model* and the *meta-contrast process* (Oakes, Haslam, & Turner, 1994) outlined by *self-categorization theory* (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987) argue that categorization into *them* and *us* provides the pre-requisite for intergroup discrimination. Models of bias-reduction, including the *common ingroup identity model* (Gaertner & Dovidio, 2000; Gaertner, Mann, Dovidio, Murrell, & Pomare, 1990; Gaertner, Mann, Murrell, & Dovidio, 1989) and *crossed categorization* (Crisp, Ensari, Hewstone, & Miller, 2002; Deschamps & Doise, 1978) are based on the idea that it is the differences between ingroups and outgroups that are the basis for bias and, correspondingly, if this perceived difference is reduced, intergroup evaluations become less biased. This contention is well supported using a variety of paradigms all of which reduce the salience of the intergroup boundaries by blurring the distinction between ingroups and outgroups (e.g., Crisp, Hewstone, & Rubin, 2001; Marcus-Newhall, Miller, Holtz, & Brewer, 1993; Vanbeselaere, 1987; for recent reviews see Crisp & Hewstone, 1999; Gaertner & Dovidio, 2000; Mullen, Migdal, & Hewstone, 2001). Rosenthal and Crisp (2006) extended this logic to the stereotype threat domain, arguing that just as categorization models of bias-reduction work by weakening the *us* versus *them* prerequisite for bias, such tasks should also reduce the tendency for participants to experience stereotype threat. One cannot conform to a stereotype based on a distinction between *us* and *them*, if *they* are not perceived as essentially different from *us*. Three studies supported this idea, finding a reduction in stereotype threat when participants thought about characteristics shared between the ingroup (women) and outgroup (men).

The experiments reported here aimed to both extend this stereotype threat-reduction model to include the role of performance expectancies, and to refine the model in terms of the shared task. Rosenthal and Crisp (2006) asked participants simply to generate *any* characteristics that could be shared between women and men. However, one may expect the effectiveness of the model to vary depending on the *relevance* of the characteristics generated. In the same way that for stereotype threat

DOI: 10.1002/ejsp

to occur, the situation, e.g., taking a math test, must be relevant to the stereotype (Steele, 1997), so too the effectiveness of interventions to reduce stereotype threat could be contingent on relevance.

EXPERIMENT 1

The current experiment compared the effectiveness of different versions of the shared characteristics task previously employed by Rosenthal and Crisp (2006; and for bias-reduction, Crisp & Beck, 2005). Specifically, the effects of asking participants to generate shared *academic* characteristics, shared *non-academic* characteristics and shared *physical* characteristics were compared. It was hypothesized that asking participants to generate shared academic characteristics would have the most dramatic effect on reducing performance expectancies. Highlighting shared characteristics between men and women that involve similarities within education (academic) should be more effective at reducing stereotype threat than thinking about similarities between the genders that exist outside of education.

This hypothesis was tested in two domains. As previously discussed, a stereotype exists that men are better at math than women, while an additional stereotype suggests that women are better at English than men. Compared to a baseline condition, it was predicted that generating shared academic characteristics between women and men would have a greater effect on improving performance expectancies, compared to non-academic and physical characteristics.

Method

Participants and Design

Two hundred and forty male ($N = 120$) and female ($N = 120$) students at the University of Birmingham (mean age = 21, $SD = 3.5$) were randomly allocated to one of four shared characteristics conditions (baseline [no task] vs. physical vs. non-academic vs. academic) and one of two test domains (math vs. English). Participants received £1.50 (€2.16; $2.63) for their participation.

Procedure

The shared academic characteristics condition was refined from instructions used by Rosenthal and Crisp (2006). Participants in the academic condition were asked to 'think of three things within the academic setting (i.e., within education) that men and women can have in common (i.e., characteristics within education that men and women share).' Below the instructions were three numbered spaces for participants to write the shared characteristics. In the non-academic condition, participants read that they were to 'think of three things outside the academic setting (i.e., outside of education) that men and women can have in common (i.e., characteristics outside of education that men and women share).' In the physical condition the participants were asked to 'think of three physical characteristics that men and women can have in common (i.e., physical characteristics that men and women share).' In the baseline condition participants did not complete a characteristics task and simply moved straight on to the second part of the questionnaire.

From this point, all participants received exactly the same instructions. The second part of the questionnaire involved the description of an academic scenario (adapted from Suen, 2006). Female

DOI: 10.1002/ejsp

participants in the math test domain read the instructions: 'Please try your best to imagine that you are in a university final exam with other students. Today, the subject of the exam is mathematics. You need to spend 60 minutes in this exam. As you look around you realise that the other eight students are all male. Please circle your answers to the following questions.' All women read that the other students were men, while all men read that the other students were women. Half of the male participants were informed that it was a mathematics exam, while the other half were informed that it was an English exam. The same was the case for female participants. Participants then completed a set of questions relating to performance expectancies, followed by some demographic questions.

Dependent Measures

The performance expectancies measurement consisted of two simple questions. Participants were asked to give the score they *themselves* expected to get (out of 100) and the score they expected the *other* people in the examination to get (out of 100). This allowed a test of whether the locus of effect stemmed not only from their own performance expectancies, but from performance relative to the relevant outgroup.

Results and Discussion

Twelve participants were removed from the analysis due to incomplete dependent measures. A 2 (participant gender: male vs. female) × 2 (test domain: math vs. English) × 4 (condition: baseline vs. physical vs. non-academic vs. academic) × 2 (target score expectancies: self vs. others) mixed ANOVA with repeated measures on the last factor was computed revealing three effects. First, there was a main effect of target. Overall, participants believed they would score lower on the test ($M = 64.16$) compared to the imagined others ($M = 67.61$), $F(1, 212) = 12.09$, $p = 0.001$. Second, there was a domain × target interaction, $F(1, 212) = 7.19$, $p = 0.008$. While participants felt that they would do equally well as the imagined others on the English test (Ms = 65.97 and 66.75, respectively), on the math test participants overall felt they would score lower ($M = 62.36$) compared to the imagined others ($M = 68.48$). These effects were, however, qualified by a significant four-way interaction between gender, test domain, condition and target, $F(3, 212) = 2.71$, $p = 0.046$, see Table 1. We decomposed this interaction by analyzing the effects within the math and English domains separately.

Table 1. Performance expectancies as a function of task, test domain, and participant gender

			Performance expectancies							
			Baseline		Physical		Non-academic		Academic	
			M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Math	Female	Self	53.20	24.33	60.62	19.72	59.20	18.43	72.21	10.17
		Others	69.00	10.32	66.15	9.82	70.20	9.81	71.36	9.79
	Male	Self	62.00	17.56	64.27	19.81	66.00	15.14	61.69	21.17
		Others	63.07	8.43	65.40	11.51	69.80	10.14	73.00	10.74
English	Female	Self	61.87	6.33	67.14	11.22	68.33	8.54	63.57	14.86
		Others	62.53	7.16	65.71	10.54	68.80	7.64	68.14	7.08
	Male	Self	71.15	5.79	65.43	12.08	65.23	10.08	65.53	13.64
		Others	71.46	7.60	65.13	8.92	66.38	5.58	66.40	10.90

DOI: 10.1002/ejsp

Separate 2 (gender) × 4 (condition) × 2 (target score) mixed ANOVAs with repeated measures on the last factor were carried out for each domain. In the English domain, there were no effects, except a theoretically uninteresting condition × gender interaction, $F(3, 106) = 3.24$, $p = 0.025$. The absence of effects in the English test domain is considered in the general discussion. However, in the math domain there was a significant gender × condition × target interaction, $F(3, 106) = 2.82$, $p = 0.043$. Decomposing this effect further, no effects were found for male participants, but a significant main effect of target was found for female participants, $F(1, 53) = 12.14$, $p = 0.001$. On the math test female participants felt overall that they would score lower ($M = 61.14$) compared to the imagined others (i.e., males, $M = 69.25$). This effect was, however, qualified by a condition × target interaction for females that approached significance, $F(3, 53) = 2.56$, $p = 0.065$.

Further analysis was therefore completed on the results for female participants by analyzing the effects of condition on the scores for self and others separately. A set of Helmert contrasts was used to incrementally test whether there was support for the specific hypotheses regarding the optimal effectiveness of generating shared academic characteristics compared to physical or non-academic characteristics, testing both predicted scores as a function of condition. With respect to self scores, a difference was first tested between the baseline and physical conditions (baseline −1, physical +1, non-academic, 0, academic, 0), revealing no significant difference, $t(53) = 1.03$, $p = 0.307$. Given no difference between these two conditions, their aggregate was compared with the non-academic condition (baseline −1, physical −1, non-academic, +2, academic, 0). This analysis also revealed no differences, $t(53) = 0.377$, $p = 0.707$. Finally, given no differences between the baseline, physical, and non-academic conditions, all three conditions were compared with the academic condition (baseline −1, physical −1, non-academic, −1, academic, +3). This analysis revealed a significant effect, $t(53) = 2.49$, $p = 0.016$. In line with predictions, it was specifically the generation of shared academic characteristics between women and men that was most effective at improving performance expectancies. Notably, in contrast, there were *no* significant differences with respect to performance expectancies for the imagined others (male).

To summarize, women who generated shared academic characteristics between the genders improved their performance expectancies within a math context. In fact, examining the comparison of the self and others scores within each condition, shared academic characteristics was the only condition in which a higher self score ($M = 72.21$) was found compared to others score ($M = 71.36$), although this trend was non-significant. Although there was also no significant difference between the self and others scores in the physical characteristics condition, the scores remained in the direction of the non-academic and baseline conditions, with the self-score ($M = 60.62$) remaining lower than the others score ($M = 66.15$). This suggests that when women are presented with the threat of a math test, they express low performance expectancies. In turn, the introduction of a task which focused on similarities within the academic domain raised these performance expectancies.

EXPERIMENT 2

Experiment 1 established that women completing a shared characteristics task within a math context expressed higher (self) performance expectations compared to all other conditions. However, the experiment looked at performance expectancies but did not go on to examine actual performance on a math test. Therefore, in order to relate this finding to an actual stereotype threat situation, Experiment 2

DOI: 10.1002/ejsp

focused on examination of actual math scores in relation to the shared characteristics task, along with the role of performance expectancies as a potential mediator of the effect.

Method

Participants and Design

Forty-eight female psychology students at the University of Birmingham (mean age = 20, $SD = 1.67$) were randomly allocated to one of four shared characteristics conditions (baseline [no task] vs. physical vs. non-academic vs. academic). Participants received course credit or £4.00 (€5.88; $6.98) for their participation.

Procedure

As in Experiment 1, participants were asked to complete one of three shared characteristics tasks (academic; non-academic; physical), before receiving instructions about the math task and the performance expectancies measure. Participants in the baseline condition did not complete a characteristics task, receiving the instructions and performance expectancies measure as their first task. The instructions relating to the math task were refined from Experiment 1. As the participants were required to take an actual math test, as opposed to an imagined test, the scenario in Experiment 1 was modified in order to retain the comparison with male students, while appearing a realistic situation. The participants were informed (both verbally and through written instructions):

'The test you are about to take assesses your current mathematical ability in comparison to other students. The aim of the test is to compare the maths performance of women and men. Therefore, your score on this test will be compared to male students'. This manipulation was similar to the manipulation used Experiment 1 in that the comparison group was not actually present in the testing situation. However, it was felt that the manipulation of Experiment 1, which involved participants *imagining* a situation, could be construed as un-realistic. Therefore, to present a situation which allowed a similar comparison to male students, but would appear more realistic as a testing situation, the imagination element was removed.

Participants were then asked to predict the score they expected to get out of 100, and the score they expected male students to get out of 100. After completion of the performance expectations items, participants received the actual math test. This consisted of ten math items (taken from Rosenthal & Crisp, 2006), which the participants were given 5 minutes to complete. To correspond with the performance expectancies measure it was stated that each correct answer was worth 10 marks. After completion of the math test, participants were given a demographic information sheet to complete before being thanked and debriefed.

Results and Discussion

One participant was removed from the analysis due to suspicion that the task was investigating stereotype threat.

DOI: 10.1002/ejsp

Performance Expectancies

The results of Experiment 1 suggested that the academic shared characteristics task improved the performance expectancies associated with the individual, while it did not affect performance expectancies associated with the male students. Therefore, it was hypothesized in Experiment 2 that the shared task would affect self-scores and not others scores.[1] As expected, a one-way ANOVA revealed no significant difference between conditions for others performance expectancies, $F(3, 43) = 1.13$, $p = 0.348$. However, a one-way ANOVA for self performance expectancies revealed a significant difference between conditions, $F(3, 43) = 2.96$, $p = 0.043$, see Table 2. As in Experiment 1, a set of Helmert Contrasts was used to test the optimal effectiveness of generating shared characteristics. No significant difference was found between the baseline and physical conditions (baseline -1, physical $+1$, non-academic 0, academic, 0), $t(43) = -1.40$, $p = 0.168$. Their aggregate was then compared to the non-academic condition (baseline -1, physical -1, non-academic $+2$, academic 0), again revealing no significant difference $t(43) = 1.44$, $p = 0.158$. Finally, all three conditions were compared to the academic condition (baseline -1, physical -1, non-academic -1, academic $+3$), revealing a significant effect, $t(43) = 2.27$, $p = 0.029$. This suggests that it is specifically listing shared academic characteristics that improves performance expectancies, which is in line with the findings of Experiment 1.

Math Test

A one-way ANOVA for accuracy (number correct/number attempted) revealed a significant difference between conditions, $F(3, 43) = 3.47$, $p = 0.024$. The same Helmert contrasts were used as for performance expectancies, revealing no significant difference between the baseline and physical conditions (baseline -1, physical $+1$, non-academic 0, academic 0), $t(43) = -1.38$, $p = 0.173$. These conditions were then compared to the non-academic condition (baseline -1, physical -1, non-academic $+2$, academic 0), revealing a difference that approached significance, $t(43) = 1.89$, $p = 0.66$. This indicated a tendency for even non-academic shared characteristics to improve performance. All three conditions were then compared to the academic condition (baseline -1, physical -1, non-academic -1, academic $+3$), revealing a significant effect, $t(43) = 2.29$, $p = 0.027$.

Therefore, it is the listing of shared characteristics which improves math performance and reduces stereotype threat. As the results for math performance are in line with the results for performance expectancies, a mediational analysis was carried out to establish the role of performance expectancies in relation to stereotype threat.

Mediational Analysis

To explore whether the performance on the math test could be explained by performance expectancies, a mediational analysis was carried out. The contrast used above acted as the predictor (baseline -1, physical -1, non-academic -1, academic $+3$), the math performance was the outcome, and the self-performance expectancy was the mediator (see Figure 1). The analysis revealed a significant association between the predictor and the outcome (math performance), $\beta = 0.308$ $p = 0.035$. The

[1]We note that a 4 (condition) $\times$ 2 (target score) mixed ANOVA with repeated measures on the second factor did not reveal a significant interaction, $F(3, 43) = 1.41$, $p = 0.254$. However, consistent with our use of contrast analysis in Experiment 1 (and our strong a priori theoretical basis for predicting a precise pattern of differences on self, but not other, expectancies), we continued to explore the data for self and other targets respectively.

DOI: 10.1002/ejsp

Table 2. Performance expectancies and math performance as a function of task

	Baseline		Physical		Non-academic		Academic	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Self score	54.17	16.07	45.73	14.66	57.33	16.73	63.33	8.88
Others score	67.50	6.57	66.36	11.85	69.00	16.10	74.17	7.33
Math accuracy	0.810	0.168	0.729	0.135	0.864	0.141	0.909	0.111

relation between the predictor and the mediator (performance expectancies) was also significant $\beta = 0.309$, $p = 0.035$, as was the path between the mediator (performance expectancies) and the outcome (math performance) controlling for the effect of task, $\beta = 0.445$, $p = 0.002$. The significant relation between the predictor and the outcome (math performance) became non-significant when the mediator (performance expectancies) was entered into the first equation, $\beta = 0.170$, $p = 0.214$, but most importantly, a Sobel test approached significance, $Z = 1.82$, $p = 0.069$.

Therefore, the analysis suggests a partially mediating role of performance expectancies in relation to stereotype threat. It appears as though the relevance of the shared characteristics task (i.e., academic characteristics within an academic stereotype threat situation) works to increase performance expectancies, which in turn reduces stereotype threat by increasing math performance.

GENERAL DISCUSSION

The experiments reported here had two aims. The first was to examine the role of performance expectancies in stereotype threat, and the second to extend recent work revealing that the generation of shared characteristics between ingroups and outgoups can reduce stereotype threat (Rosenthal & Crisp, 2006). This work was extended through specifically manipulating the *type* of generated characteristics, with the hypothesis that characteristics associated with the relevant domain (academic) would be more likely to reduce stereotype threat. The impact of three different types of characteristic were examined (academic vs. non academic vs. physical) compared to a baseline condition, finding that emphasizing academic characteristics had the greatest beneficial effect on increasing performance expectancies and reducing stereotype threat.

The first experiment examined performance expectancies within both the math and English test domains, for both women and men. Although women, when compared to men on a math test, had lower performance expectancies in the baseline compared to the shared characteristics condition, no evidence

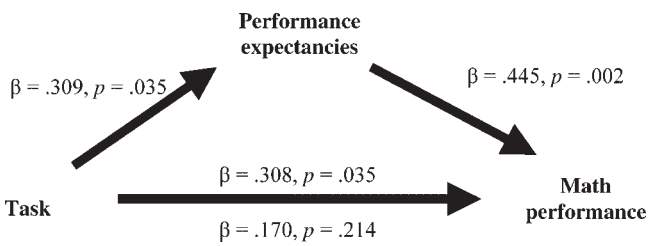

Figure 1. Mediational model of the role of performance expectancies in explaining the impact of task on math performance

DOI: 10.1002/ejsp

was found that males faced with an English exam, and a comparison to women, predicted their scores to be lower, compared to males who emphasized shared characteristics. This suggests a difference between the stereotypes for women and math and men and English. While there is much evidence in the stereotype threat literature for women coming under threat when the genders are compared on a math test, this effect has so far not been examined for males specifically within the English test domain. Stereotype threat has previously be found for men with regard to other domains, such as White males' inferiority to Asian males on mathematics tests (Aronson et al., 1999) and, specifically with regard to gender stereotypes, the effect has been found with social sensitivity (Koenig & Eagly, 2005) and affective tasks (Leyens, Désert, Croizet, & Darcis, 2000). However, the effect has not been empirically established in the English test domain. Therefore, perhaps the non-significant results suggest that the stereotype for men and English is not as strong as for women and mathematics (see also Rosenthal, 2006).

Also, much stereotype threat research has noted apparent stereotype *lift* patterns, where the non-threatened group actually improve their performance, e.g., when taking a mathematics test, men will actually perform *better* when they are informed that the genders will be compared, opposed to when they are not informed of the comparison. Although stereotype lift patterns often fail to reach significance, the pattern is a well-observed aspect of stereotype threat (for a review see Walton & Cohen, 2003). Experiment 1 contained two situations which lent well to stereotype lift conditions, although no evidence was found for the effect. Following a stereotype lift pattern, women taking an English test and men taking a math test, may have predicted their scores to be higher in the baseline condition compared to the other conditions, but this was not the case. However, it is important to note that the experiment assessed performance expectancies, not performance, which may account for the lack of evidence for stereotype lift.

As the results of Experiment 1 suggested that the shared characteristics task affected women's performance expectancies on a math test, in Experiment 2 this scenario was applied to women taking an actual math test. It was found that women completing the shared characteristics task both predicted higher performance expectancies and were more accurate on a math test than women in all other conditions. In turn, performance expectancies were found to partially mediate math performance, and therefore stereotype threat.

These findings have numerous theoretical and applied implications. The shared characteristics task employed in our previous research was found to become increasingly effective by enhancing its relevance to the threatened domain. This follows the logic that if the stereotype threat associated with women and math lies specifically within stereotypic preconceptions of differences between women and men in the academic domain, then focusing the shared characteristics task within this domain should have a greater effect. This relevance of the task to the threatened domain is perhaps not surprising, considering that one of the requirements of the emergence of stereotype threat involves the relevance of the situation to the stereotype. A person must be in a relevant situation (e.g., taking a math test) for stereotype threat to emerge (Steele, 1997). As relevance of the situation is necessary for the effect to occur, it seems appropriate that to reduce the effect the task should also be relevant to the situation.

In turn, these findings have some practical applications for applying such stereotype threat reducing techniques. When attempting to reduce stereotype threat outside of the academic domain, the task should be modified to increase its relevance to the alternative setting. For example, emphasizing shared physical characteristics may be the best application within the sporting arena (Stone, 2002; Stone, Lynch, Sjomeling, & Darley, 1999) when reducing the inferiority felt by White athletes.

To conclude, this work expands on previous research by finding that the effectiveness of emphasizing shared characteristics can be improved by increasing the relevance of the task to the threatened domain. When focusing on similarities between the ingroup (female) and outgroup (male), the shared characteristics should focus on the relevant domain, which for mathematics is academia. Such a technique was found to increase perceived performance expectations, compared to a baseline

DOI: 10.1002/ejsp

condition and conditions which emphasized other (non-academic and physical) characteristics, and in turn reduce stereotype threat. These findings therefore add to the growing literature that elaborates on the conditions that one should encourage, and avoid, to counter situations in which stereotype threat might occur, and enable people to break free from negative performance expectations.

ACKNOWLEDGEMENTS

Thanks to Mara Cadinu and three anonymous reviewers for their useful and insightful comments on a previous version of this manuscript.

REFERENCES

Ambady, N., Paik, S. K., Steele, J., Smith, A. O., & Mitchell, J. P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology, 40*, 401–408.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*, 113–125.

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When white men can't do maths: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*, 29–46.

Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology, 40*, 247–255.

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.

Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology, 33*, 267–285.

Cadinu, M. R., Maass, A., Rosabianca, A., & Kiesner, J., (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science, 16*, 572–578.

Crisp, R. J., & Beck, S. (2005). Reducing intergroup bias: The moderating role of ingroup identification. *Group Processes & Intergroup Relations, 8*, 173–186.

Crisp, R. J., Ensari, N., Hewstone, M., & Miller, N. (2002). A dual-route model of crossed categorisation effects. In W. Strobe, & M. Hewstone (Eds.), *European review of social psychology* (Vol. 13, pp. 35–73). Hove, UK & Philadelphia, PA: Psychology Press.

Crisp, R. J., & Hewstone, M. (1999). Differential evaluation of crossed category groups: Patterns, processes and reducing intergroup bias. *Group Processes & Intergroup Relations, 2*, 307–333.

Crisp, R. J., Hewstone, M., & Rubin, M. (2001). Does multiple categorization reduce intergroup bias. *Personality and Social Psychology Bulletin, 27*, 76–89.

Croizet, J.-C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin, 24*, 588–594.

Croizet, J.-C., Després, G., Gauzins, M.-E., Huguet, P., Leyens, J.-P., & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin, 30*, 721–731.

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*, 1615–1628.

Department for Education and Skills. (2006). *GCE/VCE A/AS examination results for young people in England, 2004/05 (revised)* (SFR 01/2006). London, UK: Department for Education and Skills.

Deschamps, J.-C., & Doise, W. (1978). Cross category memberships in intergroup relations. In H. Tajfel (Ed.), *Differentiation between social groups: Studies in the social psychology of intergroup relations* (pp. 141–158). London: Academic Press.

Doise, W. (1978). *Groups and individuals: Explanations in social psychology*. Cambridge: Cambridge University Press.

DOI: 10.1002/ejsp

Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The common ingroup identity model*. Philadelphia, PA: The Psychological Press/Taylor & Francis.
Gaertner, S. L., Mann, J. A., Dovidio, J. F., Murrell, A. J., & Pomare, M. (1990). How does cooperation reduce intergroup bias. *Journal of Personality and Social Psychology, 59*, 692–704.
Gaertner, S. L., Mann, J., Murrell, A., & Dovidio, J. F. (1989). Reducing intergroup bias: The benefits of recategorization. *Journal of Personality and Social Psychology, 57*, 239–249.
Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and social psychology bulletin, 28*, 659–670.
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*, 645–662.
Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*, 365–371.
Josephs, R. A., Newman, M. L., Brown, R. P., & Beer, J. M. (2003). Status, testosterone, and human intellectual performance: Stereotype threat as status concern. *Psychological Science, 14*, 158–163.
Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin, 29*, 371–381.
Koenig, A. M., & Eagly, A. H. (2005). Stereotype threat in men on a test of social sensitivity. *Sex Roles, 52*, 489–496.
Leyens, J.-P., Désert, M., Croizet, J.-C., & Darcis, C. (2000). Stereotype threat: Are lower status and history of stigmatization preconditions of stereotype threat. *Personality and Social Psychology Bulletin, 26*, 1189–1199.
Marcus-Newhall, A., Miller, N., Holtz, R., & Brewer, M. B. (1993). Cross-cutting category membership with role assignment: A means of reducing intergroup bias. *British Journal of Social Psychology, 32*, 125–146.
Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin, 28*, 1183–1193.
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*, 83–90.
McKay, P. F., Doverspike, D., Bowen-Hilton, D., & Martin, Q. D. (2002). Stereotype threat effects on the raven advanced progressive matrices scores of African Americans. *Journal of Applied Social Psychology, 32*, 767–787.
Mullen, B., Migdal, M. J., & Hewstone, M. (2001). Crossed categorization vs. simple categorization and intergroup evaluations: A meta-analysis. *European Journal of Social Psychology, 31*, 721–736.
Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality.* Oxford: Blackwell.
Pronin, E., Steele, C. M., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology, 40*, 152–168.
Rosenthal, H. E. S. (2006). Reducing stereotype threat: A category differentiation approach. Unpublished PhD thesis.
Rosenthal, H. E. S., & Crisp, R. J. (2006). Reducing stereotype threat by blurring intergroup boundaries. *Personality and Social Psychology Bulletin, 32*, 501–511.
Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*, 440–452.
Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629.
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.
Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on black and white athletic performance. *Journal of Personality and Social Psychology, 77*, 1213–1227.
Stone, J. (2002). Battling doubt by avoiding practice: The effects of stereotype threat on self-handicapping in white athletes. *Personality and Social Psychology Bulletin, 28*, 1667–1678.
Suen, M.-W. (2006). Stereotype-moderated math performance in multiple category contexts. Unpublished PhD thesis.
Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Re-discovering the Social Group: A Self-categorization Theory.* Oxford: Blackwell.
Vanbeselaere, N. (1987). The effects of dichotomous and crossed social categorizations upon intergroup discrimination. *European Journal of Social Psychology, 17*, 143–156.
Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology, 39*, 456–467.

DOI: 10.1002/ejsp