Running head: REDUCING STEREOTYPE THREAT

Reducing stereotype threat by blurring intergroup boundaries

Harriet E. S. Rosenthal and Richard J. Crisp

University of Birmingham

Word Count: 9,131

Abstract

We aimed to establish whether interventions designed to reduce intergroup bias could be applied to the stereotype threat domain. In three experiments we tested the hypothesis that blurring intergroup boundaries would reduce stereotype threat. In the first study we found that female participants who thought about characteristics shared between the genders tended to show less preference for stereotypical female careers than participants in the baseline condition. In Experiment 2 participants who thought about overlapping characteristics answered more math questions correctly compared to a baseline group and participants who though about differences between the genders. In Experiment 3 we included a specific threat manipulation. Participants who completed the overlapping characteristics task before receiving the threat completed significantly more math questions correctly than participants in the baseline and threat conditions. The findings support the idea that interventions designed to reduce intergroup bias can be applied successfully in the reduction of stereotype threat.

Keywords: STEREOTYPE THREAT, SOCIAL CATEGORIZATION

No woman in my time will be prime minister or chancellor or foreign secretary - not the top jobs. Anyway, I wouldn't want to be prime minister; you have to give yourself 100 percent.

Margaret Thatcher (British Prime Minister 1979-1990)

Sunday Telegraph, 26th October 1969

Margaret Thatcher in 1969 could not perceive herself, or another woman, in the role of British prime minister. Her declaration of disinterest can be seen as a reaction to the perception that the prime ministerial role is typically regarded as a male preserve. Thatcher was elected as prime minister from 1979 to 1990, but she remains the only woman to have held that office.

The role of women in society has diversified rapidly over the last thirty years, yet men and women in Britain continue to follow career paths stereotypical of their gender. Two-thirds of managers and senior officials are male, while four out of five people in administrative and secretarial roles are female. Likewise, 92% of people in skilled trades are male and 84% of personal service roles (e.g. healthcare; childcare; hairdressers) are filled by females (Equal Opportunities Commission [EOC], 2004). This precedent for stereotypical roles is reflected in the academic subjects chosen by 16 year olds at school, for GCSE qualifications. In the United Kingdom (with the exception of Scotland) young people are required to study certain subjects (English, math, science and a modern language), however the genders diversify in subject areas they choose; for instance, 67% of physical education students are male, while 95% of home economics students are female (Department for Education and Skills [DfES], 2004).

This choice of gender stereotypical subjects extends beyond GSCE level into further qualifications. For students who do not go on to study for A-levels (qualifications typically studied aged 16-18), young people have a number of choices in terms of qualifications and employment. One route is to take on a modern apprenticeship, which are government supported positions offering on-the-job training. The stereotypical nature of the apprenticeships chosen are very pronounced, with women only forming one percent of apprenticeships in the construction industry and men forming three percent in the childcare sector (EOC, 2004). For students who do go on to study for A-levels (which do not require the study of core subjects, as at GCSE level), these again reflect a stereotypical bias; with 63% of math

students and 77% of physics students consisting of males, and 70% of English and 67% of French language students consisting of females (DfES, 2004). Such differences in the number of women and men in certain academic subjects continue into universities. Recent debate has been sparked by the comments of Harvard President Lawrence Summer, who suggested that women's under-representation in science and engineering is due to innate differences in abilities. The research presented here dismisses a focus on biological factors, and instead examines an alternative contributor to these stereotypic trends in career and academic subject choices, that of *stereotype threat.*

Stereotype Threat

Stereotype threat is defined as the predicament felt by people in situations where they could conform to negative stereotypes associated with their own group membership (Steele, 1997). The result of this threat is that individuals may underperform on a task associated with the threatened domain, so women may underperform on a math test or African Americans may underperform on an intelligence test. Ultimately the threat could lead to people removing themselves from the domain altogether, so women may no longer consider math as important to their self-perceptions, or African Americans may leave school, and therefore no longer identify with the domain in which they are stereotyped as inferior to White Americans. Steele and Aronson (1995) examined the stereotype of African Americans and intelligence in the first research which highlighted the consequences of stereotype threat. Not only did they find that African Americans underperformed on a test when they were told it was indicative of intelligence, but they also found that simply asking African Americans to state their race before taking a test reduced the students' subsequent performance.

Since this first study, stereotype threat has become an expansive area, with research conducted in an ever increasing number of domains. Stereotype threat is an interesting phenomenon because of its continued ability to present itself in different settings; crossing race, ethnicity, gender and culture. As well as African Americans and intelligence (Steele & Aronson, 1995; Aronson, Fried, & Good, 2002) and women and math (Spencer, Steele, & Quinn, 1999; Marx & Roman, 2002; Nosek, Banaji, & Greenwald, 2002; Schmader, 2002), the threat has been examined with regard to women and career choices (Davies,

Spencer, Quinn, & Gerhardstein, 2002), White men vs. Asian men and math (Aronson, Lustina, Good, Keough, Steele, & Brown, 1999; Smith & White, 2002), White men and sport (Stone, Lynch, Sjomeling, & Darley, 1999; Stone, 2002), homosexual men and childcare (Bosson, Haymovitz, & Pinel, 2004) as well as social class (Croizet & Claire, 1998). The research has shown that all that is really needed to produce stereotype threat is to be placed in a situation where the stereotype is salient. Indeed it is defined as a situational threat or a "threat in the air" (Steele, 1997, p. 613). To come under stereotype threat people must find themselves in a situation that readily leads itself to the stereotype, so for instance, women must be given a math test, or White men must be placed in a sporting environment.

Previous research has examined possible mediating processes underlying stereotype threat, although no one factor has been shown to be exclusive. One potential explanation for the effect that has been extensively investigated is anxiety (Blascovich, Spencer, Quinn, & Steele, 2001; Osborne, 2001; Spencer et al., 1999; Steele, 1997). However, a lack of overwhelming support for anxiety's mediating role has led to the exploration of other potential candidates such as performance expectancy (Cadinu, Maass, Frigerio, Impagliazza, & Latinotti, 2003; Stangor, Carr, & Kiang, 1998; Sekaquaptewa & Thompson, 2002), dejection (Keller & Dauenheimer, 2003), working memory capacity (Schmader & Johns, 2003) and self-handicapping (Keller, 2002; Stone, 2002). All of which appear to contribute, in part, to the explanation for the effect.

<div align="center">The moderation of stereotype threat</div>

As well as the mediators of stereotype threat, focus has been placed on examining the *moderators* of stereotype threat. That is, aspects of the stereotype threat process that affect the level of performance decrement that is experienced. The literature appears to fall into three different areas. The first area includes studies which moderate stereotype threat by moderating the emotional responses to the stereotype. For example, Aronson et al. (2002) and Good, Aronson, and Inzlicht (2003) found that targeting the negative stereotype associated with task performance could reduce stereotype threat. Aronson et al. (2002) focused on the stereotype concerning African Americans in education, and informed participants (over a substantial period of time) that intelligence was malleable and not fixed. By

moderating the emotional response to the stereotype surrounding intelligence the stereotype threat effect appeared to abate. Similarly, Marx and Roman (2002) found that positive role models in the relevant domains could also moderate stereotype threat.

A second line of work has tried to moderate stereotype threat by changing perceptions of the situation. One approach is to inform participants that the stereotype does not apply in the current context (Spencer et al., 1999; Ouwerkerk, de Gilder, & de Vries, 2000). The composition of the group in relation to ingroup and outgroup members has also been shown to affect stereotype threat (Inzlicht & Ben-Zeev, 2000), as has whether the participant is solo in status (Roberson, Deitch, Brief, & Block, 2003; Sekaquaptewa & Thompson, 2002).

The third set of studies have examined the moderating role of individual differences. The level of identification that the individual has with the domain concerned has been found to affect stereotype threat (Aronson et al., 1999; Leyens, Désert, Croizet, & Darcis, 2000; Pronin, Steele, & Ross, 2004) as has identification with the group (Nosek et al., 2002; Schmader, 2002), levels of stigma-consciousness (Brown & Pinel, 2003) and even testosterone levels (Josephs, Newman, Brown, & Beer, 2003).

In our research we aimed to build on these previous findings and test a potential intervention strategy for reducing stereotype threat effects derived from work on reducing intergroup bias. In particular, we hypothesized that the positive effects of blurring intergroup boundaries on prejudicial attitudes and behavior could be extended to the stereotype threat domain.

Categorization models of reducing intergroup bias

The extent to which ingroups and outgroups are differentiated and distinct is a key determinant of how such groups are evaluated (Brewer, 1991; Tajfel & Turner, 1979; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). Indeed, merely distinguishing between people on the basis of their group affiliations is sufficient to observe ingroup favoritism (Tajfel, Billig, Bundy, & Flament, 1971). All other things being equal, the knowledge that *they* are different from *us* translates into evaluative differentiation (see Brewer, 1979; Mullen, Brown, & Smith, 1992). Models have been proposed to explain this impact of categorization on intergroup attitudes such as Doise's (1978) *category differentiation model* (CDM) and in the

form of the meta-contrast process (Oakes, Haslam, & Turner, 1994) outlined by *self-categorization theory* (SCT; Turner et al., 1987). While the emphasis of these accounts varies, what is common to all is the notion that categorization provides a psychological basis for understanding *them* to be different from *us*, and it is the emergence of this distinction between ingroups and outgroups that provides the pre-requisite for intergroup discrimination.

It is the notion that differentiation is positively related to intergroup bias that has formed the basis for multiple models of bias-reduction. Models of contact (Brewer & Miller, 1984; Hewstone & Brown, 1986; Miller, Brewer, & Edwards, 1985; Pettigrew, 1998), the formation of a *common ingroup identity* (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) all incorporate the idea that as differentiation is reduced, and groups are perceived to possess overlapping characteristics, intergroup evaluations will also become less differentiated (and bias will be reduced). This contention is well supported using a variety of paradigms all of which essentially promote the perception of intergroup overlap (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, and Hewstone, 2001).

A categorization model of reducing stereotype threat

The differentiation-reducing models outlined above have been highly successful in helping to reduce the extent to which evaluations of groups are ingroup favoring. Our aim was to test whether such interventions that promote intergroup overlap might also be successful at reducing stereotype threat. There are some good theoretical reasons why this might be the case. One can assume that awareness of relevant ingroup/outgroup differences are a precondition for observing stereotype threat effects (one can not assimilate to one's perceived ingroup inferiorities if one does not have a perception of the ingroup being a distinct entity from an outgroup comparison group). For instance, when taking an intelligence test stereotype threat effects will be observed when participants are aware of their ethnicity or class status, and when taking a math test the effects will be observed when participants are aware of

being female. Correspondingly, African Americans may compare themselves to White Americans and women may compare themselves to men. Categorization models of bias-reduction do so by encouraging a weakening of the *us* versus *them* prerequisite distinction. One cannot positively favor *us* over *them* if *we* are not perceived as psychologically different from *them*. Similarly, here we predict that a task that encourages the weakening of the *us* versus *them* distinction will reduce stereotype threat: One cannot conform to a stereotype based on expected performance differences between *us* and *them* if *we* are not perceived as psychologically different from *them*.

We focus on the well-studied stereotype threat area of women and math (Spencer et al., 1999; Marx & Roman, 2002; Nosek et al., 2002; Schmader, 2002) and women and career choices (Davies et al., 2002). Based on the model outlined above, to reduce stereotype threat in these domains we must encourage the perceived weakening of the intergroup boundary. As the salience of the gender categorical distinction is weakened, so too should be the salience of the associated stereotypes, and as a result the stereotype threat effect should be attenuated.

Experiment 1

Experiment 1 was designed to test of the potential positive effects of encouraging intergroup overlap within the stereotype threat domain of women and career choices. Previous work by Davies et al. (2002) investigated the career choices made by women after seeing either stereotypical television commercials or non-stereotypical commercials. They found that women were more likely to show an interest in stereotypically female careers after seeing commercials where women were portrayed in stereotypical roles (a stereotype threat effect), compared to women viewing neutral commercials and men in both conditions. Davies et al. found this assimilation effect only for females, not for males, ruling out the possibility that it was a simple ideomotor effect, rather than specifically threat. Here, we hypothesized that a similar orientation to female stereotypic careers for females in a baseline condition would be attenuated if participants first carried out a task designed to blur the intergroup boundary. Specifically, we predicted that encouraging intergroup overlap would lead our female participants, relative to baseline, making less stereotypically female career choices. We also hypothesized that the strength of the

stereotypicality of the career may also have an effect, More specifically, we thought it plausible that stronger stereotypic representations might be more resistant to change following our overlapping characteristics task. We expected careers that were highly stereotypical of females or males to be less open to moderation than those that were less stereotypically associated with gender. As such, we carried out a pre-test (described below) to identify careers of varying stereotypicality, and entered this factor into the analysis in our main experiment.

<div align="center">Method</div>

Participants and design

Thirty female students at the University of Birmingham (mean age = 21, SD = 2.95) were randomly allocated to one of two conditions (task: baseline vs. overlap). Participants received £1 ($1.60) for their participation.

Pre-test

The careers survey consisted of a list of eight careers, four of which had been established as stereotypically female (primary school teacher, i.e., who teach children up to the age of 11; physical therapist; registered nurse; social worker), and four as stereotypically male (dentist; accountant; military officer; mechanical engineer). The stereotypicality of these careers had been established in a pre-test which had asked 60 participants (30 male, 30 female) to judge 55 professional careers as to their typical association with males and females. Each career was judged on an eleven point scale from -5 (very atypical) to +5 (very typical), with 0 indicating neither typical nor atypical. The gender order of the careers was counterbalanced with participants either presented with a career label and asked to rate female typicality followed by male typicality, or male typicality followed by female typicality. Each of the 55 occupations were then analyzed using a 2 (participant gender: female vs. male) x 2 (order: female first vs. male first) x 2 (career rating: female vs. male) mixed ANOVA with repeated measures on the third factor. This analysis extracted careers that were perceived as stereotypically male or female, regardless of the order the careers were presented (i.e. female/male first) and regardless of the participant's own gender (all career rating main effects, $p < .0005$).

The eight careers chosen were also either *strongly* or *weakly* stereotypical on the basis of one-sample *t*-tests, test value = 0. Careers *strongly* associated with males were defined as those which were significantly stereotypical of males (p < .05), and at the same time significantly counter-stereotypical of females (p < .05) (mechanical engineer; military officer). Careers assigned a *weak* label were significantly stereotypical of males (p < .05), but *not* significantly counter-stereotypical of females (e.g., dentist; accountant). Using the same criteria we selected two strong female careers (registered nurse; primary school teacher) and two weak female careers (physical therapist; social worker).

Procedure

In order to encourage the perception of intergroup overlap (reduced differentiation) we used a task previously used successfully in the bias-reduction literature. Categorization models of bias-reduction are linked by the general notion that they reduce ingroup/outgroup differentiation and increase the perceived overlap between the ingroup and outgroup's characteristics. Crisp and Beck (2005) developed a task influenced by this basic idea. They found that simply asking participants to list characteristics shared between the ingroup and outgroup could reduce intergroup bias. This task also has effects on intergroup differentiation consistent with the above mentioned categorization models of bias-reduction (Cocker, 2004) Here we used this method of creating category overlap.

The overlap condition consisted of a task which asked participants, "to think of five things that men and women can have in common (i.e. characteristics that men and women share)". Below this were five numbered spaces for participants to write the five characteristics. This task can also be seen as conceptually similar to Mussweiler (2001) who primed similarity/dissimilarity through the use of two pictures. Mussweiler asked participants to list as many similarities or differences as they could between the two scenes. Completion of the task was found to prime an orientation to look for similarities or differences. Our task asks participants to generate shared characteristics and is therefore similar to Mussweiler's manipulation insofar as it also leads participants to focus on similarities or differences between, in our case, gender categories After completing the task, participants moved on to the careers survey. Participants in the baseline condition did not receive the overlapping characteristics task, instead

moving straight on to the careers survey. On completion of the careers survey participants in both

conditions were asked to complete a feedback form, before being thanked and debriefed.

Dependent Measures

In the careers survey the eight careers developed in the pre-test were presented to the

participants. Participants read the instructions:

We would like you to think about the career path that you would be most interested in if

you had not embarked on your chosen path. That is, please imagine that you have not yet

entered university, and you have many different options available to you. Which of the

following, using the scales below, would you find appealing?

Participants then rated the eight careers in a random order on a nine point scale, 1 = *not at all*

interested to 9 = *very interested.*

Results and discussion

Career choices

Career preference was calculated as the mean rating for each of the career types (strength: weak

male, weak female, strong male, strong female). A 2 (task: baseline vs. overlap) x 2 (career type: male vs.

female) x 2 (career strength: weak vs. strong) ANOVA with repeated measures on the last two factors

revealed a significant main effect of career type, $F(1, 28) = 65.07, p < .001, \eta_p^2 = .70$. Overall female

careers were preferred ($M = 4.61$) over male careers ($M = 2.49$). There was no main effect of task, $F(1,$

$28) = .292, p = .593$, and no interaction effects between task and strength, $F(1, 28) = 1.27, p = .269$, nor

career type x strength, $F(1, 28) = .073, p = .739$. The career type x task interaction was, however,

significant, $F(1, 28) = 7.59, p = .010, \eta_p^2 = .21$, and not qualified by strength (the three-way interaction

was not significant, $F(1,28) = .044, p = .836$), see Table 1.

Planned t-tests revealed that in the baseline condition the female participants rated stereotypically

female careers as more preferable ($M = 5.13$) compared to male careers ($M = 42.29$), $t(12) = 8.11, p <$

$.001$. This difference remained significant for participants in the overlap condition; $t(16) = 3.74, p = .002$

(Ms = 4.09 and 2.69 respectively), although the extent to which the female careers were preferred over

male careers was attenuated ($M_{Baseline}$ = 2.85, $M_{overlap}$ = 1.40). There was a trend towards there being lower preference for female careers after the overlapping task (M = 4.09) compared to baseline (M = 5.13), t (28) = 1.49, p = .148, but little evidence for moderation of preference for stereotypically male careers, t (28) = -.678, p = .503.

The results offer some initial support for the idea that encouraging a focus away from intergroup differences can alleviate stereotype threat. Although tentative, the findings suggest that female participants who carry out a task designed to reduce the salience of gender categorization subsequently show a lower preference for female stereotypical careers relative to male stereotypical careers. Notably, we did not find an effect of career strength. This suggests that the shared characteristics task is effective on preference for both weak and strong careers. This is an important finding in terms of applicability, because it suggests that if this characteristics task can be used as a basis for a real stereotype threat intervention (for example, within education), then the intervention would be effective in reducing the prevalence of threat irrespective of the level of perceived stereotypicality.

The results of Experiment 1 were encouraging. However, there are some important issues that were not addressed. The aim of the experiment was simply to ascertain whether there was any support for the basic hypothesis that thinking about shared characteristics could reduce stereotype threat. As such, the baseline condition used was simply the absence of any task. There is therefore the possibility that rather than thinking about overlapping characteristics it could be that it was the cognitive activity of simply doing *any* task that reduced the stereotype threat effect in this experiment. Therefore in Experiment 2 we included a third condition to address this issue. In this third condition participants were required to list *differences* between the genders (instead of similarities). While this offers the equivalent cognitive load as thinking about similarities, the task itself should not blur intergroup boundaries. Experiment 2 also looked to expand the generality of the task by attempting to replicate the findings in the domain of stereotype threat relating to women and math performance.

Experiment 2

In Experiment 1 participants in the overlap condition tended to show less preference for stereotypical female careers than participants in the baseline condition. In Experiment 2 we expected to find similar results with respect to women and math performance. It was hypothesized that participants in the overlap condition would answer more questions correctly than those in the baseline condition, following the same pattern as Experiment 1. In order to test the hypothesis that it was specifically thinking about intergroup overlap, and not simply thinking about *anything* that would have the positive effects, we included a third condition. Here we asked participants to list five *differences* between men and women. We expected thinking about differences to have no beneficial effects with respect to reducing stereotype threat. In fact, it was possible that the introduction of a task emphasizing differences between the genders could *accentuate* any threat inherent to the situation, resulting in these participants coming under more threat than those in the baseline condition (i.e., the math performance detriment becoming even more pronounced). However, we only tentatively make this hypothesis for the following reason. From Experiment 1 it is clear that even with no explicit threat manipulation, a bias in favor of female stereotypic careers was observed. It may therefore be that there is a ceiling effect in that when a threat is already present (as it appears to be in the baseline condition of Experiment 1), it is less likely that further demands (i.e. thinking about differences between the groups) can make that threat worse.

Previous studies (Spencer et al., 1999; Quinn & Spencer, 2001) have argued that the difficulty of math tasks is an essential element in stereotype threat effects (and that stereotype threat only occurs when math questions are of a difficult level). Therefore in Experiment 2 we presented math questions which were aimed at the ability of the participants (psychology undergraduates), while still offering a challenge. This reasoning was in line with Spencer et al. (1999, exp.3) who adjusted the difficulty of the math presented when testing psychology students, who tend to be of lower mathematical ability than students taking more math oriented degree majors.

<div align="center">Method</div>

Participant and design

Thirty-six female psychology students (mean age = 21, SD = 5.25) were randomly assigned to one of the three conditions (task: baseline vs. overlap vs. difference). Participants received £3 ($4.70) or course credit for their participation.

Procedure

The participants were tested on an individual basis and were shown to a lab by a female experimenter where they were informed that the study consisted of a selection of pre-tests. The first task consisted of either the overlapping characteristics or the different characteristics task. The overlapping characteristics task was the same as in Experiment 1, while the different characteristics task asked participants, "to think of five things that can distinguish men from women (i.e. characteristics that men and women do not share)". Numbered spaces then followed. Participants in the baseline condition did not receive such a task, moving straight on to the second task, which consisted of ten math questions. For the math test participants were given a plain piece of paper for calculating the answers, and were not allowed to use a calculator. The participants were informed that they had five minutes to complete as many of the questions as they could, but to let the experimenter know if they completed them sooner. After completing the math questions, participants were asked to provide feedback before being thanked and debriefed.

Dependent Measures

The math questions were designed to be straightforward, and constituted tests of mental arithmetic, e.g. "In hall accommodation, 23 out of 25 rooms were occupied. What was the number of occupied rooms as a percentage?". There were ten questions in total, and the participants were given five minutes to complete the task.

<div align="center">Results and discussion</div>

Math task

Only one participant, who was in the overlap condition, attempted all ten math items. A one-way ANOVA was computed on the number of correct math scores, revealing significant variation as a function of condition, $F(2, 33) = 3.79$, $p = .033$, $\eta_p^2 = .19$, see Table 2. Planned contrasts revealed no

significant difference between the baseline and difference conditions (baseline -1, overlap 0, difference +1), $t(33) = .844$, $p = .405$. Math performance was, however, better in the overlap condition ($M = 6.83$) compared to either the baseline ($M = 4.80$) or difference ($M = 5.43$) conditions (baseline -1, overlap +2, difference -1), $t(33) = 2.69$, $p = .011$.

Content analysis

In order to examine whether the exact nature of the shared and non-shared characteristics had any impact on the effects observed, two independent coders were employed to rate each of the responses given by the participants during the generation task. The coders were asked to rate each characteristic generated as being either *academic* (i.e. relating to education), non-academic and *physical* (i.e. actual physical similarities/differences between the two genders) or *other* (characteristics that were neither academic or physical). The raters' responses correlated well for all threes types of characteristic: $r(26) = .953$, ; $r = .880$; $r = .840$ respectively, all p's $< .0005$. We examined whether the extent to which each of the different types of characteristics were generated had any impact on the threat effects. We found no correlation between the number of academic, non-academic [physical], non-academic [non-physical] or non-academic [physical and non-physical combined] with math performance. This finding occurred when the correlations were carried out within the similar and different conditions respectively, and combined. Notably, most participants generated non-academic [non-physical] characteristics. Few physical or academic characteristics were generated.

Experiment 2 replicated the findings of Experiment 1, within the stereotype threat setting of women and math performance. The inclusion of the differences condition supports the notion that it is specifically emphasizing overlapping characteristics that results in the reduction of stereotype threat, ruling out an alternative explanation in terms of cognitive load. No significant difference was observed between the baseline and the difference condition. This suggest that there may be a ceiling effect, in that when a threat is already present (as it appears to be in our baseline condition) it is less likely that further demands that should instigate threat (i.e. thinking about differences) make the threat worse.

Although we may regard the difference condition used here as a kind of threat manipulation, we have nevertheless not tested the effect an explicit threat manipulation (i.e., stating that men and women will be compared on the test) will have on math performance, when combined with our characteristics task. While some studies (Smith & White, 2002; Spencer et al., 1999; McIntyre, Paulson, & Lord, 2003) have established that simply being in a relevant situation, i.e. being given a math test, is sufficient to elicit stereotype threat, other studies (Aronson et al., 1999; Brown & Pinel, 2003) have reminded participants of the stereotype threat associated with the situation they are in before completing the task. To expand the generality of our findings we therefore included a specific threat manipulation in Experiment 3. If blurring intergroup boundaries can reduce stereotype threat when an intergroup comparison is explicit, this will strengthen support for the potential benefits of this as an intervention strategy.

Experiment 3

Experiment 3 draws on the previous literature which explicitly induces stereotype threat. To examine the effects of overlapping characteristics in relation to explicit threat, the experiment included four separate conditions. The first condition was the same baseline condition as in Experiments 1 and 2. In the second condition participants were told that their results would be compared to men (specific threat condition). It was predicted that there would be no significant difference between the baseline and threat condition, as from previous work we know that simply placing participants in a math test situation can induce stereotype threat. However, as discussed above, it is also possible that some conditions, such as explicitly making the distinction between groups salient, could lead to a greater experience of threat. We therefore tentatively hypothesized a decrease in performance in the threat condition compared to the baseline condition. The third condition introduced the overlap task followed by the threat manipulation. In the fourth condition the threat manipulation was *followed* by the overlap task. The reversed ordering of the tasks in the third and fourth conditions was included to test the hypothesis that the overlapping characteristics task would work best to alleviate stereotype threat if it was presented *before* the threat manipulation.

We can explain our rationale for this prediction by conceiving a model of stereotype threat that involves three stages: categorization at stage one is a pre-requisite for instigating processes at stage two that lead to behavioral effects (e.g., detriments on math performance) at the final stage. Intervention at stage one (categorization) should prevent movement into stage two (the processes that lead to detrimental math performance). Therefore, carrying out the overlap task before the threat (i.e. intervention at stage one) should reduce the observation of stereotype threat effects. However, if threat is induced before the characteristics task (i.e. intervention at stage two), then the processes leading to detrimental math performance will already have been reached. Therefore, when the overlap task comes after the threat (albeit before the math test) we still expect to observe stereotype threat effects. We therefore tested the hypothesis that the overlapping characteristics task will work most effectively if it is used prior to the instigation of threat. If this is the case, then it would suggest that the overlapping task works effectively not by alleviating the stereotype threat, but by preventing the stereotype threat from emerging at all.

<div align="center">Method</div>

Participant and design

Sixty-two white female psychology students (mean age = 19, SD = 1.37) were randomly allocated to one of four conditions (baseline vs. threat vs. overlap-threat vs. threat-overlap). Participants received course credit for their participation.

Procedure

The participants were shown to a lab where they were tested individually by a female experimenter. Like Experiment 2, all participants were informed that they would be pretesting a selection of tasks, even though participants in the baseline and threat conditions only received one task. Participants in the baseline condition were given the same instructions as in Experiment 2. The threat condition consisted of the same task and procedure as the baseline condition with one crucial difference: before receiving the math test participants were given the threat manipulation, i.e. they were informed that the test was being administered because the experimenter was "interested in comparing the math

performance of men and women, to see if there is a difference between the two genders". In the overlap-threat condition, participants were asked to complete the overlapping characteristics task, after which they were given the threat manipulation, followed by the math test. In the threat-overlap condition, the participants were informed that the second task they would be receiving was a selection of math questions and they were given the threat manipulation. Before receiving the math questions they were given the overlapping characteristics task. They were *not* reminded of the threat immediately before being given the math task. On completion of the math test all participants completed a feedback questionnaire before being thanked and debriefed.

Dependent Measures

The same math questions were used as in Experiment 2. However, a question was added to the feedback form to indicate whether the participant could recall the reason for the math test.

<div align="center">Results and discussion</div>

Math task

Because the addition of the threat manipulation was the only difference between the baseline and threat conditions, we included a manipulation check to assess whether participants in the threat condition could remember the reason for the math test (i.e. to examine if there was a difference in performance between men and women). In order for the conditions to be perceived as distinct, seven participants who could not recall the reason were omitted from the analysis.

A one-way ANOVA was carried out on the number of math questions answered correctly, revealing a significant effect of task, $F(3, 51) = 3.07$, $p = .036$, $\eta_p^2 = .15$, see Table 3. We used a set of Helmert contrasts to incrementally test whether there was support for our hypotheses. First, we tested whether there was any difference between the baseline and threat conditions (baseline -1, threat +1, overlap-threat 0, threat-overlap, 0), this analysis revealed no significant difference, $t(51) = 1.32$, $p = .192$. Given no difference between these two conditions, we then compared their aggregate with the threat-overlap condition (baseline -1, threat -1, overlap-threat 0, threat-overlap +2). This analysis also revealed no differences, $t(51) = -.448$, $p = .656$, Finally, given no differences between the baseline, threat and

threat-overlap conditions, we compared all three conditions with the overlap-threat condition baseline -1, threat -1, overlap-threat +3, threat-overlap -1). This analysis revealed that math performance was significantly better when the overlap task preceded the threat compared to all other conditions, t (51) = 2.71, p = .009. These findings support the idea that compared to the baseline and threat conditions (Ms = 5.21 and 6.20), completing a task that blurs gender boundaries reduces stereotype threat and improves math performance (M = 7.58), although this is not the case if the overlap task is completed after the threat has been induced (M = 6.00).

As in Experiment 2, two independent coders rated the characteristics generated as academic, physical (and non-academic) or other (non-academic and non-physical). The raters' responses correlated well for all threes types of characteristic: r (26) = .923, ; r = .858; r = .845 respectively, all p's < .0005. As in Experiment 2, however, there was no correlation between the number of different characteristics generated and math performance, across or within categorization conditions.

Experiment 3 furthered supported the notion that emphasizing overlapping characteristics can be beneficial for reducing stereotype threat by including a specific threat manipulation. There was no significant difference between the baseline and threat condition suggesting that the threat was implicitly activated in Experiments 1 and 2. As expected, when participants received the threat manipulation after first having completed the overlapping characteristics task, math performance improved. However, participants who received the threat manipulation *before* the overlapping characteristics measure did not improve their performance on the math test. This finding has important implications for the effectiveness of the overlapping characteristics task as an intervention technique. As the overlapping task was most effective in reducing stereotype threat when it was presented before the threat manipulation, this suggests that the task may prevent the stereotype threat from emerging at all, as apposed to it simply alleviating the threat. It is in line with our predictions that the task is more effective before the instigation of threat rather than after. If categorization instigates processes that lead to detriments on math performance, intervention at this stage should prevent such processes occurring (and so prevent them leading to stereotype threat effects). If it is the case, however, that the threat has already been perceived,

and the processes set in motion that lead to performance detriments, then encouraging the perception of blurred boundaries will have little effect. It is the processes behind stereotype threat that lead to performance detriment, *not* the stereotypes themselves. Therefore, so working against the stereotype (i.e. emphasizing similarities between females and males) *after* it has set these processes in motion will have little positive effect.

General discussion

The experiments presented here tested whether the application of an overlapping characteristics task derived from bias-reduction literature would have beneficial effects on stereotype threat. Experiment 1 established the usefulness of the task for reducing bias in career choices. Female participants who generated overlapping characteristics between males and females subsequently showed a lower preference for stereotypically female careers relative to stereotypically male careers, compared to a baseline condition. The benefit of emphasizing overlapping characteristics was reinforced in Experiment 2 within a second domain, the highly examined stereotype threat arena of women and math performance. Participants who completed the overlapping characteristics task completed more questions correctly than participants in both the baseline and difference conditions. Finally, in Experiment 3 we explicitly activated threat, by informing the participants that their results would be compared to men, yet those who received the overlapping characteristics task, before receiving the threat, still completed significantly more questions correctly than either a threat alone or baseline condition. We discuss the theoretical and practical implications of these findings below.

Theoretical Implications

We observed a reduction of stereotypical career preference and greater performance on math tests for participants who carried out the overlapping characteristics task. The task appeared to be most successful when it was presented *before* an explicit threat, which suggests that the task's use may lie in preventing stereotype threat from emerging in the first place, rather than alleviating threat once it is instigated. To our knowledge, this is the first intervention in the stereotype threat literature that focuses on changing the perceived distinction between relevant social categories.

The overlapping characteristics task is derived from models developed to encourage reductions in prejudice, discrimination, and bias between groups. This indicates that the benefits of such interventions used to reduce prejudicial attitude, which focus on reducing the salience of intergroup boundaries, can be generalized to the domain of stereotype threat. In turn, this suggests that categorization, and specifically the salience of the categories involved, is central to the emergence of stereotype threat. This finding supports previous stereotype threat research which has noted the importance of the salience of the stereotype (Keller, 2002). As such, the findings help support the meta-theoretical centrality of category salience as a key predictor of attitudes and behavior. Importantly, as in the bias-reduction domain, this does not rule out other potential explanations. Rather, the categorization model outlines a pre-requisite role for observing biased attitudinal (prejudice) or behavioral (stereotype threat) effects. Quite simply biases resulting from thinking categorically depend upon category salience. The findings here support the notion that anything that weakens the salience of distinctive social categorization can correspondingly weaken category-related effects. Stereotype threat will be less likely to occur if the categories that embody the stereotype become less salient. From a wider perspective, these findings help to integrate work in the stereotype threat domain with work on prejudice and discrimination, providing a theoretical link between attitudinal and behavioral assimilation literature.

With respect to future work, while the overlapping characteristics task used here appears to have been successful, it would be useful to test whether other similar models for reducing categorical differentiation (and bias) can also be applied to the stereotype threat domain. These might include using the *crossed categorization* approach (where a single overlapping category is made salient) or the *common ingroup identity model* (where a single overlapping superordinate category is made salient). Until then, the current findings lend strong support for the basic idea that blurring intergroup boundaries may be a beneficial tool for reducing stereotype threat.

Practical Implications

Experiment 3 established that participants in our baseline conditions experienced threat effects to the same level as when we explicitly induced threat. This supports previous studies which have

concluded that stereotype threat is inherent in relevant situations, i.e. simply being in a situation associated with a negative stereotype can result in stereotype threat. This, of course, reinforces the importance of developing interventions designed to alleviate such negative behavioral tendencies. Also, it has previously been stated that math tasks need to be *difficult* to induce stereotype threat, however, in comparison to questions on the Scholastic Aptitude Test (SAT) and Graduate Record Examination (GRE), which many stereotype threat studies use, the questions the participants received here were relatively easy. However, the participants had a limited amount of time to complete the questions, which may have added urgency and pressure to the task, thus making it more difficult. Likewise, there were no multiple choice responses which should have added a level of difficulty, by reducing the possibility of correctly guessing. As such, the questions used appear to be of sufficient difficulty for the participants, which is illustrated by the absence of any notable ceiling effects. However, even if one could argue that the questions were of a relatively easy level, if such biases can occur at such a level this again reinforces the need to develop means to attenuate stereotype threat effects. The overlapping characteristics task employed here, and by extension other existing models for reducing categorical differentiation, offer such a means for reducing a stereotype threat. The reported studies are a first step, and future work will need to refine such methods and extent them beyond the laboratory to test their applicability to policy and practice.

Conclusion

In this research we aimed to examine the applicability of categorization models for reducing prejudice in the stereotype threat domain. The results of the three experiments presented here offer support for the notion that interventions used for reducing bias can be applied to stereotype threat research with success. The experiments found that participants who focused on characteristics which overlapped between the genders, compared to those who did not, completed more math questions correctly and preferred significantly less stereotypical careers. These findings support the notion that such categorization models can be generalized to the stereotype threat domain and that blurring intergroup boundaries can offer the potential for reducing stereotype threat.

References

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 math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*, 29-46.

Blascovich, J., Spencer, S.J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science, 12*, 225-229.

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.

Brewer, M.B. (1979). Ingroup bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86*, 307-324.

Brewer, M.B. (1991). The social self – on being the same and different at the same time. *Personality and Social Psychology Bulletin, 17*, 475-482.

Brewer, M.B., & Miller, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller & M.B. Brewer (Eds.), *Groups in contact: The psychology of depression* (pp. 281-302). Orlando, FL: Academic Press.

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.

Cocker, R.M. (2004). Social identification and shared characteristics: Routes to bias-reduction. Unpublished MPhil dissertation.

Crisp, R., & 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.

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 (2004). *GCSE/GNVQ and Key Stage 3 to GCSE/GNVQ

value added measures for young people in England, 2002/03 (final)* (SFR 24/2004). London, UK: Department

for Education and Skills.

Department for Education and Skills (2004). *GCE/VCE A/AS examination results for young people

in England, 2002/03 (final)* (SFR 24/2004). 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.

Equal Opportunities Commission (2004). *Facts about women and men in Great Britain.* Manchester,

UK: Equal Opportunities Commission.

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.

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.

Hewstone, M., & Brown, R. (1986). Contact is not enough: An intergroup perspective on the 'contact hypothesis'. In M. Hewstone & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 1-44). Oxford: Blackwell.

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. (2002). Blatant stereotype threat and women's math performance: Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles, 47*, 193-198.

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.

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.

Miller, N., Brewer, M.B., & Edwards, K. (1985). Cooperation interaction in desegregated settings: A laboratory analog. *Journal of Social Issues, 41*, 63-79.

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An interaction. *European Journal of Social Psychology, 22*, 103-122.

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.

Mussweiler, T. (2001). 'Seek and ye shall find': Antecedents of assimilation and contrast in social comparison. *European Journal of Social Psychology*, *31*, 499-509.

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.

Oakes, P.J., Haslam, S.A., & Turner, J.C. (1994) *Stereotyping and social reality.* Oxford: Blackwell.

Osborne, J.W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology, 26*, 291-310.

Ouwerkerk, J.W., de Gilder, D., & de Vries, N.K. (2000). When the going gets tough, the tough get going: Social identification and individual effort in intergroup competition. *Personality and Social Psychology Bulletin, 26*, 1550-1559.

Pettigrew, T.F. (1998). Intergroup contact theory. *Annual Review of psychology, 49*, 65-85.

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.

Quinn, D.M., & Spencer, S.J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues, 57*, 55-71.

Roberson, L., Deitch, E.A., Brief, A.P., & Block, C.J. (2003). Stereotype threat and feedback seeking in the workplace. *Journal of Vocational Behavior, 62*, 176-188.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38*, 194-201.

Schmader, T., & Johns, M. (2003) Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*, 440-452.

Sekaquaptewa, D., & Thompson, M. (2002). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology, 39*, 68-74.

Smith, J.L., & White, P.H. (2002) An examination of implicitly activated, explicitly activated, and nullified stereotypes on mathematical performance: It's not just a woman's issue. *Sex Roles,47,* 179-191.

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.

Stangor, C., Carr, C., & Kiang, L. (1998). Activating stereotypes undermines task performance expectations. *Journal of Personality and Social Psychology, 75*, 1191-1197.

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. (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.

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.

Tajfel, H., Billig, M.G., Bundy, R.P., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology, 1*, 149-178.

Tajfel, H., & Turner, J.C. (1979). An integrative theory of intergroup conflict. In W.G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks/Cole.

Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D., & Wetherell, M.S. (1987). *Re-discovering the social groups: 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.

Author note

Table 1. *Career preference as a function of task, Experiment 1.*

	Female stereotypical careers		Male stereotypical careers		Relative preference for female stereotypical careers	
	M	*SD*	*M*	*SD*	*M*	*SD*
Baseline	5.13	1.34	2.29	1.24	2.85	1.26
Overlap	4.09	2.24	2.69	1.84	1.40	1.54

Table 2.*Number of math questions answered correctly, as a function of task, Experiment 2.*

	Number of correct math questions	
	M	*SD*
Baseline	4.80	1.55
Overlap	6.83	1.80
Difference	5.43	1.95

Table 3.*Number of math questions answered correctly, as a function of task, Experiment 3.*

	Number of correct math questions	
	M	*SD*
Baseline	5.21	2.22
Threat	6.20	1.90
Overlap-threat	7.58	1.78
Threat-overlap	6.00	2.08