Group Processes & Intergroup Relations
G P I R

Article

Group Processes & Intergroup Relations
1–15

DOI: 10.1177/1368430218811250
journals.sagepub.com/home/gpi
SAGE

Diversity may help the uninterested: Evidence that exposure to counter-stereotypes promotes cognitive reflection for people low (but not high) in need for cognition

Ekaterina Damer,[1] Thomas L. Webb[1] and Richard J. Crisp[2]

Abstract

Previous theorizing and research has linked exposure to counter-stereotypical diversity (e.g., an Oxford-educated bricklayer) to enhanced cognitive performance and creativity. However, it is unclear whether people's *motivation* to cognitively engage with the counter-stereotypical information (i.e., need for cognition [NFC]) influences this effect. Across three experiments (N = 887) we found consistent support for the idea that exposure to counter-stereotypes (CSTs) promotes cognitive reflection for people low in NFC (d_+ = .34). In contrast, people high in NFC showed decreased cognitive reflection after being exposed to CSTs (d_+ = −.18), although the evidence for the latter effect was weak. These findings suggest that exposure to CSTs can promote cognitive reflection *unless* people have a strong desire to understand and predict outcomes and events, in which case exposure to CSTs may backfire. Taken together, we conclude that motivation to engage in cognitive activity may be an important consideration for research and interventions involving social and cultural diversity.

Keywords

diversity, cognitive reflection, counter-stereotypes, motivation, need for cognition

Paper received 9 November 2017; revised version accepted 15 October 2018.

Introduction

When the music legend Prince passed away, his fans honored him by writing that he "defied narrow stereotypes about race and gender" and thus "opened the minds of others" ("Fans Thank Prince"; New York Times, 2016). Indeed, it is notable that contemporary societies increasingly bring people into contact with complex combinations of social, religious, and cultural identities. These days, it is not uncommon to learn about female chief executives spearheading technology firms, Muslims advocating for liberal values

[1]The University of Sheffield, UK
[2]Durham University, UK

Corresponding author:
Ekaterina Damer, Department of Psychology, The University of Sheffield, Sheffield, UK.
Email: ek.damer@gmail.com

such as abortion rights, disabled people winning sports competitions, Catholics coming out as gay, or successful entrepreneurs who dropped out of university. Such individuals are exemplars of counter-stereotypical diversity: They do not conform to traditional stereotypes and thus are termed counter-stereotypes ("CSTs"; Crisp & Turner, 2011). People encounter counter-stereotypical individuals when they commute to work, attend school, and pursue leisure activities, or when they travel, change jobs, or move to new places. The media, films, and books frequently feature significant achievements of underrepresented minorities, such as the film *Hidden Figures* that tells the story of three Black female engineers who helped send U.S. American rockets into space. What is more, grassroots social media campaigns have been launched in recent years with the goal to dismantle clichés, for example the #ILookLikeAnEngineer campaign started by a female engineer on Twitter. The question, then, is how do people respond to new forms of social and cultural diversity? And to what extent can exposure to social and cultural diversity affect broader cognitive functioning?

The CPAG Model

Crisp and Turner's (2011) categorization–processing–adaptation–generalization (CPAG) model suggests that people try to make sense of CSTs by engaging in "inconsistency resolution" (Hutter & Crisp, 2005). For example, someone who meets an individual that challenges their stereotypes may wonder: "Why is this Muslim a hipster?"; "How did this woman achieve the rank of a CEO?"; or "What made this Harvard-educated man become a carpenter?" By seeking answers to these questions, judgments are less likely to be based on existing stereotypes and more likely on impressions of individuals. According to the CPAG model, at least two conditions are necessary for people to try to resolve apparent inconsistencies: The perceiver needs to (a) be motivated to resolve the inconsistency, and (b) have sufficient cognitive resources to do so (see also Fazio, 1990). If these conditions are met, then the perceiver will seek to resolve the inconsistency by suppressing their existing stereotypical knowledge and reconstruing the target with individualized attributes (e.g., by thinking about the Harvard-educated carpenter as "nonconformist"). Crucially, Crisp and Turner (2011) predict that the process of resolving inconsistencies will stimulate greater cognitive flexibility in the short term, and if repeated over time, in the long term as well. In this area of research, cognitive flexibility is typically defined as the "capacity to 'break set,' go beyond the established and mentally accessible ways of thinking in favor of thinking differently from other people or differently from what is habitual" (e.g., Goclowska & Crisp, 2013, p. 218).

There is some support for the CPAG model. For example, Goclowska, Crisp, and Labuschagne (2012) found that thinking of a gender CST (e.g., a female mechanic) boosted creative performance within a short experimental session. In another line of research, Prati, Vasiljevic, Crisp, and Rubini (2015) showed that thinking of CSTs pertaining to gender (e.g., a female mechanic) decreased dehumanization (i.e., the tendency to consider others as less human than ourselves). Importantly, this change was mediated by a reduced reliance on heuristic thinking, lending support to the model. Finally, research indicates that exposure to CSTs reduces intergroup bias by evoking surprise (Prati, Crisp, & Rubini, 2015), suggesting that affective-motivational states may play a role in the process of resolving inconsistencies following exposure to CSTs. However, the premise that perceivers need to be motivated to engage in *cognitive activity*[1] in order for CSTs to promote cognitive performance has not been tested to date. More precisely, some initial work suggests that personal need for structure (PNS; i.e., preferences for the desired *outcome* of cognitive activity) moderates the effects of exposure to CSTs on cognitive flexibility (e.g., Goclowska, Baas, Crisp, & De Dreu, 2014). In this research, individuals low (vs. high) in PNS showed improved (vs. decreased) cognitive flexibility after exposure to CSTs. However, it is currently unknown whether need for cognition (i.e., preferences for the desired *amount* of cognitive activity) moderates this effect. Need for cognition (or NFC), also known as epistemic/intellectual curiosity (Mussel, 2010), can be seen as an individual

difference variable reflecting the extent to which people desire to understand and predict outcomes or events. This desire manifests itself as "an individual's tendency to engage in and enjoy effortful cognitive endeavors" (Cacioppo, Petty, Feinstein, & Jarvis, 1996, p. 197), and it seems likely that contexts that challenge traditional stereotypes pose a challenge to people's ability to understand and predict outcomes and events.

In the present work, we investigated a close relative of cognitive flexibility—cognitive reflection—which is defined as "the ability or disposition to resist reporting the response that first comes to mind" (Frederick, 2005, p. 36). Given that the CPAG model postulates *inconsistency resolution* to be the critical process that is triggered when people are exposed to CSTs (and after certain necessary conditions are met), it was important to investigate whether individual differences in the desired *amount* (rather than *outcome*) of cognitive activity may play a moderating role in the effect on cognitive reflection. More precisely, because inconsistency resolution itself is a process rather than outcome, it seems likely that people's desire to think about CSTs (i.e., individual differences in NFC) may play a more important role in exposure to CSTs than their desire for certain cognitive outcomes (i.e., individual differences in PNS). The present paper aimed to test this prediction of the CPAG model, that is, the role of individual differences in NFC in the experience of counter-stereotypical diversity.

How Might Need for Cognition Moderate the Effect of Exposure to CSTs on Cognitive Reflection?

There are two plausible—but competing—predictions of how exposure to CSTs might affect intellectually more (vs. less) curious people. First, people *high* in NFC may be more likely to show cognitive flexibility in response to CSTs because they are more motivated to resolve the inconsistencies than people low in NFC, which might make them more likely to expend resources in the face of expectancy-violating experiences (Gocłowska, Damian, & Mor, 2018; Leung & Chiu, 2008). In turn, this could mean that people high in NFC form more cross-cutting explanations for the inconsistent social categories, which may activate more distal cognitive associations and networks (Greenwald et al., 2002) and ultimately enhance cognitive reflection. In other words, people high in NFC should be more willing to resolve the inconsistencies than people low in NFC and, consequently, they should be more likely to switch from a heuristic, category-based mode of processing to a systematic, individuating mode (Evans, 2008; Fiske & Neuberg, 1990; Strack & Deutsch, 2004; Tversky & Kahneman, 1974). In contrast, people low in NFC are by definition less intellectually curious and thus, less likely to be motivated to cognitively resolve CSTs. As a result, they are likely to remain in the heuristic processing mode, both when being exposed to CSTs and in subsequent cognitively challenging tasks. Taken together, one can predict that exposure to CSTs makes people high in NFC switch from heuristic to systematic processing (thus boosting cognitive reflection), whereas exposure to CSTs may not affect people low in NFC.

Alternatively, exposure to CSTs, being surprising and unexpected (Prati, Crisp, et al., 2015), may spark interest and curiosity in individuals low in NFC, which in turn may increase their levels of cognitive reflection. In other words, exposure to CSTs may motivate individuals low in NFC (rather than those high in NFC) to seek to resolve the stereotypical inconsistencies, which in turn might make these individuals switch from a heuristic, category-based mode of processing to a systematic, individuating mode. This idea is consistent with the findings of Allen, Sherman, Conrey, and Stroessner (2009), who found that when people have low processing capacity and stereotypes are strong (e.g., a violent Black person, a warm and friendly woman), they pay more attention to information that is inconsistent with their preexisting stereotypes than to information that is consistent. In contrast, people high in NFC (who already engage in relatively systematic modes of processing by default) might not be sufficiently surprised by CSTs and thus not engage in more cognitive reflection than they already engage in (i.e., a ceiling effect).

The Present Research

The present research involved three experiments that tested the competing predictions described before by exposing participants to different CSTs and subsequently measuring their cognitive reflection. As such, this research was exploratory rather than confirmatory in nature. We developed and validated two paradigms to elicit CST experiences, and measured cognitive reflection using the seven-item Cognitive Reflection Test (CRT; Frederick, 2005; Toplak, West, & Stanovich, 2013). NFC was measured using the 18-item Need for Cognition Scale (Cacioppo, Petty, & Kao, 1984). Sample sizes and participant inclusion criteria were specified a priori and we report all measures, manipulations, and exclusions. The data sets of all three experiments and the *R* code used to run all analyses can be found on the Open Science Framework (https://osf.io/tcfau/).

Experiment 1

Pretest

To manipulate exposure to CSTs, participants were asked to read a short paragraph, which described a CEO (chief executive officer) named David. Participants in the control condition were asked to imagine that they read the following paragraph on the Internet: "David is a CEO. He's also a college graduate (Harvard), born and raised in the US, and happily married to his wife Linda." Participants in the experimental condition were asked to imagine that they read the following paragraph about David: "David is a CEO. He's also a college dropout (Harvard), a Mexican immigrant, and happily married to his husband Michael." We established that the description of David was counter-stereotypical by recruiting 41 U.S. American participants (16 female; M_{age} = 31.51, SD_{age} = 11.53) through the participant recruitment platform Prolific (www.prolific.ac; Peer, Brandimarte, Samat, & Acquisti, 2017) and randomly assigning them to the two conditions described before.

After reading the paragraph about David, participants were asked, "To what extent do you feel surprised?" and "To what extent do you feel astonished?" on a scale from 0 (*not at all*) to 100 (*very much*). Next, to reinforce the manipulation, participants were instructed to imagine what David and his life are like and to describe (in as much detail as possible) their thoughts as to what characteristics he might possess. We checked that this manipulation was successful by asking participants to indicate their agreement with four statements: "David is a typical CEO" (reverse-coded), "Reading about David challenged some of my beliefs," "There isn't anything puzzling about David's life" (reverse-coded), and "Imagining David's life made me think 'outside the box'"; again on a scale from 0 (*strongly disagree*) to 100 (*strongly agree*). The manipulation check was followed by an attention check because it is often difficult to ascertain whether or not participants pay attention to the study materials (Oppenheimer, Meyvis, & Davidenko, 2009; see Appendix A). We created a measure of counter-stereotypicality by calculating the mean of six items (i.e., the two items reflecting surprise and the four items reflecting counter-stereotypicality; $\alpha = .80$). Lastly, participants were asked to indicate their sex, nationality, ethnicity, and English-speaking ability, before being thanked and debriefed.

As expected, participants in the experimental condition perceived David as significantly more counter-stereotypical ($M = 48.09$, $SD = 16.67$) than participants in the control condition ($M = 16.24$, $SD = 9.95$); $t(29) = -7.20$, $p < .001$,[2] Cohen's $d = 2.34$. These findings confirm the adequacy of the manipulation.

Method

Participants. Following previous findings exploring PNS as a moderator of the effect of CSTs on cognitive reflection (Gocłowska & Crisp, 2013), we reasoned that the moderating effect of NFC on the same relation would be medium-sized ($d = 0.50$). Power analysis, conducted using G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007), with an alpha of .05 suggested that 210 participants would provide 95% power to detect an effect of this magnitude. We recruited 397

participants via social media (www.reddit.com) and the crowdsourcing platform Prolific to take part in an online experiment on "imagination and problem solving." Participants either volunteered their time or were compensated with GB£1.30/US$1.80. We determined participant inclusion criteria a priori (see Appendix A) and 315 participants (177 male, 134 female, 3 other, 1 prefer not to say; M_{age} = 29.87, SD_{age} = 10.57; 86% U.S. American nationality, 14% other) were included in the analyses.

Procedure and materials. The experiment comprised three parts and participants completed all tasks online using the survey software Qualtrics (Version 2015). Part 1 was identical to the pretest in that participants were randomly assigned to imagine a stereotypical or a counter-stereotypical CEO named David. We recorded the amount of time that participants spent on this task and also asked them to rate their surprise and astonishment. Next, to reinforce the manipulation, participants were instructed to imagine what David and his life could be like and to describe what characteristics he might possess.

In Part 2, we measured participants' cognitive reflection using the seven-item CRT. The items are designed such that an incorrect solution to each of the seven questions initially comes to mind. Cognitive reflection is demonstrated when the incorrect response is overridden and, upon further reflection, the correct solution is determined. For example, one item states that "Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?" The intuitive, but incorrect, answer is "30," while the correct answer is "29." Participants were presented with seven such problems in a counterbalanced order and were given up to 2 minutes to solve each problem. They were automatically redirected to the next page when the time was up. If they solved the problem in less than 2 minutes, then they were allowed to proceed. The problem-solving task was followed by the manipulation check and attention check, which were identical to the pretest. The manipulation check items had good internal consistency (α = .82).

In Part 3 of the experiment, participants were first asked how vividly they imagined the CST individual (i.e., David) and his life, and several questions about their motivation to engage in the imagination task and CRT, how easy/difficult they found doing so, their feelings about David and his life, as well as the extent to which they are prone to experience awe (Shiota, Keltner, & Mossman, 2007; see Appendix A for all items used). Participants were then asked to indicate whether they were suspicious at any point that the researchers were investigating something other than what was stated, and if so, they were asked to describe what they thought the real purpose of the study was. Next, participants indicated their age, sex, sexual orientation, nationality, ethnicity, English-speaking ability, and their average marks in core high school subjects (namely, English, mathematics, social sciences, and science). Finally, participants completed the 10-item version of the Big Five Inventory (Gosling, Rentfrow, & Swann, 2003), the 10-item Curiosity and Exploration Inventory (Kashdan et al., 2009), and the 18-item NFC scale (Cacioppo et al., 1984). Upon completing these questionnaires, participants were thanked and debriefed.[3]

Analytic approach. The data were analyzed using moderated regression analyses with the pequod package in the programming language *R* (Mirisola & Seta, 2016). Conditions were contrast-coded (−1 = control, +1 = experimental), and we computed a mean score reflecting NFC by averaging the 18 items (reverse-coded where appropriate; α = .95).

Results and Discussion

Manipulation check. To check the adequacy of the CST manipulation, we regressed the mean CST score (i.e., the index of the manipulation check items) on condition, NFC, and their interaction. The predictor variables were centered prior to computing the interaction term. As expected, there was a main effect of condition, b = 12.11,

$t(310) = 12.12$, $p < .001$, such that participants in the experimental condition perceived David as more counter-stereotypical than participants in the control condition. The effect of NFC, $b = -0.65$, $t(310) = -0.94$, $p = .35$, and the interaction term were not statistically significant, $b = 0.41$, $t(310) = 0.60$, $p = .55$. Thus, our manipulation of counter-stereotypicality was successful, regardless of participants' level of NFC.

Randomization check. Prior to exploring whether NFC moderates the effect of exposure to CSTs on cognitive reflection, we checked whether NFC differed across conditions. This is because we had measured NFC as part of the same experimental session and, although unlikely,[4] participants' responses to the measure of NFC may have been affected by the experimental manipulation. A Welch two-sample *t* test revealed that NFC did not significantly differ across conditions, $t(312) = 0.08$, $p = .93$, Cohen's $d = 0.01$, suggesting that the manipulation did not affect NFC scores.

The effects of condition, NFC, and their interaction on cognitive reflection. To explore the role of NFC, we repeated the previous analysis, but this time regressed the number of correctly solved CRT items on condition, NFC, and their interaction. The analyses revealed no main effect of condition, $b = 0.01$, $t(311) = 0.11$, $p = .92$, but a statistically significant main effect of NFC, $b = 0.39$, $t(311) = 4.81$, $p < .001$, such that participants high in NFC consistently outperformed participants low in NFC on the cognitive reflection task. This is not surprising, as previous research has demonstrated that NFC predicts cognitive performance (Cacioppo et al., 1996).

The main effect of NFC was, however, qualified by a marginally statistically significant two-way interaction between condition and NFC, $b = -0.15$, $t(311) = -1.87$, $p = .06$. To understand the nature of the interaction, we inspected the effect of condition (experimental vs. control) on cognitive reflection at different levels of NFC using simple slopes analysis (Aiken & West, 1991). We defined "low NFC" as 1 *SD* below the mean and "high NFC" as 1 *SD* above the mean.

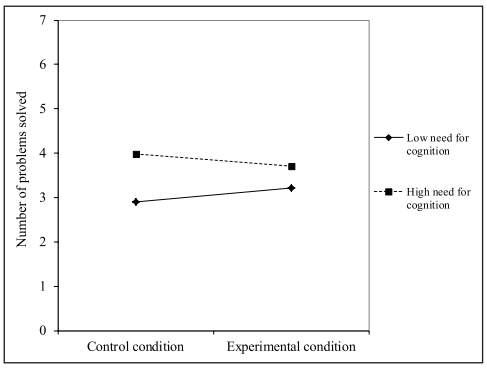

Figure 1. Cognitive reflection as a function of exposure to CSTs at different levels of NFC (Experiment 1).

As Figure 1 illustrates, our analysis revealed two trends: a trend towards a positive effect of condition on cognitive reflection among people *low* in NFC, $b = 0.24$, $t(311) = 1.40$, $p = .16$, and a trend toward a negative effect of condition on cognitive reflection among people *high* in NFC, $b = -0.21$, $t(311) = -1.24$, $p = .21$.

Discussion. The findings of Experiment 1 provide preliminary evidence in support of the second prediction outlined in the Introduction; namely, that exposure to CSTs may benefit people low but not high in NFC. Experiment 2 aimed to replicate the preliminary findings of Experiment 1 with an alternative manipulation of exposure to CSTs.

Experiment 2

Pretest

As before, participants were asked to read a paragraph, but this time describing a person named Mary. Participants in the control condition were asked to imagine that they read the following paragraph on the Internet: "Mary is a secondary school teacher (married, two children), a university graduate (English literature), and UK native. Mary has a positive outlook on life." Participants in the experimental condition were asked to imagine reading the following paragraph instead:

"Mary is a political leader (remarried, two children), a scientist (quantum physics), and a Polish immigrant. Mary has a positive outlook on life." A pretest was used to establish the extent to which these new stimulus materials were deemed to run counter to conventional stereotypes. Specifically, 51 British participants (25 female; $M_{age} = 34.06$, $SD_{age} = 10.15$) were recruited via Prolific and randomly assigned to imagine Mary as a (stereotypical) female teacher or as a (counter-stereotypical) female political leader. After reading the paragraph about Mary, participants were asked how *surprised* and *astonished* they felt and were instructed to imagine what Mary and her life could be like. Following this task, participants indicated their agreement with four statements: "Mary is a typical woman" (reverse-coded), "Reading about Mary challenged some of my beliefs," "There isn't anything puzzling about Mary's life" (reverse-coded), and "Imagining Mary's life made me think 'outside the box'"; all on a scale from 0 (*strongly disagree*) to 100 (*strongly agree*). We created a composite measure of counter-stereotypicality by calculating the mean of the six items, that is, the items measuring surprise and astonishment and the four items measuring counter-stereotypicality. The internal consistency of these items was acceptable ($\alpha = .69$). Lastly, participants were asked to indicate their sex, age, nationality, ethnicity, and English-speaking ability, before being thanked and debriefed.

In support of the adequacy of the manipulation, participants in the experimental condition perceived Mary as significantly more counter-stereotypical ($M = 37.29$, $SD = 17.35$) than participants in the control condition ($M = 26.47$, $SD = 10.14$), $t(40) = -2.73$, $p = .009$, Cohen's $d = 0.76$.

Method

Participants. Based on the calculation of statistical power used in Experiment 1, we again aimed to recruit a minimum of 210 participants for Experiment 2. We recruited 616 participants via a UK university mailing list to take part in an online experiment on "imagination and problem solving." All participants who completed the experiment were entered into a prize draw to win one of two GB£50.00 shopping vouchers or one of five GB£20.00 shopping vouchers. The attention check and participant inclusion criteria were identical to Experiment 1. The final sample consisted of 302 participants (90 male, 206 female, 3 other, 3 prefer not to say; $M_{age} = 24.21$, $SD_{age} = 8.12$; 81% British nationality, 19% other nationality).[5]

Procedure and materials. The procedure and materials were identical to Experiment 1 except for the new manipulation (i.e., Mary the female teacher vs. political leader, rather than David the CEO) and the addition of the Brief Mood Introspection Scale (Mayer & Gaschke, 1988)[6] after the attention check. The manipulation check items had acceptable internal consistency ($\alpha = .69$) and all instructions are reported verbatim in Appendix B.

Results and Discussion

Manipulation check. To check the adequacy of the manipulation of CSTs, we used multiple regression to examine the effect of condition on the mean CST score. We entered condition, NFC, and their interaction term as predictor variables and the mean CST score as the dependent variable. This produced a highly statistically significant main effect of condition, $b = 6.10$, $t(283) = 7.53$, $p < .001$, such that participants in the experimental condition viewed Mary as more counter-stereotypical than participants in the control condition. There was no statistically significant effect of NFC on the mean CST score, $b = 0.38$, $t(283) = 0.47$, $p = .64$, but there was a marginally statistically significant interaction between NFC and condition on the mean CST score, $b = -1.44$, $t(283) = -1.82$, $p = .07$. The main effect of condition on the CST score suggests that the CST manipulation was successful.

Randomization check. A Welch two-sample t test revealed that NFC did not significantly differ between the conditions, $t(277) = -0.87$, $p = .38$, Cohen's $d = 0.10$, suggesting that the randomization to the experimental versus control condition was successful.

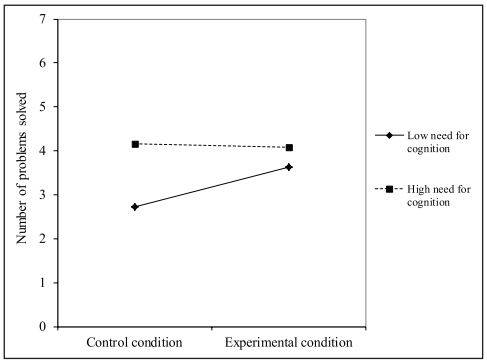

Figure 2. Cognitive reflection as a function of exposure to CSTs at different levels of NFC (Experiment 2).

The effects of condition, NFC, and their interaction on cognitive reflection. To examine the role of NFC, we regressed the number of correctly solved CRT items on condition, NFC, and their interaction. The analyses revealed a trend for the experimental condition to influence CRT performance, $b = 0.20$, $t(284) = 1.68$, $p = .09$, such that participants in the experimental condition outperformed participants in the control condition. There was also a highly significant main effect of NFC on CRT performance, $b = 0.47$, $t(284) = 3.99$, $p < .001$, such that participants high in NFC outperformed participants low in NFC. The main effects were, however, qualified by a statistically significant two-way interaction between condition and NFC, $b = -0.25$, $t(284) = -2.10$, $p = .04$. As Figure 2 illustrates, simple slopes analyses revealed a positive effect of the experimental condition on performance for participants low in NFC, $b = 0.46$, $t(284) = 2.67$, $p = .008$, but no effect of the experimental condition on performance for participants high in NFC, $b = -0.05$, $t(284) = -0.30$, $p = .76$.[7, 8]

Discussion. The findings of Experiment 2 support those of Experiment 1 and provide further evidence in support of the second prediction outlined in the Introduction, which is that exposure to CSTs benefits people low, but not high, in NFC. However, one limitation of Experiments 1 and 2 is that we manipulated counter-stereotypicality and measured cognitive reflection and NFC in the same experimental session. Although our randomization checks showed that NFC did not differ across conditions, Experiment 3 aimed to provide a more rigorous test of the competing predictions by separating the measure of NFC from the experimental manipulation by a week.

Experiment 3

Method

Participants. As in Experiments 1 and 2, we aimed to recruit a minimum of 210 participants. We recruited 344 participants via Prolific to take part in an online experiment on "imagination and problem solving" in return for GB£1.60. The attention check and participant inclusion criteria were identical to Experiments 1 and 2. Our final sample consisted of 270 participants (96 male, 171 female, 2 other, 1 prefer not to say; $M_{age} = 31.59$, $SD_{age} = 10.77$; 99% British, 1% other).

Procedure and materials. The procedure and materials were identical to Experiment 2 except for the following changes. In Part 1 of the experiment, participants answered the questions designed to assess NFC, Big Five personality traits, curiosity, dispositional differences in proneness to awe, and demographic characteristics (sex, age, nationality, ethnicity, English language ability). Part 2 was then administered 1 week later and involved the CST manipulation and CRT. Experiment 3 also incorporated two attention checks. The first attention check was the same as in Experiments 1 and 2 and was placed in Part 1 of the experiment. The second attention check was placed after the question "I was motivated to solve the 7 problems" in Part 2 of the experiment. In addition, we included exploratory items measuring self-relevance of/similarity to the CST individual (see Appendix C).

Results and Discussion

Manipulation check. To check the adequacy of the CST manipulation, we again used multiple regression to examine the effect of condition on the mean CST score. We entered condition, NFC,

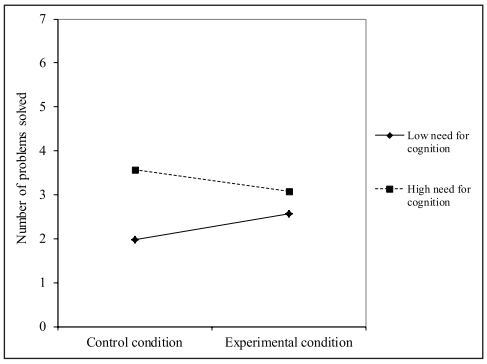

Figure 3. Cognitive reflection as a function of exposure to CSTs at different levels of NFC (Experiment 3).

and their interaction term as predictor variables, and the mean CST score at the dependent variable. This produced a highly statistically significant main effect of condition, $b = 8.35$, $t(256) = 8.88$, $p < .001$, such that participants in the experimental condition viewed Mary as more counterstereotypical than participants in the control condition. There was no effect of NFC on the mean CST score, $b = -0.83$, $t(256) = -1.00$, $p = .32$, but there was a marginally statistically significant interaction between NFC and condition, $b = -1.50$, $t(256) = -1.82$, $p = .069$. The main effect of condition on the CST score suggests that our CST manipulation was successful.

The effect of condition, NFC, and their interaction on cognitive reflection. To examine the effect of condition, NFC, and their interaction on cognitive reflection, we repeated the previous moderated regression analysis, but this time regressed the number of correctly solved CRT items on condition, NFC, and their interaction. We found no effect of condition on CRT performance, $b = 0.03$, $t(258) = 0.19$, $p = .85$, and a highly statistically significant main effect of NFC on CRT performance, $b = 0.53$, $t(258) = 4.18$, $p < .001$, such that participants high in NFC outperformed participants low in NFC. The main effects were qualified by a statistically significant interaction between NFC and condition on CRT performance, $b = -0.27$, $t(258) = -2.44$, $p = .016$. As Figure 3 illustrates, simple slopes analyses revealed a positive effect of condition on CRT performance for participants low in NFC, $b = 0.36$, $t(250) = 1.95$, $p = .064$, and a trend for condition to have a negative effect on CRT performance for participants high in NFC, $b = -0.34$, $t(250) = -1.88$, $p = .113$.[9]

Discussion. The findings of Experiment 3 provide further evidence in support of the second competing prediction; namely, that people low, but not high, in NFC benefit from exposure to CSTs. In the next two sections, we combine the insights from the three experiments meta-analytically in order to gain an estimate of the overall magnitude of the effect.

Metasummary of Effect Sizes Across the Experiments

Because all three experiments investigated the effect of exposure to CSTs on measures of cognitive reflection, we employed a random-effects meta-analysis model (using the metafor package in R; Viechtbauer, 2010) to estimate the average effect of exposure to CSTs on cognitive reflection. Specifically, we computed the sample-weighted (main) effects of condition on cognitive reflection, respectively, across the sample as a whole and also the effect of condition among participants low (vs. high) in NFC separately. The average effect of condition (i.e., exposure to CSTs vs. control condition) on cognitive reflection across the three experiments was $d_+ = 0.08$, 95% CI [−0.06, 0.22]. The fact that the 95% CI includes zero suggests that exposure to CSTs does not generally boost cognitive reflection (or at least not in our sample). This finding stands in contrast to previous research that has reported main effects of exposure to CSTs on cognitive performance (e.g., Gocłowska et al., 2012; Prati, Vasiljevic, et al., 2015). Recall, however, that our primary goal was to test whether exposure to CSTs would change cognitive reflection depending on people's levels of NFC. The meta-analysis across the three experiments showed that, on average, exposure to CSTs had a small to medium-sized positive effect on the cognitive reflection of participants low in NFC, $d_+ = 0.34$, 95% CI

[0.15, 0.54], while exposure to CSTs had a small negative effect on the cognitive reflection of participants high in NFC, $d_+ = -0.18$, 95% CI [−0.38, 0.02]. Note, however, that the 95% confidence interval for the effect on participants high in NFC includes zero, so the evidence for an effect among people high in NFC is weak.

General Discussion

Three experiments explored how exposure to CSTs affects cognitive reflection among US American and British participants who are low (vs. high) in NFC. The findings revealed that participants low in NFC performed better on the CRT following exposure to CSTs than did participants low in NFC who were not exposed to CSTs. Across the three experiments, the average effect of exposure to CSTs among participants low in NFC was small to medium in magnitude ($d_+ = .34$). Interestingly, exposure to CSTs also influenced the performance of participants high in NFC on the CRT. However, unlike participants low in NFC, the cognitive performance of participants high in NFC tended to decrease following exposure to CSTs—an effect that was, on average, small in magnitude ($d_+ = -0.18$). Taken together, these findings provide converging evidence that the effects of interventions based on exposure to CSTs depend on, or are moderated by, individual differences in NFC.

Practical and Theoretical Implications

There has been a surprising dearth of research on whether and how the effects of exposure to diversity on cognitive outcomes differ between individuals. By identifying one moderating variable—namely, NFC—and how it influences the effect of exposure to CSTs on cognitive reflection, the present research represents an important advance in understanding this process. Specifically, the findings of the present research suggest that a simple "one size fits all" explanation of how exposure to CSTs influences performance may be overly simplistic. Failing to consider individual differences in NFC in the effects of exposure to CSTs may, for example, unintentionally give rise to adverse consequences for people high in NFC, although the evidence for a negative effect among people high in NFC was weak in the present research. The practical implications are that both researchers and practitioners need to consider individual differences in NFC when designing and delivering interventions that involve exposing people to counter-stereotypical diversity.

Recall that previous research on exposure to CSTs and PNS suggests that individual differences in PNS moderate the effect of exposure to CSTs on creativity. According to Gocłowska and Crisp (2013), people *high* in PNS seek to organize information in relatively simple ways and therefore dislike experiences that challenge their mental representations. In contrast, people *low* in PNS approach tasks in a more open-minded manner and are less inclined to overgeneralize, which predisposes them to embrace inconsistencies. In line with this reasoning, Gocłowska and Crisp (2013) found that exposure to a CST (a female mechanic) only enhanced creative performance among individuals low in PNS (see also Gocłowska et al., 2014). Also recall that NFC and PNS are typically construed as relatively orthogonal, independent constructs (e.g., see Neuberg & Newsom, 1993, who reported only a weak, negative correlation between NFC and PNS). That is, while NFC represents preferences for the *amount* of cognitive activity, PNS represents preferences for the desired *outcome* of cognitive activity. From this perspective, our findings complement research on the moderating effects of PNS by suggesting that people may require *both* a low level of PNS *and/or* a low level of NFC in order to reap cognitive benefits from exposure to CSTs.

On a theoretical level, the present research extends previous work on how exposure to CSTs affects emotional, motivational, and cognitive outcomes. Specifically, the three experiments reported in this article suggest that exposure to CSTs can sometimes be remarkably powerful, which has theoretical implications for models specifying the psychological effects of exposure to CST diversity (Crisp & Turner, 2011; Gocłowska et al., 2018). That is, it appears that high levels of motivation to

engage in cognitive activity may *not* be required in order for people to engage with CSTs, but instead exposure to CSTs may actually have larger (and more positive) effects among people with relatively low levels of motivation to engage in cognitive activity. However, an important caveat is that too much motivation to engage in cognitive activity can potentially backfire.

Limitations and Future Directions

Several psychological mechanisms may explain why NFC moderates the effect of exposure to CSTs on cognitive performance. On the one hand, it may be that exposure to CSTs triggered interest and curiosity in participants low in NFC, thus boosting their cognitive performance. This idea is in line with research on the emotion of *interest*, which suggests that interest is a "knowledge emotion" that motivates people to learn and explore (Silvia, 2008; see also von Stumm, Hell, & Chamorro-Premuzic, 2011). It also seems likely that people low in NFC have more "headroom"—that is, more potential to open up and be cognitively stimulated—than those high in NFC, which renders a higher capacity to become interested in the first place. With respect to the (small) negative effect of exposure to CSTs among people high in NFC, it is possible that participants high in NFC found making sense of CSTs (i.e., the process of inconsistency resolution) depleting. This is because inconsistency resolution has been reported to be a resource-consuming psychological process wherein people need to suppress existing stereotypes and then generate new impressions of expectancy-violating individuals (Hutter & Crisp, 2006; Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999). People high in NFC may therefore have had less capacity to engage in cognitive activity after exposure to CSTs than people low in NFC, who may have found the process less depleting. Future research needs to test these possible psychological mechanisms to elucidate why exposure to CSTs sometimes has beneficial effects and why it sometimes may backfire.

One limitation of the present research is that some of the conditions postulated by the CPAG model may not have been met, which may have resulted in a failure to replicate the direct effect of exposure to CSTs on performance reported in prior research (e.g., Vasiljevic & Crisp, 2013). For example, whether or not participants actually engaged in inconsistency resolution is unknown. Future research using similar counter-stereotype paradigms is advised to measure this process (e.g., by content-coding the imagery descriptions and using text mining or linguistic analysis) in order to determine when or why it happens or fails to happen. More broadly, developing a method to analyze participants' text responses may help reveal to what extent participants are engaged in the experiments, and whether different types of engagement may influence the findings.

In addition, a limitation but also a strength of the reported research is that the experiments drew on different sources to recruit participants. While all experiments were conducted online, some of the participants were recruited via Prolific and Reddit, and others were recruited via a UK-based university. On the one hand, it is remarkable that the reported patterns of results were relatively comparable across the three experiments and the different recruitment methods, suggesting that the findings are robust. On the other hand, it is difficult to ascertain to what extent the different recruitment methods influenced the reported effects because the experiments also differed from each other in other ways. For example, across the experiments we recruited participants from different countries and tested different manipulations, so it is unclear which of the factors influenced the strength of the effect. It is therefore important that in future research different participant recruitment methods (e.g., in the lab vs. online)—and their potential implications for the hypothesized effects—are taken into consideration.

Finally, we recognize that the present research used a limited range of CSTs and only one measure of cognitive reflection. Future research needs to test whether the effects discovered in the present research can be replicated with different manipulations of CSTs and alternative measures of cognitive reflection/flexibility. Moreover, it

will be important to explore whether the reported effects can arise in different contexts and cultures in order to better understand how universal and generalizable (vs. local and specific) they are.

Conclusion

The role of people's motivation to engage in cognitive activity (i.e., intellectual curiosity) in the effect of exposure to counter-stereotypical diversity on cognitive performance has been relatively neglected in social psychological research to date. Three experiments ($N = 887$ participants) support the idea that exposure to CSTs has a positive effect on cognitive reflection among people low (but not high) in NFC. Taken together, this research contributes to a more nuanced understanding of the effects of exposure to CSTs on cognitive reflection, which in turn could help to maximize the gains and minimize the pains of diversity.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by a Departmental Scholarship from the Department of Psychology, University of Sheffield.

Notes

1. Here we define *cognitive activity* as information processing that enables inconsistency resolution.
2. All *p* values in this paper are two-tailed.
3. We included the measures of vividness, motivation and ability to engage with experimental materials, feelings about David and his life, awe proneness, sex, sexual orientation, age, ethnicity, nationality, English-speaking ability, average marks in core high school subjects, the Big Five personality traits, and trait curiosity purely for exploratory purposes—they were not central to our hypotheses and are not further analyzed. However, Appendix D reports the results of statistical analyses examining the moderating role of the Big Five personality traits, trait curiosity, and sex.
4. It is unlikely that the experimental manipulation affected responses to the NFC scale because the latter is a trait measure. Furthermore, to minimize the possibility that the experimental condition affected responses to the NFC scale, we temporally separated the main part of the experiment from the questionnaires designed to measure individual differences by instructing participants to first answer the demographic questions.
5. Experiment 2 was different from Experiment 1 because we distributed a Qualtrics link via an email server to all students and employees at a large UK university, inviting them to take part in our online experiment. As a result, the experiment was not as constrained as it normally would be in a laboratory setting or on an online platform like Prolific, where you have to set a maximum allowed time to complete a study. Thus, because participation in Experiment 2 was completely voluntary and done remotely, many participants did not have an incentive to finish it or read the instructions carefully. As a result, 151 participants took longer than 30 minutes to complete the study and 163 participants did not pass our attention check.
6. We also included a measure of mood (namely, the Brief Mood Introspection Scale; Mayer & Gaschke, 1988) in Experiment 2 because both positive and negative moods have previously been linked to enhanced cognitive and creative performance (Baas, De Dreu, & Nijstad, 2008; Cheng, Leung, & Wu, 2011; Isen, Daubman, & Nowicki, 1987). Including the Brief Mood Introspection Scale allowed us to investigate whether the effect of exposure to CSTs on cognitive performance holds when controlling for different mood states, and thus to rule out mood as a potential explanation for the effect. To examine whether exposure to CSTs altered mood states related to cognitive reflection, we computed variables representing positive activating moods (7-point Likert-type scale; lively, happy, peppy, loving, caring, and active; $\alpha = .74$), negative activating moods (jittery, nervous, fed up, gloomy, grouchy, and sad; $\alpha = .80$), positive deactivating moods (content and calm; $\alpha = .52$), and negative deactivating moods (tired and drowsy; $\alpha = .66$).
7. Adding the mood variables as covariates to the regression model yielded a more clear-cut pattern of results. Again, there was a trend for experimental condition to influence CRT performance, $b = 0.20$, $t(274) = 1.68$, $p = .10$, and there was a highly statistically significant main effect of NFC, $b = 0.48$, $t(274) = 3.98$, $p < .001$. The two-way interaction between condition and NFC also

remained statistically significant, $b = -0.28$, $t(274) = -2.29$, $p = .02$. Simple slopes analyses revealed a positive effect of the experimental condition on the performance of participants low in NFC, $b = 0.49$, $t(274) = 2.79$, $p = .006$, but no effect of the experimental condition on the performance of participants high in NFC, $b = -0.08$, $t(274) = -0.45$, $p = .65$.

8. To illustrate what happens when less restrictive participant inclusion criteria are applied, we reran the analyses with 35 (instead of 30) minutes as an inclusion criterion. Doing so meant that the sample size increased from 302 to 328 participants. The interaction effect between condition and NFC became marginally significant, $b = -0.19$, $t(310) = -1.68$, $p = .09$. As with the more restrictive inclusion criterion (i.e., 30 minutes), simple slopes analyses revealed a positive effect of the experimental condition on performance for participants low in NFC, $b = 0.37$, $t(310) = 2.25$, $p = .025$, but no effect of the experimental condition on performance for participants high in NFC, $b = -0.02$, $t(310) = -0.14$, $p = .89$.
9. Adding the four mood variables as control variables to the previous regression equation resulted in a more clear-cut pattern of results. The main effect of condition on CRT performance remained nonsignificant, $b = 0.01$, $t(250) = 0.08$, $p = .94$, and the effect of NFC on CRT performance remained highly statistically significant, $b = 0.53$, $t(250) = 4.77$, $p < .001$. Like before, the interaction effect between NFC and condition on CRT performance was statistically significant, $b = -0.30$, $t(250) = -2.67$, $p = .008$. Simple slopes analyses revealed a positive effect of condition on CRT performance for participants low in NFC, $b = 0.36$, $t(250) = 1.95$, $p = .05$, and a negative effect of condition on CRT performance for participants high in NFC, $b = -0.34$, $t(250) = -1.88$, $p = .06$.

Supplemental Material

Supplemental Material for this article is available online.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: SAGE.

Allen, T. J., Sherman, J. W., Conrey, F. R., & Stroessner, S. J. (2009). Stereotype strength and attentional bias: Preference for confirming versus disconfirming information depends on processing capacity. *Journal of Experimental Social Psychology*, *45*, 1081–1087. doi:10.1016/j.jesp.2009.06.002

Baas, M., De Dreu, C. K., & Nijstad, B. A. (2008). A meta-analysis of 25 years of mood–creativity research: Hedonic tone, activation, or regulatory focus? *Psychological Bulletin*, *134*, 779–806. doi:10.1037/a0012815

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, B. W. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197–253. doi:10.1037/0033-2909.119.2.197

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*, 306–307. doi:10.1207/s15327752jpa4803_13

Cheng, C. Y., Leung, A., & Wu, T. Y. (2011). Going beyond the multicultural experience–creativity link: The mediating role of emotions. *Journal of Social Issues*, *67*, 806–824. doi:10.1111/j.1540-4560.2011.01729.x

Crisp, R. J., & Turner, R. N. (2011). Cognitive adaptation to the experience of social and cultural diversity. *Psychological Bulletin*, *137*, 242–266. doi:10.1037/a0021840

Evans, J. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278. doi:10.1146/annurev.psych.59.103006.093629

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. doi:10.3758/bf03193146

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology*, *23*, 75–109. doi:10.1016/S0065-2601(08)60318-4

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, *23*, 1–74. doi:10.1016/S0065-2601(08)60317-2

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25–42. doi:10.2307/4134953

Gocłowska, M. A., Baas, M., Crisp, R. J., & De Dreu, C. K. W. (2014). Whether social schema violations help or hurt creativity depends on need for

structure. *Personality and Social Psychology Bulletin*, *40*, 1–13. doi:10.1177/0146167214533132

Gocłowska, M. A., & Crisp, R. J. (2013). On counter-stereotypes and creative cognition: When interventions for reducing prejudice can boost divergent thinking. *Thinking Skills and Creativity*, *8*, 72–79. doi:10.1016/j.tsc.2012.07.001

Gocłowska, M. A., Crisp, R. J., & Labuschagne, K. (2012). Can counter-stereotypes boost flexible thinking? *Group Processes & Intergroup Relations*, *16*, 217–231. doi:10.1177/1368430212445076

Gocłowska, M. A., Damian, R. I., & Mor, S. (2018). The diversifying experience model: Taking a broader conceptual view of the multiculturalism–creativity link. *Journal of Cross-Cultural Psychology*, *49*, 303–322. doi:10.1177/0022022116650258

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*, 504–528. doi:10.1016/S0092-6566(03)00046-1

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3–25. doi:10.1037/0033-295X.109.1.3

Hutter, R. R. C., & Crisp, R. J. (2005). The composition of category conjunctions. *Personality & Social Psychology Bulletin*, *31*, 647–657. doi:10.1177/0146167204271575

Hutter, R. R. C., & Crisp, R. J. (2006). Implications of cognitive busyness for the perception of category conjunctions. *The Journal of Social Psychology*, *146*, 253–256. doi:10.3200/SOCP.146.2.253-256

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, *52*, 1122–1131. doi:10.1037/0022-3514.52.6.1122

Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The Curiosity and Exploration Inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality*, *43*, 987–998. doi:10.1016/j.jrp.2009.04.011

Leung, A., & Chiu, C.-Y. (2008). Interactive effects of multicultural experiences and openness to experience on creative potential. *Creativity Research Journal*, *20*, 376–382. doi:10.1080/10400410802391371

Macrae, C. N., Bodenhausen, G. V., Schloerscheidt, A. M., & Milne, A. B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology*, *76*, 200–213. doi:10.1037/0022-3514.76.2.200

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology*, *55*, 102–111. doi:10.1037/0022-3514.55.1.102

Mirisola, A., & Seta, L. (2016). pequod: Moderated regression package. R package (Version 0.0–4) [Computer software]. *Rerieved* from http://CRAN.R-project.org/package=pequod

Mussel, P. (2010). Epistemic curiosity and related constructs: Lacking evidence of discriminant validity. *Personality and Individual Differences*, *49*, 506–510. doi:10.1016/j.paid.2010.05.014

Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology*, *65*, 113–131. doi:10.1037/0022-3514.65.1.113

New York Times. (2016, April 21). Fans thank Prince for defying narrow stereotypes about race and gender, and opening the minds of others [Twitter feed]. Retrieved from https://twitter.com/nytimes/status/723233362228465664

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*, 867–872. doi:10.1016/j.jesp.2009.03.009

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. doi:10.1016/j.jesp.2017.01.006

Prati, F., Crisp, R. J., & Rubini, M. (2015). Counter-stereotypes reduce emotional intergroup bias by eliciting surprise in the face of unexpected category combinations. *Journal of Experimental Social Psychology*, *61*, 31–43. doi:10.1016/j.jesp.2015.06.004

Prati, F., Vasiljevic, M., Crisp, R. J., & Rubini, M. (2015). Some extended psychological benefits of challenging social stereotypes: Decreased dehumanization and a reduced reliance on heuristic thinking. *Group Processes & Intergroup Relations*, *18*, 801–816. doi:10.1177/1368430214567762

Qualtrics (Version 2015) [Computer software]. Provo, UT: Qualtrics.

Shiota, M. N., Keltner, D., & Mossman, A. (2007). The nature of awe: Elicitors, appraisals, and effects on self-concept. *Cognition and Emotion*, *21*, 944–963. doi:10.1080/02699930600923668

Silvia, P. J. (2008). Appraisal components and emotion traits: Examining the appraisal basis of trait curiosity. *Cognition and Emotion*, *22*, 94–113. doi:10.1080/02699930701298481

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247. doi:10.1207/s15327957pspr0803_1

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, *20*, 147–168. doi:10.1080/13546783.2013.844729

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. doi:10.1126/science.185.4157.1124

Vasiljevic, M., & Crisp, R. J. (2013). Tolerance by surprise: Evidence for a generalized reduction in prejudice and increased egalitarianism through novel category combination. *PLoS ONE*, *8*, 1–9. doi:10.1371/journal.pone.0057106

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48. doi:10.18637/jss.v036.i03

Von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science*, *6*, 574–588. doi:10.1177/1745691611421204