No longer a threat: a failure to evoke stereotype threat using the race IAT

Bethany S. Jurs
Hannah E. Weber
Mya E. Bowen
William H. Hazelwood
Kathryn D. Cooksey
Iva Katzarska-Miller
Transylvania University

Previous research has argued the Implicit Association Test (IAT) can be used to demonstrate stereotype threat, specifically the fear of white individuals to be perceived as racist. To establish the extent to which this tool can evoke physiological anxiety surrounding the threat, participants completed the IAT under three separate threat conditions while both behavioral and arousal measures were recorded. Results did not show evidence of the stereotype threat, nor any changes in arousal. However, participants did show implicit biases favoring European Americans. These findings suggest several experimental considerations before attempting to use the IAT to evoke a stereotype threat.

Keywords: stereotype threat, racism, implicit association test (IAT), physiological measures

The presence of stereotypes can be found in societal outlets ranging from popular media to subtle passing comments, resulting in difficulty avoiding exposure to stereotypes in regard to one's own identifying or surrounding group. The awareness of potentially negative and false constructs surrounding your social group can lead to a concept defined as stereotype threat; specifically that individuals may feel increased pressure to perform in a certain way so as to avoid validating the perceived negative stereotype of their group (see review by Spencer, Logel, & Davies, 2016, but also Flore & Wicherts, 2015 for a critical meta-analysis). For example, Steele and Aronson (1995) examined the proposed stereotype threat present when African American students took an intellectual exam. In the condition where participants were made to feel that the exam was a direct reflection of their intellectual ability, the threat of confirming a negative stereotype about their racial group led to decrease performance. This is in comparison to the higher performance by the group that was given the same exam but told it was *not* a direct representation of their intellectual abilities.

This stereotype threat has been illustrated for a variety of topics, including gender roles (Spencer, Steele, & Quinn 1999), socioeconomic status (Croizet & Claire 1998), and other racial factors (Devine & Monteith, 1993). It is hypothesized that the presence of cues may trigger stereotype threat by simply reminding targets of culturally held stereotypes, even if the target is presumably unaware of the source of the threat (Spencer et al., 2016; Steele, 1997). Thus, the reminder that the target group is perceived within a certain context could, in turn, affect their performance (Emerson & Murphey, 2015).

Anxiety has been the prevailing explanation as to why performance suffers when experiencing stereotype threat. Specifically, the anxiety people experience when they fear they will confirm a negative stereotype about their group increases their cognitive load to levels that impede performance (Fratz et al., 2004). Physiological studies have supported this theory by demonstrating significant changes in heart rate (Etgen & Rosen, 1993) and skin conductance measurements (Harmon-Jones, Brehm, Greenberg, Simon, & Nelson, 1996; Losch & Cacioppo, 1990) in situations associated with potentially perpetuating a false negative stereotype about one's group. For example, Osborne (2007) assigned participants to groups of either high or low stereotype threat and observed a significant increase in physiological recordings including blood pressure, skin temperature, and skin conductance in the high threat condition. These studies all demonstrate a significant connection between anxiety (such as that experienced in situations of stereotype threat) and physiological arousal.

The Implicit Association Test (IAT) has often been cited as a tool that assesses one's internal associations with prevailing societal stereotypes. Specifically, the test quantifies how easily participants make associations between groups of interest and varying evaluations of the respective group (see Greenwald, McGhee, & Schwartz, 1998 for a complete description of IAT, but Singal, 2017 for a review of the controversy surrounding this measure). Furthermore, Frantz, Cuddy, Burnett, Ray, & Hart (2004) argue that the IAT could be explicitly used as a tool to measure stereotype threat. Specifically, they examined whether (a) white participants performed worse on the race IAT when told the test was a direct reflection of their own racial beliefs and (b) whether self-affirmation can guard against this threat. Their Experiment 1 results illustrate white participants in the threatened condition performed significantly worse on the IAT while Experiment 2 and 3 demonstrate both motivation and self-affirmation are influential factors in modulating the threat impact.

While Franz et al. (2004) did not measure physiological arousal throughout the study, the authors attribute the behavioral findings to heightened anxiety. This explanation aligns with stereotype threat research demonstrating anxiety responses via tasks with a heavy cognitive load (see Croizet Despres, Gauzin, Huguet, Leyen, & Meot, 2004, for example). However, as argued by Frantz et al. (2004), the IAT requires a relatively light cognitive load yet still evokes a behavioral response motivated by stereotype threat. As such, it is uncertain if a lower cognitive load will still elicit a similar anxiety response. By establishing a corresponding physiological anxiety response, the IAT can be further validated as a tool to evaluate this stereotype threat.

Thus, the goal of this study is to demonstrate the reliability of the IAT as a tool to evoke stereotype threat, as well as examine the relationship between stereotype threat and physiological anxiety for a relatively low cognitive demand medium. Specifically, the current study had participants complete the race IAT with the corresponding stereotype threat manipulations as described in Frantz et al.'s Experiment 1 (2004), while simultaneously measuring skin conductance and heart rate. Based on the Frantz et al.'s (2004) behavioral findings, we predict to find similar behavioral impacts across threat conditions as well as see heighted sweat response and increased heart rate in the threat condition, indicating participants feel anxiety about possibly being perceived as racist.

Methods

Participants

Participants were 40 white individuals ranging in age from 18-22. Twelve participants identified as men, while 28 identified as women. All were undergraduate students receiving research participation credit toward an introductory psychology course at Transylvania University, a small liberal arts college in central Kentucky. Data was collected between 2015-2017. Given that Frantz et al. (2004) reported moderate effect sizes, an a priori power analysis using the GPower computer program (Erdfelder, Faul, & Buchner, 1996) indicated that a total sample size of 40 would be needed to detect a medium effect size of .5 (Cohens, 1988) with 80% power using a mixed-factors ANOVA at alpha of .05.

Apparatus

All physiological recordings for this experiment were performed via Biopac MP36 system (Biopac Systems Inc., USA). Changes in electrodermal activity (EDA) were recorded continuously from a two lead electrode set (SS2L) attached to the planar on the participant's left hand. The EDA signal was sampled at .05 Hz High Pass filter with the Biopac *AcqKnowledge* software (Version 4.0.1, BIOPAC Systems, Inc) allowing the baseline reading to always settle to 0 Hz in accordance with the Biopac Systems Inc., USA manufacturer guidelines. EDA change was computed via alternating current (AC) through disposable electrodes (hypoallergenic, monitoring electrode with foam tape and Ag/AgCl solid gel). EDA change was continuously recorded and averaged within each specific condition.

Electrocardiography (ECG) was measured continuously using a 3 lead wire ECG set (SS2LB). The ECG set was placed in a modified Einthoven's triangle with VIN- on right ankle, VIN + on the left wrist, and GND on left ankle. This modification was necessary to avoid data corruption due to the biomechanics of movement when responding to stimuli

(pressing buttons). Alcohol swabs were used to clean participants' skin prior to the attachment of the same type of electrodes as in the EDA recording. ECG signals were displayed on a laptop, using *AcqKnowledge* software (Version 4.0.1, BIOPAC Systems, Inc). Heart rate was calculated from the ECG signal using the Biopac's built-in functions. Similar to EDA, ECG and heart rate were continuously recorded and averaged within each specific condition.

Stimuli & Procedure

Each participant was separately administered the IAT on a Dell computer with a 15-inch monitor through "E-prime" software (Version 2.0; Psychology Software Tools, Pittsburgh, PA, and Schneider, Eschman, & Zuccolotto, 2002). For the IAT, a similar model to that developed by Greenwald et al. (1998) was utilized, in which participants' relative strengths of association were measured based on response latencies and accuracy. Participants were presented with black and white images of clearly African American or European American faces (from the forehead to the nose) as well as words that corresponded to "good" and "bad", and then asked to categorize them either separately (i.e. just words or just faces) or as pairs (i.e. words & faces), dependent on condition. Responses were entered using two different computer keys. The paired conditions had participants categorize either African American faces and good words together or African American faces and bad words together. The face images and words were the same as those used by Greenwald et al. (1998). Of note, Frantz et al. (2004) used different face images in their study.

After electrode placement, Frantz et al.'s (2004) experimental design was explicitly followed. Specifically, participants were randomly assigned to one of three experimental instruction conditions. General verbal instructions were read to all participants, with additional on screen instructions varying based on the experimental condition. Researchers were physically in close proximity while participants read the on screen instructions. This was intended to increase the perceived threat of being exposed confirming a negative stereotype. Identical to Frantz et al. (2004), participants in the *explicit threat* condition were given the following instructions on screen:

> *The IAT compares your attitudes toward two different racial groups. It is a measure of racial bias. In this experiment, we are interested in measuring your unconscious racial attitudes toward Blacks and Whites as accurately as possible. Research shows that a high proportion of Whites show a preference for White people. This is a challenging task, but it's necessary for the aim of this study. Please try hard to help us in our analysis of individuals' racial attitudes.*

In the explicit *no threat* condition, the following instructions appeared:

> *The IAT is a measure of knowledge of cultural stereotypes. In this study, we are interested in measuring the extent to which people are aware of cultural stereotypes. Research shows that knowledge of cultural stereotypes is not related to (1) personal belief in cultural stereotypes or (2) inter-racial attitudes and behaviors. This is a challenging task, but it's necessary for the aim of this study. Please try hard to help us in our analysis of people's knowledge of cultural stereotypes.*

In the no information section, participants were provided with these instructions:

"This is a challenging task, but it's necessary for the aim of this study. Please try hard."

After participants read their respective condition instructions in the presence of the research technician, they were free to individually proceed through the IAT experiment in privacy.

Results

Similar to the Frantz et al. (2004) study, reaction times were established following Greenwald et al. (1998); any response time below 300 ms was changed to 300 ms and over 3000 ms was changed to 3,000 ms.

To parallel the behavioral findings from the Franz et al. (2004) study, we computed a 2x3 mixed-factors ANOVA comparing reaction time and accuracy changes across the paired conditions (Bad Word-African American versus Good Word-African American) and between threat-level groups. However, there were no significant group x accuracy or group x reaction time interactions ($F(2,37) = .92$, $p = .47$ and $F(2,37) = .147$, $p = .86$, respectively). Similar to the behavioral findings, there were no significant interactions between arousal and group threat level for either heart rate or sweat response ($F(2,37) = .56$, $p = .57$ and $F(2,37) = 2.93$, $p =. 07$, respectively).

When collapsing across groups, there were significant main effects across the two paired conditions for both the behavioral (accuracy: $F(1,37) = 7.6$, $p = .009$, $\eta_p^2 = .17$ and reaction time: $F(1,37) = 46.01$, $p < .001$, $\eta_p^2 = .55$) but for neither of the arousal measures. Further analyses of these main effects using a repeated-measures t-test shows overall participants were significantly faster ($M_D = -132.24$) and more accurate ($M_D = .015$) in the condition that paired "bad" words and African American compared to paired "good" words and African American, $t(39) = -6.96$, $p < .001$, $d = -1.1$, 95% CI [-170.68, -93.81] and $t(39) = 2.7$, $p = .01$, $d = .42$, 95% CI [.0037, .026], respectively. This result corresponds to findings repeatedly illustrated in the literature showing white participants tend to more easily associate positive with European American (for a review, see Greenwald et al., 2002). Table 1 shows the averaged group as well as each individual group's mean differences between the Good Word-African American and the Bad Word-African American conditions as calculated via a repeated measures t-statistic. In addition, there were no significant correlations between behavioral measures or arousal responses, $p > .05$.

Discussion

The goal of the current study was to further establish the reliability of the IAT as a tool to evoke a stereotype threat (as shown in Frantz et al., 2004), as well as examine the relationship between the stereotype threat and anxiety through physiological arousal. Although this study was not a direct replication of Frantz et al. (2004) due to the different face images used and addition of the physiological measures, we predicted comparable behavioral results given the similarities in experimental designs. However, we were unable to evoke an instance of stereotype threat using the race IAT. Specifically, while each condition did elicit the standard IAT effect illustrating a preference for European American, there

Table 1. *Behavioral and physiological mean differences between Bad Word-African American and Good Word-African American paired conditions*

Group	Reaction Time (ms)	Accuracy (%C)	Heart Rate (bpm)	Change in Sweat (µS)
Explicit Threat	-120.51 (130.72)*	0.54 (2.49)	-0.84 (3.6)	-1.46 (2.25)
No Threat	-145.54 (91.75)*	1.71 (2.83)*	-2.48 (7.54)	0.71 (2.52)
No Information	-130.41 (143.74)*	2.38 (5.0)	-0.49 (2.78)	-0.46 (2.33)
Combined Group	-132.24 (120.18)*	1.5 (3.52)*	-1.31 (5.12)	-0.40 (2.49)

Table 1 displays each group's mean difference score across the four measures when comparing the Bad Word-African American minus the Good Word-African American paired conditions. In addition, the combined group data is also displayed. Repeated measures t-tests for each group and measure shows replication of classic race IAT findings for reaction time. * = significant at p < .05.

were no significant differences in the IAT data among the different threat group conditions. This contradicts the results outlined in Frantz et al.'s (2004) study that illustrated distinct performance impacts dependent on threat group condition. Furthermore, we were not able to establish any arousal changes across the three group conditions, again indicating that the IAT and threat condition descriptions were not sufficient to elicit feelings of threat.

After examining our results, there are potential explanations for why we were unable to see a change in either variable. The most obvious deviation from Frantz et al. (2004) regards our incorporation of the physiological measurements. Being connected to equipment via multiple electrodes and wires may have served as a distraction to participants resulting in decreased attention to the instructions containing the experimental condition descriptions. In short, if participants in the explicit threat condition were not fully aware of what the test was stated to measure (i.e. this test will measure your racial attitudes), they would not feel threatened to perpetuate a negative stereotype. While this is a candidate explanation and we cannot definitively ascertain the degree to which participants fully read the instructions containing the experimental condition description, we argue against distraction and thus ignorance as a potential factor. To ensure participants were aware of the study description, researchers were physically within close proximity while participants read the instructions containing the experimental condition descriptions. As discussed in the methods, this proximity was intended to heighten the experienced threat of confirming a negative stereotype. In addition, the incorporation of physiological recording has not influenced the expression of stereotype threat in other studies on the topic. Therefore, while the addition of physiological measurements did deviate from the original Frantz et al. (2014) design, we do not see it as a contributing factor for our lack of behavioral or physiological effects.

Given the robustness of the findings produced by Frantz et al. (2004) the most likely explanation for why we were unable to obtain similar results is the possibility of previous exposure. Our participants were undergraduates fulfilling a research participation requirement as part of their introductory psychology course. While none of the offered introductory courses explicitly discussed the IAT as part of the course curriculum while the study was being conducted, it is possible that many students were already aware of it via other forums (e.g. popular media, high school psychology courses, etc.). When Frantz et al. (2004) study was conducted, the IAT was still relatively novel and thus it is more unlikely their participants had any prior exposure to it. As the IAT has gained popularity and become a

regularly used tool across disciplines in and out of academia, many students have arguably been exposed to the task prior to participating in the study. If participants already had prior understanding of how the IAT works, they could have disregarded our explicit threat condition descriptions. This would alleviate any potential threat entering into the study. Without acquiring information regarding their prior exposure to the IAT, it is impossible to rule out this factor. As such, it is recommended that any future researcher attempting to use the IAT to evoke stereotype threat should collect this additional information and restrict their analysis to naive participants.

Despite these potential confounding factors, if one were to assume attentiveness to experiment instructions and naivety surrounding the IAT, a potential theoretical explanation for not evoking a stereotype threat could surround bias blindspot: the decreased likelihood to identify bias in oneself than in others (Scopelliti et al., 2015). It is possible that our participants did not consciously acknowledge their own potential biases nor potential for themselves to be deemed racist. This is despite the fact that they demonstrated an implicit bias preferring European American. This lack of self-identification of a negative stereotype could negate any potential stereotype threat as evidenced by the lack of IAT and physiological response differences; our students arguably did not feel as threatened at being perceived as racist as those in the previous studies. If so, these findings would deviate from previous research conducted in the 90s and early 2000s supporting the idea that many white individuals experience anxiety regarding situations in which they believe their racial attitudes are being evaluated and could be perceived as racist (Devine & Monteith, 1993; Dunton & Fazio, 1997; Greenwald et al., 1998).

To validate this theory, we propose participants naive to the IAT be given an initial race IAT with no instruction for the purpose of demonstrating their internal biases. An explanation of their results should eliminate any bias blindspot the participants may have. After receiving these results, the participants could then be verbally given one of the three threat condition explanations of the race IAT and asked to take the test again. This in turn could demonstrate stereotype threat dependent on the condition and realign results to be more consistent with those found by Frantz et al. (2004).

Conclusion

Despite following a similar experimental design as Frantz et al. (2004), the current study was unable to evoke a stereotype threat using the race IAT. Given changing demographics and the growing discussions of implicit biases and the IAT within both academia and the popular media, the field would benefit from a re-examination of how the IAT can continue as a tool to evoke a stereotype threat.

References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed).* Hillsdale, NJ: Lawrence Earlbaum Associates.

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

Croizet, J. C., Després, G., & Gauzins, M. E. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin, 30*(6), 721-31.

Devine, P. G., & Monteith, M. J. (1993). The role of discrepancy-associated affect in prejudice reduction. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in group perception* (pp. 317-344). San Diego, CA: Academic Press.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23,* 316-326.

Emerson K. T., & Murphy, M. C. (2015). A company I can trust? Organizational lay theories moderate stereotype threat for women. *Personality and Social Psychology Bulletin, 41*(2), 295-307.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28,* 1-11.

Etgen, M. P., & Rosen, E. F. (1993). Cognitive dissonance: Physiological arousal in the performance expectancy paradigm. *Bulletin of the Psychonomic Society, 31,* 229-231.

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology, 53*(1), 25-44.

Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race implicit association test as a stereotype threat experience. *Personality and Social Psychology Bulletin, 30,* 1611-1624.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnum, S. E., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109,* 3-25.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464-1480.

Harmon-Jones, E., Brehm, J. W., Greenberg, J., Simon, L., & Nelson, D. E. (1996). Evidence that the production of aversive consequences is not necessary to create cognitive dissonance. *Journal of Personality and Social Psychology, 70,* 5-16.

Losch, M. E., & Cacioppo, J. T. (1990). Cognitive dissonance may enhance sympathetic tonus, but attitudes are changed to reduce negative affect rather than arousal. *Journal of Experimental Social Psychology, 26,* 289-304.

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from http://www.pstnet.com.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide.* Pittsburge, PA: Psychology Software Tools.

Singal, J. (2017). The creators of the implicit association test should get their story straight. Retrieved from http://nymag.com/intelligencer/2017/12/iat-behavior-problem.html.

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology, 67,* 415-437.

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.

Scopelliti, I., Morewedge, C. K., McCormick, E., Min, H. L., Lebrecht, S., Kassam K. S. (2015). Bias blind spot: Structure, measurement, and consequences. *Management Science, 61*(10), 2468-2486.

Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist,* 52(6), 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.