Title

Stereotype Threat Favors Accuracy Over Speed: Dismantling the Influence of Gender Stereotypes on the Speed-Accuracy Tradeoff in Mathematical Tasks

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Keller, Kim Sophie

Voss, Andreas

Publication Date

2026

Stereotype Threat Favors Accuracy Over Speed: Dismantling the Influence of Gender Stereotypes on the Speed-Accuracy Tradeoff in Mathematical Tasks

Kim Keller (kim.keller@psychologie.uni-heidelberg.de) & Andreas Voss
Institute of Psychology, Heidelberg University

Abstract

Stereotype Threat describes the negative impact on cognitive performance caused by the activation of negative social stereotypes. This study investigates the underlying cognitive mechanisms of gender-specific Stereotype Threat effects. Four online experiments were conducted with 1,164 participants, randomly assigned to either a Stereotype Threat or a control condition. To model the underlying cognitive processes, parameters of the drift diffusion model were estimated and effects on accuracy and response times were reported. Results showed no differences in accuracy as an effect of Stereotype Threat. However, women in the Stereotype Threat condition consistently responded more slowly and, in one experiment, exhibited a higher threshold separation when completing the mathematical task, indicating more conservative response tendencies, compared to men and the control group. The study addresses the need for potential interventions such as stereotype awareness to mitigate these effects and calls for further research into cultural and social influencing factors.

Keywords: Stereotype Threat; drift diffusion model analysis; gender stereotypes; speed-accuracy tradeoff

Introduction

Stereotypes play a fundamental role in shaping social interactions and cognitive processes. Stereotype Threat, a phenomenon first described by Steele and Aronson in 1995, occurs when individuals underperform due to the fear of confirming a negative stereotype about their social group. This effect has been widely studied, particularly in mathematics, where women are often stereotyped as less competent than men (Nguyen & Ryan, 2008; Spencer et al., 1999). Research has demonstrated that even subtle cues are sufficient to induce Stereotype Threat and women tend to perform worse on tasks where gender-related stereotypes are relevant (Nguyen & Ryan, 2008). A meta-analysis reported an effect size for Stereotype Threat effects on women's math performance of d = 0.21 and found that the effect is often moderated by task difficulty and the degree to which individuals identify with the domain in question (Nguyen & Ryan, 2008). Specifically, Stereotype Threat tends to have a stronger influence when tasks are more challenging and when individuals place a higher personal value on their performance in the given domain.

A widely studied explanation for Stereotype Threat effects is its connection to working memory capacity. Schmader and Johns (2003) found that Stereotype Threat impaired women's math performance, mediated by working memory capacity. Additionally, research indicates that practicing tasks can mitigate Stereotype Threat effects, as repeated exposure reduces the working memory demands of a given task, thereby diminishing performance impairment (Beilock et al., 2007). The impact of Stereotype Threat on working memory could also explain why its effects tend to be stronger for complex tasks, where cognitive load is higher (Beilock et al., 2007). Schmader et al. (2008) proposed a process model explaining how Stereotype Threat depletes cognitive resources through physiological stress, self-monitoring, and suppression of negative thoughts, all of which limit working memory capacity.

However, some studies have found no direct link between Stereotype Threat and working memory capacity. For instance, Jamieson and Harkins (2007) examined Stereotype Threat in a working-memory-intensive task with distracting stimuli and found no evidence of interference. Instead, they observed that higher motivation to perform well led participants to choose dominant but incorrect answers more frequently. This finding may explain why Stereotype Threat effects are less pronounced in simpler tasks, where initial intuitive responses are often correct. These results suggest that Stereotype Threat may exert its influence through motivational mechanisms rather than directly impairing working memory (Jamieson & Harkins, 2007).

The Regulatory Focus Theory (Higgins, 1997, 2000) provides another framework for understanding how Stereotype Threat affects cognitive and behavioral responses. This theory distinguishes between two motivational orientations: promotion focus, characterized by eagerness and striving for gains, and prevention focus, marked by vigilance and an emphasis on avoiding losses and errors (Higgins, 1997, 2000). Research suggests that Stereotype Threat induces a prevention focus. Seibt and Förster (2004) showed that negative stereotypes promote avoidance-oriented thinking and lead to slower but more accurate responses. This aligns with findings on loss aversion, error monitoring, and performance-avoidance goals under Stereotype Threat (Carr & Steele, 2010; Forbes et al., 2008; Barber & Mather, 2013; Brodish & Devine, 2009). Regulatory focus manipulations confirm that Stereotype Threat effects appear only under prevention focus (Gaillard et al., 2011).

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

Alternate explanations for Stereotype Threat effects are possible. Individuals may engage in self-handicapping (Shepperd & Arkin, 1989) or vary in their competition motivation, with men excelling in competitive environments while women avoid competition in male-stereotyped tasks (Gneezy et al., 2003; Günther et al., 2010). Additionally, Stereotype Threat reduces self-assessment accuracy and self-control (Bonnot & Croizet, 2007; Inzlicht et al., 2006; Tellhed & Adolfsson, 2018). Physiologically, Stereotype Threat increases arousal and stress responses, such as higher blood pressure, skin conductance, and heart rate (Blascovich et al., 2001; Ben-Zeev et al., 2005; Osborne, 2007). Affective responses, including nonverbal anxiety (Bosson et al., 2004) and test anxiety (Tempel & Neumann, 2014), further mediate performance effects.

The drift diffusion model

All in all, there are many possible cognitive explanations for Stereotype Threat effects. However, it is not always possible to disentangle cognitive mechanisms. One possible method to do so is to use statistical modelling. In this study we make use of the drift diffusion model.

The drift diffusion model is a mathematical framework that captures latent cognitive components underlying rapid binary decision-making, allowing for statistical differentiation of individual parameters (Laming, 1968; Ratcliff, 1978; Voss et al., 2013). It describes evidence accumulation processes and speed-accuracy trade-offs (Ratcliff, 1978). In the drift diffusion model, binary decisions are based on continuous information gathering until enough evidence is accumulated (i.e., an upper or lower threshold is reached), triggering a decision (Ratcliff, 1978; Voss et al., 2013). Information sources can be external stimuli or internal memories, and evidence accumulation follows stochastic sampling with normally distributed noise, explaining response time variability and error rates (Voss et al., 2013).

By analyzing parameters such as drift rate (information accumulation efficiency) and threshold separation (conservatism in decision-making; Ratcliff, 1978; Voss et al. 2004), we can differentiate between competing theories explaining Stereotype Threat. Regulatory Focus Theory (Higgins, 1997) posits that Stereotype Threat induces a prevention focus leading to more conservative response patterns that should be seen in increased threshold separation (Voss et al., 2004). On the other hand, a reduced working memory capacity as theorized by Schmader et al. (2008) would manifest in decreased drift rates as information would be accumulated more slowly under threat.

This study aims to distinguish between these two possible cognitive mechanisms by applying a drift diffusion model analysis to gender Stereotype Threat effects while controlling for motivational effects and gender differences.

Experiment 1

A total of 291 women (mean age = 23.48, *SD* = 7.32) were recruited in the US and Great Britain via Prolific (Palan & Schitter, 2018) and randomly assigned to either a Stereotype Threat or a control condition. Participants completed 120 trials of the Mean Value Computation Task (Lerche et al., 2020). In this task, participants were presented with 16 numbers ranging from 100 to 900 and were asked to estimate whether the mean value was below or above 500.

Participants first received general instructions about the experiment, followed by specific task instructions. They completed a brief training phase before starting the actual task. The Stereotype Threat condition included a priming statement emphasizing gender differences in mathematical performance ("With this task, we assess your mathematical abilities. In addition to your personal skills, we are also interested in whether there are gender differences in the mathematical performance of men and women."), whereas the control condition framed the task neutrally ("With this task, we assess your perceptual abilities."). Additionally, participants in the experimental condition indicated their gender immediately before starting the task. In both conditions, participants were asked to complete the tasks as fast and thoroughly as possible.

Response times and accuracy (percentage of correct trials) were recorded, and post-task surveys assessed participants' identification with the domain, personal evaluation of their skill level, belief in and knowledge of the gender stereotype, motivation, and stress during the experiment. All variables were assessed with a single item on a 7-point Likert scale. A drift diffusion model analysis was conducted to estimate threshold separation, non-decision time, and drift rates as a function of stimulus type. In addition, Bayesian model comparisons were performed using Bayesian regression models to evaluate the effects of Stereotype Threat on accuracy, reaction times, and drift diffusion model parameters. To assess the effect of Stereotype Threat on participants in Experiment 1, we compared a Bayesian regression model including the effect of experimental manipulation with an intercept-only model.

Results and Discussion

Descriptive results for women and mathematical tasks for all four experiments can be seen in Table 1. Control variables had no substantial influence on any dependent variable in any of the experiments.

Regression analyses and model comparisons showed no Stereotype Threat effects on accuracy (difference in accuracy in the threat condition: 0.04, *95%-CI* = [0.00, 0.07], *BF* = 0.33). Women in the Stereotype Threat condition performed as well as women in the control condition. However, women who were presented with the Stereotype Threat manipulation answered more slowly compared to the control group (difference in reaction times in the threat condition:

201.29ms, *95%-CI* = [-5.83, 408.43], *BF* > 100). The model that included the manipulation as a factor was substantially better than the intercept-only model (BF_{10} > 100). All Bayes factors as well as WAIC and LOO for model comparisons can be seen in Table 2.

Table 1: Descriptive mean values of accuracy, reaction times (RTs) and diffusion model parameters in Experiments 1 to 4 for women depending on condition (Stereotype Threat (ST) and control group (Ctrl)).

Experiment		1	2	3	4
Accuracy	ST	0.79	0.78	0.84	0.73
	Ctrl	0.77	0.77	0.85	0.76
RTs in s	ST	1.91	2.05	2.90	27.0
	Ctrl	1.71	1.90	2.85	19.7
Drift rate	ST	1.28	2.98	1.36	-
	Ctrl	1.28	2.93	1.32	-
Threshold separation	ST	2.41	4.59	3.48	-
	Ctrl	2.24	3.64	3.32	-

No substantial differences were found in the diffusion model parameters (differences in the threat condition: drift rate: 0.00, *95%-CI* = [-0.11, 0.11], *BF* = 0.14; threshold separation: 0.17, *95%-CI* = [-0.04, 0.39], *BF* = 0.95). Descriptively, the threshold separation was higher under threat.

Contrary to many studies on Stereotype Threat, we did not find differences in performance defined as accuracy (percentage of trials correct). However, women who received the Stereotype Threat manipulation took longer to respond to trials and were slightly more careful and error-avoiding in their responses. This finding raises the question whether Stereotype Threat effects are not just seen in accuracy as a performance measure. It might be that women under threat change their response strategies and focus on accuracy rather than speed. Stereotype threat effects are a not consistently found in every study (Doyle & Voyer, 2016; Nguyen &Ryan, 2008). That might be because the prioritizing of accuracy over speed might not be detrimental to performance on every kind of task. For tasks with no time limit, performance differences might just not be detected.

Experiment 2

As a next step, we first wanted to replicate our response times results and secondly, make sure that the effects we see are specifically Stereotype Threat effects. To do so, we added two more control conditions. First, men were included in the analysis. If we were to find the same effects for men, the effect seen would not be a Stereotype Threat effect but rather a general effect of the gender manipulation. Secondly, we added a task in which women stereotypically perform better to see whether effects were actually based on a negative stereotype.

Table 2: Model comparisons using Watanabe-Akaike Information Criterion (WAIC), Leave-One-Out Cross-Validation (LOO), and Bayes factors (*BF*) (Gelman et al., 2014; Kass & Raftery, 1995) in favor of the model including the expected Stereotype Threat effects compared to a model that does not. Stereotype Threat effects were defined as follows: Experiment 1: effect of condition; Experiment 2: effect of the three-way interaction of gender, task and condition; Experiments 3 and 4: effect of the interaction of gender and condition.

Model	WAIC difference	LOO difference	BF_{10}
Experiment 1			
Accuracy	-1.9	-1.9	0.33
RTs	-1.6	-1.7	> 100
Drift rate	2.2	2.2	0.14
Threshold separation	0.4	-0.6	0.95
Experiment 2			
Accuracy	2.2	2.3	0.09
RTs	0.2	0.2	> 100
Drift rate	1.0	0.9	1.35
Threshold separation	-4.8	-4.8	29.87
Experiment 3			
Accuracy	1.5	1.5	0.09
RTs	2.0	2.0	> 100
Drift rate	0.9	0.9	0.77
Threshold separation	1.3	1.3	1.78
Number of trials	2.4	4.5	23.66
Experiment 4			
Accuracy	0.5	0.5	0.30
RTs	1.5	1.4	> 100

The Emotion Recognition Task we designed for this purpose required participants to classify facial expressions as positive or negative, an ability often stereotypically attributed more to women. 316 participants were shown 104 pictures of different people taken from the diverse RADIATE stimulus set (Conley et al., 2018). Similarly to Experiment 1, another group of 296 participants completed 104 trials of the Mean Value Computation Task. A total of 612 participants (49.48% female, mean age = 38.78, *SD* = 13.69) were recruited with the same search criteria as in Experiment 1 via Prolific (Palan & Schitter, 2018).

We assessed the same control variables and used the same settings for the drift diffusion model analysis. Again, the goal was to estimate Stereotype Threat effects on accuracy, response times and the drift diffusion model parameters. Here, a regression model with the three-way interaction of gender, Stereotype Threat condition and experimental task was compared with a model without the three-way interaction. Stereotype Threat impairments were specifically expected to be found for women under threat in the mathematical task.

Results and Discussion

Descriptive results for the added control conditions for Experiments 2 to 4 can be seen in Table 3. Again, we observed no differences in accuracy (0.01, *95%-CI* = [-0.05, 0.08], *BF* = 0.09) but women under threat who completed the mathematical task were slower than in the control condition (346.87ms, *95%-CI* = [-171.25, 859.89], *BF* > 100).

Table 3: Descriptive mean values of accuracy, reaction times (RTs) and diffusion model parameters in Experiments 2 to 4 for control conditions. Specifically, men are listed as well as the emotion recognition task from Experiment 2 depending on condition (Stereotype Threat (ST) and control group (Ctrl)).

Experiment		2		2	3	4
Task		Emotion Recognition		Mathematical Task		
Gender		f	m	m	m	m
Accuracy	ST	0.75	0.74	0.80	0.86	0.79
	Ctrl	0.76	0.74	0.80	0.87	0.75
RTs in s	ST	1.44	1.69	2.11	3.02	23.0
	Ctrl	1.48	1.59	2.18	2.93	18.2
Drift rate	ST	1.56	1.51	2.65	1.23	-
	Ctrl	1.56	1.48	2.89	1.37	-
Threshold separation	ST	1.61	1.78	4.33	3.25	-
	Ctrl	1.65	1.71	4.45	3.46	-

Contrary to Experiment 1 we observed differences in the threshold separation parameter (1.17, *95%-CI* = [0.24, 2.09], *BF* = 29.87). Women under threat completing the mathematical task had higher values of the threshold separation parameter, therefore collecting more information before reaching a decision and deciding more conservatively. No differences in drift rate were observed (0.32, *95%-CI* = [-0.29, 0.93], *BF* = 1.35).

As in Experiment 1 we see a possible speed-accuracy tradeoff. Under threat, women seem to avoid incorrect responses and instead respond more slowly. We do not observe the same pattern for men or for a task that is not affected by negative gender stereotypes for women, indicating that our findings are indeed Stereotype Threat effects.

Experiment 3

Because of the possible speed-accuracy tradeoff observed in the previous experiments, we included a time limit to actively induce speed-accuracy tradeoff considerations in participants. 170 participants (74.12% students, 73.53% female, mean age = 29.46, *SD* = 14.77) completed between 47 and 156 trials (mean = 80.59, *SD* = 25.70) of the Mean Value Computation task within the given time limit of two minutes.

We again compared Stereotype Threat effects on accuracy, response times and the diffusion model parameters. This time, we were also interested in differences in the number of trials participants were able to finish. As a second step we compared a model with the interaction of experimental condition and gender with a model with only the main effects.

Results and Discussion

As in Experiment 1, we did not observe any Stereotype Threat effects for women in accuracy (-0.02, *95%-CI* = [-0.08, 0.04], *BF* = 0.09), drift rate (-0.18, *95%-CI* = [-0.57, 0.22], *BF* = 0.77) or threshold separation (-0.37, *95%-CI* = [-1.5, 0.77], *BF* = 1.78). However, we observed not only differences in response times (43.43ms, *95%-CI* = [-726.59, 827.54], *BF* > 100) but also in the number of trials participants were able to complete (1.61, *95%-CI* = [-16.71, 20.00], *BF* = 23.66). Specifically, women under the influence of Stereotype Threat took longer to respond to each trial and subsequently completed fewer trials within the two-minute time limit.

Experiment 3 was able to more specifically examine the speed-accuracy tradeoff by adding a time limit for the participants. However, Experiment 3 did not find the same effects on the threshold separation as Experiment 2. One reason for this might be differences in the sample. The participant sample for this experiment was a convenience sample and comprised mainly of students (74.12%), whereas the sample of Experiment 2 was recruited via an online survey platform. Students in Psychology might not be as susceptible to mathematical threat manipulations as different samples. First, Psychology students learn about the effect at some point in their studies and second, they get in contact with mathematical contexts outside of school, where gender stereotypes about mathematical ability might not be as pronounced. Additionally, the gender proportion could not be collected as equally as via Prolific (73.53% female).

Still, all experiments showed similar patterns. No differences in accuracy could be observed but differences in response times were apparent. A reason for this might be that the task is too easy and accuracy does not show enough variance. Additionally, the Stereotype Threat effect is more pronounced in difficult tasks. We originally chose a simple task with two response options to make a diffusion model analysis feasible. Still, we now wanted to examine whether we find the same response times patterns for more difficult tasks.

Experiment 4

In the fourth experiment, we wanted to test whether the speed-accuracy tradeoffs observed in Experiments 1 and 2 could also be found in more difficult tasks, in which, according to meta-analyses, Stereotype Threat effects are more pronounced. Therefore, we presented 91 participants (23.08% students, 61.54% female, mean age = 39.95, *SD* =

16.60) 50 text-based math questions with two response options to keep the answer format consistent with Experiments 1 to 3.

The same analyses for accuracy and response times were conducted as in Experiment 3. As response times had a mean of 22.7 seconds (*SD* = 20.7 s) a robust estimation of diffusion model parameters was not possible.

Results and Discussion

Like in Experiments 1 to 3, there was no effect of Stereotype Threat for women on mean accuracy in the task (0.07, *95%-CI* = [-0.04, 0.19], *BF* = 0.30). Still, as in the other experiments, women under threat took longer to respond (2476.87ms, *95%-CI* = [-8166.07, 12888.74], *BF* > 100). Therefore, the same pattern of accuracy and response times could be seen for another and a more difficult task, supporting the robustness of our speed-accuracy tradeoff findings.

General Discussion

This study explored the cognitive mechanisms underlying gender-specific Stereotype Threat effects using a diffusion model analysis. Stereotype Threat refers performance impairment caused by the fear of confirming negative stereotypes about one's group. While prior research has associated Stereotype Threat with diminished accuracy due to factors such as working memory depletion and shifts in response strategies, the present findings aim to distinguish between those theories and question the common definition of performance in this domain.

To induce Stereotype Threat, about half of the 1,164 participants across four experiments were alerted to gender differences in mathematical performance before completing their task. Contrary to expectations, the Stereotype Threat manipulation did not substantially affect accuracy in any of the experiments. However, response times were consistently higher for women in the Stereotype Threat condition when completing a mathematical task.

Only Experiment 2 showed differences in the diffusion model parameters that were included to distinguish between theoretical accounts of Stereotype Threat. Women under threat exhibited higher threshold separation values, indicating a more conservative response style. This aligns with Regulatory Focus Theory, which suggests that Stereotype Threat triggers a prevention focus, leading to cautious decision-making (Higgins, 1997). Interestingly, under threat no substantial changes were observed in drift rate, suggesting no change in the speed of evidence accumulation. These findings contrast with earlier research linking Stereotype Threat to slowed processing due to cognitive load. This may be because our chosen tasks do not require a lot of working memory capacity or the decline of working memory capacity cannot be fully mapped on the drift rate of the diffusion model.

The absence of Stereotype Threat effects on accuracy is also notable, given the extensive literature reporting performance declines under Stereotype Threat. Several explanations may account for this discrepancy. First, the narrow variability in accuracy (ranging between 73% and 87%) may have limited this study's ability to detect subtle differences. Second, publication bias may play a role, as studies failing to find accuracy impairments may be underreported. Prior meta-analyses indicate substantial heterogeneity in effect sizes, suggesting that contextual factors influence outcomes (Doyle & Voyer, 2016; Nguyen &Ryan, 2008). Third, response times and the diffusion model may be more sensitive to nuanced effects of Stereotype Threat that conventional accuracy analyses overlook. It may therefore be fruitful to redefine performance in this research domain. Studies that do not report Stereotype Threat effect might have detected them if response times had been included in the analysis. Performance should not only be conceptualized as accuracy; the time required to complete a task should be considered. For example, in timed testing contexts, slower responses constitute a clear disadvantage.

Independent of experimentally induced Stereotype Threat effects, women performed worse than men on the mathematical tasks. Since no actual gender differences are generally found in mathematical performance (Hyde & Mertz, 2009), this difference may reflect differing performance expectations or motivation rather than experimentally induced Stereotype Threat. This suggests that merely presenting a mathematical task may be sufficient to activate concerns related to gender stereotypes, even without explicit reminders. This finding aligns with prior research demonstrating that implicit cues can trigger Stereotype Threat effects (Nguyen & Ryan, 2008). It also implies that passive mitigation strategies, such as avoiding explicit gender references, may be insufficient to prevent Stereotype Threat effects. Instead, proactive interventions that reshape ability perceptions and task framing may be required.

Overall, this study underscores the importance of examining speed-accuracy tradeoffs in Stereotype Threat research. Traditional accuracy-based measures may fail to capture the cognitive adjustments individuals make under threat. The consistent increase in response times and the increased threshold separation observed in Experiment 2 suggest that individuals adopt more cautious decision strategies rather than exhibiting outright performance declines.

Limitations and future research

Across all four experiments, Bayesian regression models and model comparisons were conducted. However, WAIC, LOO, Bayes factor, and credible intervals did not always converge in their indication of model fit and substantial effect size. Therefore, the robustness of the reported findings should be interpreted with caution. In this paper, results based on

Bayes factor were emphasized, while the other criteria were reported but not discussed in detail.

Another direction for future research concerns the impact of Stereotype Threat on learning rather than immediate performance. Prior research has found that Stereotype Threat can impair learning and long-term retention even when immediate test performance remains unaffected (Rydell, 2010). Statistical models incorporating learning processes could help determine whether Stereotype Threat has long-term cognitive effects on skill acquisition in stereotyped domains.

Cultural and societal influences on Stereotype Threat should also be examined. Most studies, including the present one, are conducted in Western contexts with relatively strong gender equality norms. Investigating Stereotype Threat in more diverse cultural settings could clarify how societal attitudes shape its impact. Additionally, examining Stereotype Threat experiences among non-binary and gender-diverse individuals would broaden the theoretical framework beyond binary gender comparisons.

Finally, practical interventions aimed at reducing Stereotype Threat should be developed and tested in real-world settings. This includes educational environments, where gender disparities in STEM fields persist. Explicit strategies, such as emphasizing the absence of gender-based performance differences, have been shown to improve women's mathematical performance (Spencer et al., 1999; Schmader, 2002). Increasing awareness of Stereotype Threat can help individuals reattribute test anxiety and improve performance (Johns et al., 2005), while self-affirmation (Martens et al., 2006), humor (Ford et al., 2004), and intergroup contact (Abrams et al., 2006) can further mitigate its effects.

Conclusion

This study contributes to the growing body of research on Stereotype Threat by providing empirical evidence that gender-based performance differences in mathematics may stem from altered decision-making strategies rather than impaired cognitive efficiency. The findings highlight the importance of considering speed-accuracy tradeoffs when evaluating performance and underscore the need for targeted interventions to reduce stereotype-induced performance gaps and promote gender equality in academic and professional contexts.

References

Abrams, D., Eller, A., & Bryant, J. (2006). An age apart: the effects of intergenerational contact and stereotype threat on performance and intergroup bias. Psychology and aging, 21(4), 691. https://doi.org/10.1037/0882-7974.21.4.691

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype Threat and Working Memory : Mechanisms, Alleviation, and Spillover. Journal of experimental psychology. General, 136, 21 S. https://doi.org/10.1037/0096-3445.136.2.256

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. Journal of experimental social psychology, 41(2), 174-181. https://doi.org/10.1016/j.jesp.2003.11.007

Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. Psychological science, 12(3), 225-229. https://doi.org/10.1111/1467-9280.00

Bonnot, V., & Croizet, J.-C. (2007). Stereotype internalization and women’s math performance: The role of interference in working memory. Journal of experimental social psychology, 43(6), 857-866. https://doi.org/https://doi.org/10.1016/j.jesp.2006.10.006

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(2), 247-255. https://doi.org/https://doi.org/10.1016/S0022-1031(03)00099-4

Carr, P. B., & Steele, C. M. (2010). Stereotype threat affects financial decision making. Psychological science, 21(10), 1411-1416. https://doi.org/10.1177/0956797610384146

Conley, M. I., Dellarco, D. V., Rubien-Thomas, E., Cohen, A. O., Cervera, A., Tottenham, N., & Casey, B. J. (2018). The racially diverse affective expression (RADIATE) face stimulus set. Psychiatry research, 270, 1059-1067. https://doi.org/https://doi.org/10.1016/j.psychres.2018.04.066

Ford, T. E., Ferguson, M. A., Brooks, J. L., & Hagadone, K. M. (2004). Coping Sense of Humor Reduces Effects of Stereotype Threat on Women’s Math Performance. Personality and Social Psychology Bulletin, 30(5), 643-653. https://doi.org/10.1177/0146167203262851

Gaillard, M., Desmette, D., & Keller, J. (2011). Regulatory focus moderates the influence of age-related stereotypic expectancies on older adults’ test performance and threat-based concerns. European Review of Applied Psychology, 61(1), 23-29. https://doi.org/10.1016/j.erap.2010.11.001

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. Statistics and Computing, 24(6), 997-1016. https://doi.org/10.1007/s11222-013-9416-2

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. The quarterly journal of economics, 118(3), 1049-1074. https://doi.org/10.1162/00335530360698496

Günther, C., Ekinci, N. A., Schwieren, C., & Strobel, M. (2010). Women can’t jump?—An experiment on competitive attitudes and stereotype threat. Journal of Economic Behavior & Organization, 75(3), 395-401. https://doi.org/https://doi.org/10.1016/j.jebo.2010.05.003

Higgins, E. T. (1997). Beyond pleasure and pain. American psychologist, 52(12), 1280.

Higgins, E. T. (2000). Making a good decision: value from fit. American psychologist, 55(11), 1217. https://doi.org/10.1037/0003-066X.55.11.1217

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. Proceedings of the National Academy of Sciences - PNAS, 106, 8801-8807. https://doi.org/10.1073/pnas.0901265106

Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as Ego Depletion: How Being the Target of Prejudice Affects Self-Control. Psychological science, 17, 262-269. https://doi.org/10.1111/j.1467-9280.2006.01695.x

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. Journal of personality and social psychology, 93(4), 544. https://doi.org/10.1037/0022-3514.93.4.544

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. Psychological science, 16(3), 175-179. https://doi.org/10.1111/j.0956-7976.2005.0079

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. Journal of the American Statistical Association, 90(430), 773-795. https://doi.org/10.1080/01621459.1995.10476572

Laming, D. R. J. (1968). Information theory of choice-reaction times.

Lerche, V., von Krause, M., Voss, A., Frischkorn, G. T., Schubert, A.-L., & Hagemann, D. (2020). Diffusion Modeling and Intelligence: drift Rates Show Both Domain-General and Domain-Specific Relations With Intelligence. Journal of experimental psychology. General, 149, 2207-2249. https://doi.org/10.1037/xge0000774

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. Journal of experimental social psychology, 42(2), 236-243. https://doi.org/https://doi.org/10.1016/j.jesp.2005.04.010

Nguyen, H.-H. D., & Ryan, A. M. (2008). Does Stereotype Threat Affect Test Performance of Minorities and Women? A Meta-Analysis of Experimental Evidence. Journal of applied psychology, 93, 1314-1334. https://doi.org/10.1037/a0012702

Osborne, J. W. (2007). Linking Stereotype Threat and Anxiety. Educational Psychology, 27(1), 135-154. https://doi.org/10.1080/01443410601069929

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. Journal of behavioral and experimental finance, 17, 22-27. https://doi.org/10.1016/j.jbef.2017.12.004

Ratcliff, R. (1978). A theory of memory retrieval. Psychological review, 85, 59-108.

Rydell, R. J., Rydell, M. T., & Boucher, K. L. (2010). The effect of negative performance stereotypes on learning. Journal of personality and social psychology, 99(6), 883–896. https://doi.org/10.1037/a0021139

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. Journal of experimental social psychology, 38(2), 194-201. https://doi.org/10.1006/jesp.2001.1500

Schmader, T., & Johns, M. (2003). Converging Evidence That Stereotype Threat Reduces Working Memory Capacity. Journal of personality and social psychology, 85, 440-452. https://doi.org/10.1037/0022-3514.85.3.440

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. Psychological review, 115(2), 336. https://doi.org/10.1037/0033-295X.115.2.336

Seibt, B., & Förster, J. (2004). Stereotype threat and performance: how self-stereotypes influence processing by inducing regulatory foci. Journal of personality and social psychology, 87(1), 38. https://doi.org/10.1037/0022-3514.87.1.38

Shepperd, J. A., & Arkin, R. M. (1989). Self-Handicapping:The Moderating Roles of Public Self-Consciousness and Task Importance. Personality and Social Psychology Bulletin, 15(2), 252-265. https://doi.org/10.1177/0146167289152012

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. Journal of experimental social psychology, 35(1), 4-28. https://doi.org/10.1006/jesp.1998.1373

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. https://doi.org/10.1037/0022-3514.69.5.797

Tellhed, U., & Adolfsson, C. (2018). Competence and confusion: How stereotype threat can make you a bad judge of your competence. European Journal of Social Psychology, 48(2), O189-O197. https://doi.org/10.1002/ejsp.2307

Tempel, T., & Neumann, R. (2014). Stereotype threat, test anxiety, and mathematics performance. Social Psychology of Education, 17, 491-501. https://doi.org/10.1007/s11218-014-9263-9

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology. Experimental psychology. https://doi.org/10.1027/1618-3169/a000218

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. Memory & cognition, 32, 1206-1220. https://doi.org/10.3758/BF03196893