Received: 19 January 2021 | Revised: 25 February 2021 | Accepted: 23 May 2021

DOI: 10.1111/jasp.12802

ORIGINAL ARTICLE

Journal of Applied Social Psychology WILEY

The effects of activating a "baby brain" stereotype on pregnant women's cognitive functioning

Madeleine Pownall | Mark Conner | Russell R. C. Hutter

School of Psychology, University of Leeds, Leeds, UK

Correspondence
Madeleine Pownall, School of Psychology, University of Leeds, Leeds LS2 9JT, UK.
Email: M.V.Pownall@leeds.ac.uk

Funding information
This project was funded by a University of Leeds Postgraduate Research award to Madeleine Pownall. Study two was supported by a Gorilla Experiment Builder Researcher Grant

Abstract

Throughout pregnancy and into the immediate postpartum period, women are generally perceived to be incompetent, stressed, and forgetful. However, the neuropsychological "baby brain" literature remains unclear and contradictory. Across two studies, we provide the first experimental tests of whether perceived cognitive impairment in pregnancy can be explained by stereotype threat theory, which proposes that awareness of negative stereotypes about one's ingroup can harm performance. In Study 1 ($N = 364$), we tested stereotype threat effects in a 2 (stereotype threat versus no threat) × 3 (pregnant women versus new mothers versus never-pregnant female control) design. We observed a main effect of group on memory performance (pregnant women and new mothers performed worse than controls), but no other main or interactive effects. Study 2 ($N = 409$) aimed to extend these research questions with mathematics ability, memory, and attention as the dependent variables. Again, we found that a stereotype threat manipulation did not impair pregnant women and new mothers' cognitive performance, nor was there any interactive effects. Groups also did not differ in their performance. We discuss these results in the context of stereotype threat mechanisms, calling into question whether a stereotype threat paradigm can be applied effectively to pregnancy-related stereotypes. This work has implications for the advancement of stereotype threat as a theory and contributes to the reappraisal of the utility of stereotype threat as a way of understanding how stereotypes affect performance.

1 | INTRODUCTION

> "Most mothers, on some level, feel torn between the pleasures, responsibilities, and pressures of children and their own need for financial or emotional resources. We know that the female brain responds to this conflict with increased stress, increased anxiety, and reduced brainpower for the mother's work and her children"
>
> (Brizendine's "The Female Brain", 2007, p. 112).

As Brizendine illustrates, pregnant women are regularly framed by society to be stressed, overworked, anxious, and cognitively less able through their pregnancy (e.g., Brett & Baxendale, 2001; Crawley et al., 2008). This means that throughout pregnancy women are perceived to be less intelligent (Morgan et al., 2013), less committed employees (Correll et al., 2007; Jones, 2017), and at the continual mercy of their hormones (Longhurst, 1997). In line with this, social psychologists have also provided empirical evidence which shows the pervasive levels of maternity prejudice (Longhurst, 1999; Masser et al., 2007) benevolent sexism (Hebl et al., 2007; Sutton et al., 2011) and discrimination (Halpert et al., 1993; Johnson, 2008; Kitroeff &

Silver-Greenberg, 2019) that women face as a result of their pregnancy. This is fueled by the existence and promotion of negative stereotypes about pregnant women (Green et al., 1990), which suggest pregnant women are warm but incapable (Fiske et al., 2002; Glick & Fiske, 1997). These perceptions have important consequences; for example, research shows that pregnancy discrimination leads to negative treatment in the workplace (Bragger et al., 2002; Fox & Quinn, 2015; Little et al., 2015, 2018) and increased postpartum depression symptoms (Hackney et al., 2020).

These negative reactions to pregnant women are largely aligned with the notion that women are incompetent throughout pregnancy (e.g., Brett & Baxendale, 2001). However, it is unclear whether the cognitive changes that women experience occur due to biological reasons (e.g., hormones and changes in sleep patterns), or whether this is a social perception that is fueled by the societal positioning of pregnant women as inferior beings. For example, up to 81% of women self-report experiencing cognitive changes in pregnancy, colloquially referred to as maternal amnesia, "preg head", or, more commonly, "baby brain" (Brett & Baxendale, 2001). In the cognitive literature, this has been presented as poor concentration, worse memory recall, and difficulties with motor coordination (Casey et al., 1999). However, while self-report measures of this phenomenon suggest large effects (Sharp et al., 1993), there are contradictory reports (see a comprehensive meta-analysis by Davies et al., 2018). For example, some studies report large changes in pregnant women's memory on both explicit (Henry & Sherwin, 2012) and implicit (Brindle et al., 1991) tests, whereas others find no differences (McDowall & Moriarty, 2000; Schneider, 1989). Consequently, it is unclear whether "baby brain" results from conformation to socially derived stereotypes or genuine physiological changes, during and immediately following pregnancy.

There is a small but growing body of literature which considers how social explanations may contribute to the so-called "baby brain" phenomena (Crawley et al., 2008; Hurt, 2011; Pownall, 2019). Given the implications that pregnancy perceptions have on women and baby's health (Hackney et al., 2020), it is important to continue assessing the social explanations which explain pregnancy-related cognition. As Bleier (1978) stresses, supposed biological explanations of social phenomena are generally widely accepted, because they instill a sense of order and structure to an otherwise subjective world. There is, therefore, scope to reappraise that which is considered entirely biological, to assess the extent to which social and societal influences play a role. For example, Shahvisi (2020) provides a useful social reappraisal of "nesting" behaviors during pregnancy, calling into question the supposedly hormonally determined behavior of preparing a space for a baby during pregnancy. In her reappraisal, Shahvisi (2020) offers various social explanations of this pregnancy behavior that go beyond the purely biological, including stressing the contribution that gender stereotyping and pervasive gender norms play in pregnancy related behaviors. Here, we suggest that the "baby brain" perception may too be exacerbated by distinctly social phenomena rather than being a sole product of biological determents.

Crawley et al. (2008) attempted to empirically assess whether perceived cognitive impairments during pregnancy are a product of stereotyping, or a real decline due to organic, physiological changes. The authors concluded, following use of both self-report measures and cognitive tasks, that pregnant woman generally rate their cognitive abilities as worse than pre-pregnancy despite only mild differences between the groups on the memory tasks. However, the pregnant participants only performed worse than non-pregnant participants in two performance measures (speed of language processing and attentional switching) from a total of 13 objective measures (including immediate, prospective, and delayed memory recall, efficacy of language processing, planning ability, organization, and selective and divided attention). This suggests that while there is a mild performance difference in cognitive tasks, favoring non-pregnant participants, the perceived difference is greater than the actual effect. This provides initial evidence for the contribution of social stereotypes to the "baby brain" phenomena, and echoes Fine's (2008) notion of "neurosexism" in the field of pregnancy-related cognition. Therefore, in the present work, we are interested in assessing the impact that social stereotypes about pregnant women and new mother's cognitive abilities have on performance. To investigate this, we focus on an area of stereotyping that investigates the impact of negative stereotypes on performance – stereotype threat. To our knowledge, stereotype threat research has not previously investigated "baby brain" related performance reduction.

1.1 | Stereotype threat theory

Here, we posit that stereotype threat theory may provide insights into the contradictory findings in the "baby brain" literature. A body of research suggests that exposure to information concerning self-applicable negative stereotypes undermines performance on associated tasks (Schmader et al., 2008; Steele, 1997; Steele & Aronson, 1995; Wakefield et al., 2012). According to stereotype threat theory, when people fear that their behaviors will confirm a negative stereotype about a group that they are a member of, this worsens performance in the stereotyped domain (Steele, 1997; Steele & Aronson, 1995). For example, exposure to explicit gender-mathematics stereotype-based information, including "*women are poorer at mathematics*" or "*men are better at mathematics*," (Martens et al., 2006; Spencer et al., 1999) can induce women's performance concerns (e.g., Doyle & Voyer, 2016; Shapiro & Neuberg, 2007), prompting a "state of self-evaluative threat" (Koenig & Eagly, 2005, p. 489), which worsens performance in this domain, compared with controls not under stereotype threat (Schmader et al., 2008).

There are generally two competing hypotheses related to the mechanisms that underpin stereotype threat effects (see Pennington et al., 2016, for a useful review). For example, Schmader and Johns (2003) suggest that stereotype threat effects occur due to overload on *working memory*. The working memory approach posits that a stereotype threat manipulation prompts threat-related negative thoughts, which compete for working memory resources with the

task at hand, which harms performance. This is related to the *cognitive load* theory, which suggests that stereotype threat places higher demand on mental resources, which depletes performance (Rydell et al., 2014). Contrastingly, the *mere effort* approach (e.g., Jamieson & Harkins, 2007, 2009) suggests that stereotype threat increases motivation to disprove the negative stereotype, energizing performance, and increasing reliance on *pre-potent* (i.e., dominant and learned; McFall et al., 2009) responding. In other words, when under a stereotype threat, the goal to overcome the stereotype potentiates a well-learned and habitual response (Grandjean & Collette, 2011; McFall et al., 2009). Therefore, if the performance measure is not conducive to reliance on pre-potent responding, participants under a stereotype threat underperform (Davies et al., 2016). An integrated process model developed by Schmader et al. (2008) attempted to blend several accounts for stereotype threat into one single overarching model (including working memory interference, motivation and anxiety). However, the dominant factor in the model, which lies most proximal to performance itself, is working memory interference.

Researchers have found stereotype threat effects related to race and academic performance (Gonzales et al., 2002; Steele & Aronson, 1995), social class and standardized tests (Spencer & Castano, 2007), age and memory (Hess et al., 2003) and gender and math performance (Ambady et al., 2004; McIntyre et al., 2003; Pronin et al., 2004; Spencer et al., 1999). These findings have also been applied to threatened performance in tasks such as memory (Beilock et al., 2007; Hess et al., 2003; Levy, 1996), mental rotation (Moè & Pazzaglia, 2006, 2010) and anagram completions (Wakefield et al., 2012). Stereotype threat in memory performance is of direct relevance to the present work. Further, stereotype threat effects have been found when stereotype information is both explicit (Spencer et al., 1999) and when the stereotype is communicated more implicitly (Smith & White, 2002).

Stereotype threat theory may have important real-world implications for stigmatized groups. For example, research has evidenced a plethora of negative consequences associated with exposure to a stereotype threat, such as lower aspirations in counter-stereotypical careers (Davies et al., 2002), higher task-related anxiety (Bosson et al., 2004), feelings of dejection (Keller & Dauenheimer, 2003), and dissociation from the stereotyped group (e.g., Davies et al., 2005; Major et al., 1998). Therefore, understanding the social mechanisms at play in stereotype threat theory enables targeted interventions to improve performance and reduce stereotype effects.

1.1.1 | Stereotype threat and publication bias

In recent years, there have been concerns of reported publication bias that pollutes the stereotype threat literature (Flore & Wicherts, 2015; Zigerell, 2017), such that only positive effects in support of the stereotype threat effect are published, and null effects are relegated to the "file drawer". These concerns have been amplified given the recent reappraisal of stereotype threat theory in relation to the replication crisis (e.g., Finnigan & Corker, 2016; Forscher et al., 2019; Ganley et al., 2013; Jurs et al., 2019; Pennington et al., 2019), with some researchers finding null effects for replication attempts of classic stereotype threat studies (Ganley et al., 2013; Stoet & Geary, 2012).

As Lewis and Michalak (2019) note, many of the recent failed replications in stereotype threat work have focused specifically on the gender-math performance effect (e.g., Finnigan & Corker, 2016; Ganley et al., 2013; Stoet & Geary, 2012). In response to this, Lewis and Michalak (2019) suggest that stereotype threat effects may be difficult to replicate in modern times due to the dissipation of pervasive negative stereotypes about women's math abilities in comparison to men. Therefore, we suggest that the recent failed stereotype threat experiments may reflect the *endorsement* of stereotypes by contemporary society, rather than a lack of theoretical grounds of the theory itself. To date, the majority of stereotype threat re-appraisal work has concentrated on replicating the math-gender and race-intelligence stereotype-performance combinations (although, for notable exceptions, see Kaye et al., 2018; Pennington et al., 2018). However, there has been a lack of studies that *extend* the enquiries of the experimental paradigm to test other equally pervasive stereotypes, such as the "baby brain" stereotype.

With this logic, there remains a need for experimental social psychology to assess the negative performance effects that may exist due to activation of other more persistent stereotypes, beyond these classic stereotype-performance domains. For example, while general gender stereotype effects related to women's math performance may be dissipating over time, research should further stratify this demographic group to assess more nuanced gendered stereotypes and their effect on women's performance. This will contribute to the ongoing reappraisal of the utility of stereotype threat as a theory that can explain underperformance in stereotyped tasks.

1.2 | The present work

While stereotype threat has been retested extensively in recent years with a particular focus on the "girls are poorer at mathematics" stereotype, to date no research has investigated stereotype threat as a potential explanation of pregnancy-related stereotypes, more specifically, the "baby brain" perception. Memory performance has been found to be affected by a stereotype threat manipulation in previous experimental work (Beilock et al., 2007), which suggests that this paradigm may be useful in understanding "baby brain" effects. Therefore, here, across two studies, we aim to test whether the inconsistencies in seemingly objective inquiries into "baby brain" can be explained by stereotype threat theory. In theory, stereotype threat is a wholly compatible explanation to the perception of women as cognitively less able throughout pregnancy and into motherhood, because it makes the crucial connection between social perceptions and observable, quantitative performance. However, importantly, this theory has not yet been tested. Indeed, any social psychological work which assesses any potential social, cultural, or

societal explanations to "baby brain" remains entirely in its infancy (see Crawley et al., 2008 for an exception).

Therefore, the present work aims to explore empirically whether activation of the "baby brain" stereotype itself can harm objective performance on cognitive tasks, in both pregnant women and new mothers, when compared with a never-pregnant female control. Even if biological differences exist, a stereotype threat manipulation should, in theory, exacerbate performance decrements, thus having an *additive* effect on performance (Sekaquaptewa & Thompson, 2003; Stone & McWhinnie, 2008; Woodcock et al., 2012). In doing so, this research develops the understanding of the principles underlying social cognition throughout pregnancy and continues the reappraisal of stereotype threat theory in a contemporary context, following principles of open and transparent science.

2 | STUDY 1

According to stereotype threat theory (Steele & Aronson, 1995), when a negative performance stereotype about a self-applicable group is made salient, a performance inhibition effect often occurs. One explanation of stereotype threat effects is that exposure to negative stereotype- related information consumes working memory, which thus harms performance (Schmader & Johns, 2003). These effects have been widely reported across studies in relation to girls' math performance (Doyle & Voyer, 2016; Martens et al., 2006; Shapiro & Neuberg, 2007). Pregnant women and new mothers are also stereotypically linked to poorer cognitive performance relative to other groups (Pownall, 2019). Therefore, this gendered stereotype may also have potential to threaten performance when the stereotype is explicitly activated to relevant groups.

To explore this further, Study 1 tested whether a stereotype threat manipulation containing information related to pregnant women's cognitive abilities (i.e., the "baby brain" stereotype) significantly disrupts performance in a memory task. For the purpose of this initial study, memory recall was selected as the outcome measure, due to evidence that suggests the "baby brain" stereotype is most strongly associated with performance in this domain (e.g., Sharp et al., 1993) and previous cognitive research has shown pregnancy-related memory deficiencies (Davies et al., 2018). Immediate recall memory is most frequently cited as a symptom associated with "baby brain" in both social (Crawley et al., 2008) and cognitive research (Brett & Baxendale, 2001; Casey et al., 1999; Cuttler et al., 2011; de Groot et al., 2006; Sharp et al., 1993), therefore, while many facets of "cognitive functioning" are seemingly debilitated by pregnancy, memory recall is the most common and the most compatible with the stereotype threat paradigm. This study tested pregnant women, as well as mothers with children under two years old, due to neurological evidence to suggest "baby brain" effects dissipate after two years postpartum (Hoekzema et al., 2017).

In this initial novel work, we hypothesized that pregnant women and women with young children in the stereotype threat condition will show weaker cognitive abilities than those who are not exposed to threatening information. We also predicted that pregnant women and new mothers would not significantly differ from one another; that is, "baby brain" stereotype will affect both motherhood groups equally. We also hypothesized a significant interaction between group and stereotype threat condition, such that pregnant women and new mothers in the threat condition would perform worse than (i) those in the never-pregnant control group (ii) those who do not receive a stereotype threat.

TABLE 1 Number of participants allocated to condition by group in Study 1

	Condition		
Group	Stereotype threat	Control	Total
Pregnant women	43	59	102
New mothers	64	62	126
Never pregnant (control)	65	71	136
Total	172	192	364

2.1 | METHOD

2.1.1 | Participants and design

A priori power calculation using software G*Power (Faul et al., 2007) indicated that in order to achieve 80% statistical power with an effect size of 0.17 (the lower bound effect size found in Nguyen & Ryan's meta-analysis for blatant stereotype threat manipulations[1]) and an alpha level of 0.05, the sample size should be 337, with six independent groups. This is to test for a 2 × 3 ANOVA, testing for both main effects and interactions.

Participants were primigravid (first-time) pregnant women ($N = 102$; $M_{age} = 27.97$, $SD = 6.05$), new mothers (defined as women with a child less than two years old) ($N = 126$; $M_{age} = 30.98$, $SD = 5.70$), and a never pregnant female control ($N = 136$; $M_{age} = 20.62$, $SD = 4.2$). The study was implemented on Qualtrics. Control participants were mainly recruited from the University of Leeds participant recruitment pool and survey share platforms (e.g., Survey Circle). Pregnant women and new mothers were recruited from Prolific Academic and social media, to allow for more targeted sampling, given the specificities of the participant group. The majority of participants were White British (68%) and currently in paid employment (60.1%).

Participants were randomly assigned to one of two conditions: stereotype threat or a control, using the "evenly present elements" survey flow function in Qualtrics. Breakdown of participant allocation to each condition is displayed in Table 1. The study therefore

[1]Note that there have been concerns over publication bias inflating effect sizes of Nguyen and Ryan's (2008) meta-analysis (Zigerell, 2017) but, given that there is no recent meta-analysis of stereotype threat effects, we have used the lower bound effect size of interest as reported in the 2008 analysis.

followed a 2 (condition: stereotype threat versus control) × 3 (group: pregnant women versus new mothers versus never pregnant controls) between-groups design. Ethical approval was granted by the University of [*redacted for anonymous peer review*] School of Psychology Research Ethics Committee. (Ref: PSC-673) on 16th April 2019.

2.1.2 | Procedure

Participants completed one online testing session[2] that lasted approximately 20 minutes. Participants first read an information sheet detailing the study's procedure and gave informed consent, before answering demographic questions (age, occupation, marital status, education, and ethnicity). Pregnant participants were asked to identify which trimester of pregnancy they were currently in. Participants were randomly assigned to receive either the stereotype threat information or a control, before completing two memory measures (detailed below). In exchange for participation, participants were entered into a prize draw for shopping vouchers. All materials and data for this study are openly available in this Open Science Framework page: https://osf.io/y2m53/.

2.1.3 | Measures

All measures are described in sequential order in the study. All participants received materials in the same order.

Experiences of "baby brain"

Pregnant women and new mothers were asked to report whether they noticed changes in their memory and mood throughout their (current or recent) pregnancy, by way of answering "yes" or "no" to the items "*did you notice any changes to your memory/mood so far in your pregnancy?*". This wording was adapted for each group, to refer to either current or recent pregnancy. Participants were provided with a free text box to provide examples. This textual data will not be reported or analyzed further here.

Stereotype threat

In the experimental condition, participants were told "*Research shows that due to a phenomenon known as 'baby brain', pregnant women and women with young children tend to perform worse on this kind of tasks*". In this condition participants were asked to identify whether they were aware of the term "baby brain" and were asked to give a brief definition. In the control condition, participants were told "*Research has not found any differences between pregnant women and non-pregnant women on these kinds of tasks*". This constituted the stereotype threat manipulation. This type of manipulation has been used in previous stereotype threat literature (e.g., Lesko & Corpus, 2006).

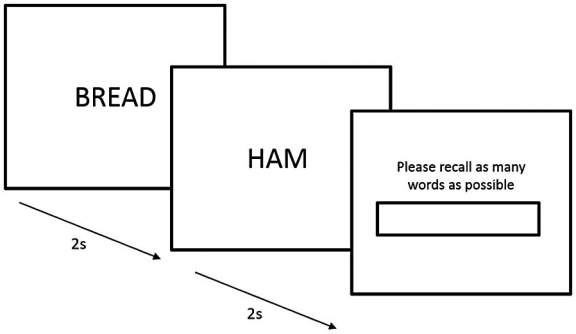

FIGURE 1 Verbal word learning task procedure in Study 1

Verbal word learning tasks

Participants were asked to complete two modified versions of Verbal Word Learning Task (VWLT; adapted from de Groot et al., 2006). The first VWLT showed participants a list of 14 common monosyllabic shopping-list items (e.g., ham, eggs, bread, jam, rice). Each word was presented in isolation to participants in the middle of a white screen and were automatically replaced after two seconds. After viewing all 14 words, participants were asked to free recall as many as they could using a text box using the computer keyboard (see Figure 1).

The second VWLT was adapted for the purpose of this study. This replicated the previous VWLT, but with unrelated monosyllabic nouns rather than food items (e.g., house, gate, hand, yarn). The second task was included due to evidence that suggests that pregnant women differ in their memory performance when tested through applied versus laboratory-based measures (Cuttler et al., 2011). Additionally, this measure allowed comparison with the VWLT shopping list above because shopping lists often consist of finite and thematically similar items; thus, the unrelated nouns list acted a control for guessing on the VWLT shopping list. The procedure was identical to the shopping list VWLT.

Domain identification

Participants were asked to identify the extent to which they value their memory ("*it is important to me that I have a good memory*", "*I have a good memory*") on a 5-point Likert scale, (1 = *Strongly disagree*, 5 = *Strongly agree*). Pregnant women and women with young children were asked to identify the extent to which they agreed with the statement "*it important to me that I am not perceived to have 'baby brain'*". This is due to evidence that suggests stereotype threat effects are only found when individuals under threat value the targeted domain (Appel et al., 2011).

Motherhood social identity

Pregnant women and women with young children were asked to complete a brief 4-item motherhood social identity identification measure, which included items such as "*Being a mother is an important*

[2]Online testing was deemed appropriate for this study largely due to feasibility concerns and evidence that participants tested online do not differ substantially from participants tested in the lab (Casler et al., 2013). We discuss the theoretical implications of this in more detail in our general discussion.

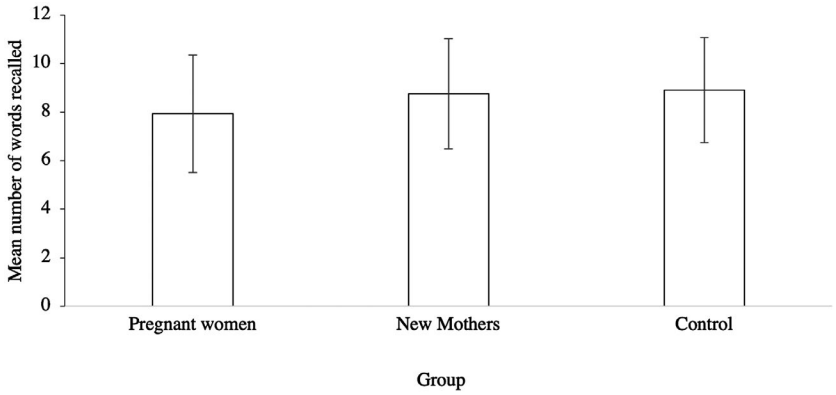

FIGURE 2 Mean number of words recalled in the VWLT shopping list task per group in Study 1. Error bars represent standard deviation

part of my self-image" and reverse-score items "*Being a mother has little very little to do with how I feel about myself*". Participants indicated their agreement on a 5-point Likert Scale (1 = *Strongly disagree*, 5 = *Strongly agree)*. This measure was included to assess the extent to which participants align themselves with the social identity under threat, which is a key factor in stereotype threat theory (Shih et al., 1999).

Dependent measures

The main dependent measures in this experimental study were the number of items correctly recalled on both the Verbal Word Learning Tasks.[3]

2.2 | Results

2.2.1 | Exploratory results

Two separate between-subjects ANOVAs were conducted to explore the experimental hypotheses. The first was a 2 × 3 ANOVA. The Independent Variables were "Group" which had three levels: primigravid pregnant women, new mothers, and non-pregnant controls. Secondly, "Condition" had two levels: Stereotype Threat and Control. The dependent variables were the number of correctly recalled words in the VWLT Tasks.

2.2.2 | Verbal word learning task 1

Memory performance (indexed by number of correctly recalled words) was not affected by participant's demographic factors. An ANOVA confirmed that demographic factors did not impact recall on VWLT1; there was no significant main effects of ethnicity ($p = .26$), education level ($p = .33$), employment status ($p = .45$) or marital status ($p = .09$) on recall in this task. A linear regression also confirmed that participants' age did not independently predict recall on VWLT1 ($p = .76$).

A 2 (Condition: stereotype threat versus control) × 3 (Group: pregnant women versus new mothers versus never pregnant control) ANOVA showed a significant main effect of group, $F(2, 543) = 5.67$, $p = .004$, $\eta p^2 = 0.033$. Follow-up post-hoc tests with Bonferroni corrections indicated that pregnant women ($M = 7.94$, $SD = 2.42$) recalled significantly fewer words than both new mothers, ($M = 8.75$, $SD = 2.27$), $p = .029$, 95% CIs [−1.57, −0.06] and controls, $M = 8.91$, $SD = 2.17$, $p = .005$, CIs [−1.71, −0.23]. New mothers and controls did not significantly differ from one another, $p = 1.00$, CIs [−0.85, 0.54] (see Figure 2).

Contrary to hypotheses, there was no significant main effect of condition, $F(1, 343) = 0.38$, $p = .537$, $\eta p^2 = 0.001$, nor was there a significant interaction between group and condition, $F(2, 343) = 0.67$, $p = .514$, $\eta p^2 = 0.004$.

2.2.3 | Verbal word learning task 2

As with VWLT1, memory performance (indexed by number of correctly recalled words) in VWLT2 was not affected by participant's demographic factors. An ANOVA confirmed that there were no significant main effects of ethnicity ($p = .32$), education level ($p = .79$), employment status ($p = .86$) or marital status ($p = .26$) on recall in this task. A linear regression also confirmed that participants' age did not independently predict recall on VWLT2 ($p = .13$).

Across groups, participants recalled significantly fewer words in the unrelated Verbal Word Learning Task (VWLT2, $M = 8.70$, $SD = 2.17$) compared with the shopping list Verbal Word Learning task (VWLT1, $M = 6.69$, $SD = 2.35$), $t(330) = 15.4$, $p < .001$.

An identical 2 × 3 ANOVA was conducted, this time with performance on the unrelated noun VWLT as the dependent variable (DV2). Again, there was a significant main effect of group, $F(2, 333) = 8.43$, $p < .001$, $\eta p^2 = 0.049$ (Figure 3). Follow-up post-hoc tests with Bonferroni corrections indicated that pregnant women ($M = 6.08$, $SD = 1.96$) recalled significantly fewer words than controls, ($M = 7.31$, $SD = 2.51$), $p < .001$, 95% CIs [−1.99, −0.45] but this time performed similarly to new mothers ($M = 6.41$, $SD = 2.34$), $p = .975$, CIs [−1.10, 0.46]. New mothers performed significantly worse compared with

[3]As an exploratory measure, participants also completed the short-form Feminist Identity Development Scale (Bargad & Hyde, 1991) and a short-form version of the Big 5 personality scale (Rammstedt & John, 2007). These measures are not further reported, due to our lack of clearly defined hypothesis in relation to these constructs.

FIGURE 3 Mean number of words recalled in the VWLT unrelated nouns task per group in Study 1. Error bars represent standard deviation

controls $p = .008$, CIs [−1.62, −0.18]. Contrary to hypotheses, there was no significant main effect of condition, $F(1, 333) = 0.66$, $p = .419$, $\eta p^2 = 0.002$, nor was there a significant interaction between group and condition, $F(2, 333) = 0.078$, $p = .926$, $\eta p^2 = 0.000$.

2.3 | Exploratory results

2.3.1 | Experiences of "baby brain"

In order to assess self-reports of pregnancy-related change in memory, we first assessed the frequencies of self-reported memory and mood changes in pregnancy, for the pregnant women and new mother respondents. Overall, the majority (65.1%) of new mothers reported that they noticed negative changes to their memory capacity while they were pregnant. 83.2% of this group also reported changes to their mood during pregnancy. For the currently pregnant women, 51.5% reported noticing explicit changes to their memory so far in their pregnancy, whereas 70.9% reported noticing differences in mood. A Chi-squared test revealed that these self-reported changes in memory and mood were not associated with the current trimester of the participant's pregnancy, X^2 (2, N = 103) = 1.265, $p = .53$.

2.3.2 | Age effects

As an exploratory analysis, we tested whether participant's age predicted number of items correctly recalled on the Verbal Word Learning Task. A stepwise multiple regression analysis was used to test whether group and age predicted recall on VWLT1. The regression indicated that Model 1, Group explained 2.3% of overall variance, $F(1, 340) = 9.05$, $p = .003$ in word recall on the first task. Changes in word recall were significant predicted with group ($t = 3.00$, $p = .003$) In Model 2, when age was entered, adjusted R squared increased to 3.2%, $F(2, 340) = 6.59$, $p = .002$. Changes in word recall were thus significantly associated with the addition of age as a factor $t = 2.01$, $p = .045$.

2.3.3 | Social identity as a moderator

To test whether motherhood social identity moderated the relationship between stereotype threat condition, we conducted a moderation analysis using Haye's PROCESS Macro v3.4 on SPSS. This was to test whether identification with the identity under threat (i.e., "motherhood", in this instance) affects susceptibility to a stereotype threat (Nosek et al., 2002). The moderation model, which included condition as the independent variable, motherhood social identity as a moderator, and recall on VWLT as the outcome, explained 9.75% of overall variance in the outcome. Overall, the moderation model was not significant, $F(3, 104) = 0.33$, $p = .80$. The effect of condition on recall was not significant $\beta = 1.82$, SE = 1.93, $p = .35$, nor was motherhood social identity $\beta = 0.61$, SE = 0.80, $p = .45$. The addition of the interaction term increased the R^2 by 0.007, which was not statistically significant, $F(1, 104) = 0.73$, $p = .39$, $\beta = -0.43$. Therefore, moderation did not occur.

2.4 | Study 1 discussion

To summarize, we have extended the application of stereotype threat theory and used the theory to test a novel stereotype-performance dyad, the "baby brain" stereotype and associated memory ability. Study 1 demonstrated that pregnant women generally underperformed in memory-based tasks compared with new mothers and female controls. This result replicated in two objective memory measures. However, contrary to our hypotheses, the presence of a negative performance stereotype did not significantly impact memory performance. There was also no interaction between the stereotype threat condition and group, which was also not aligned with our original hypotheses.

Moreover, on the surface, these results could support the notion that "baby brain" is a genuine, physiological phenomenon, owing to, for example, hormonal changes in pregnancy (e.g., Henry & Rendell, 2007). However, the lack of stereotype threat effects in our results may be due to a number of factors. For example, research shows that stereotype threat effects may only be present if

certain moderators (e.g., Shapiro & Neuberg, 2007) or mediators (Pennington et al., 2016) are accounted for. Moreover, given the novelty of the present work, these results must be interpreted with caution. This is the first study that assesses the capacity of stereotype threat to explain cognitive decline in pregnancy, and therefore, further work should extend this line of questioning in order to paint a more complete picture of the utility of the theory. It is also important to fully stratify the concept of "baby brain" in testing concepts. Immediate memory recall, as used in this work, is only one of the various cognitive areas thought to be associated with cognitive decline in pregnancy and new motherhood (Davies et al., 2018). This must be investigated fully in future studies.

Finally, our exploratory analysis showed that, while age alone was not a significant predictor of participants' memory recall, it did contribute a small but significant portion of variance in recall when coupled with group effects. Therefore, future studies should aim to consider how age may be a factor in other applications of stereotype threat, given the link between memory and cognitive ability (e.g., Finkel et al., 2005).

3 | STUDY 2

Study 2 was a conceptual replication of the stereotype threat paradigm in Study 1. This is in line with Crandall and Sherman's (2016) notion that conceptual replications in social psychology are required to enhance confidence in theoretical hypotheses. In this second study, we again tested the stereotype threat experimental paradigm with an identical stereotype threat manipulation and with pregnant women, new mothers, and a never-pregnant female controls as participant groups. However, this time, we aimed to extend and diversify the outcome measure of "baby brain" to include more diverse facets of cognitive functioning, rather than focusing on memory ability alone. This is due to research which suggests that the "baby brain" concept extends beyond memory capacity alone, and also sees deficits in a wider range of cognitive abilities (Davies et al., 2018), such as executive control and general cognition (e.g., de Groot et al., 2006; Onyper et al., 2010).

In this extension of Study 1, we aimed to extend the original theoretical hypotheses, to test whether group effects replicate in wider cognitive domains: memory, executive control, and general cognition. These outcome measures have all been used to objectively test pregnant women's cognitive functioning from a neurocognitive perspective in previous research (e.g., memory; Henry & Rendell, 2007; general cognition; Brindle et al., 1991; and executive functioning; Raz, 2014). Therefore, we expected that pregnant women's and new mothers' potential (under)performance in these wider cognitive domains will also fit within the "baby brain" social stereotype.

In directly replicating the stereotype threat manipulation with a broader range of outcome variables, this allows us to investigate whether stereotype threat or group effects, or indeed interactions between the two, are affected by the nature of threat itself or the measure of performance. Therefore, in this study we predicted a significant main effect of group, such that pregnant women and new mothers would both underperform compared with a never-pregnant female control, but not when compared with one another. We also predicted that pregnant women and new mothers in the stereotype threat group would perform worse than those who do not receive the threat; never-pregnant female controls will not be affected by the stereotype threat content. In other words, stereotype threat would have an *additive* effect (Sekaquaptewa & Thompson, 2003; Stone & McWhinnie, 2008); that is, the presence of a threat will *exacerbate* existing deficiencies in pregnant women and new mothers' cognitive ability that occur due to biological changes. In this sense, we expect that pregnant women will do the worst, followed by new mothers, followed by control group, who should by unaffected by the contents of the stereotype threat manipulation.

3.1 | Method

3.1.1 | Participants & design

A priori power calculation using software G*Power (Faul et al., 2007) indicated that to achieve 85% power and to detect an observed partial η^2 of 0.033 (converted to a Cohen's *f* of 0.18, as per the original study of which this replication is based) and an alpha of 0.05, that the sample size should be at least 341, with six independent groups. Again, this allows us to test a 2 × 3 ANOVA with main effects and interactions.

As with the previous study, participants were new mothers with a child less than two years old ($N = 110$), primigravid pregnant women ($N = 130$), and female never-pregnant control ($N = 169$), see Table 2 for breakdown of participants by condition. The mean age was 30.59 ($SD = 6.13$). The majority of participants were employed (75.1%), White British (60.9%) and educated to a bachelor's degree level (71.7%). Ethical approval was granted by the University of Leeds School of Psychology Research Ethics Committee (Ref: PSC-878) on 31st January 2020.

Participants were all recruited on Prolific Academic and custom pre-screening was applied. Control group participants were 25 years or older in order to roughly age-match this group with the pregnant women and new mothers. Participants registered with Prolific Academic must have indicated that they are female and answered "yes" to the "I am currently pregnant" screening item on Prolific (for

TABLE 2 Number of participants allocated to condition by group in Study 2

	Condition		
Group	Stereotype threat	Control	Total
Pregnant women	61	69	130
New mothers	58	52	110
Never pregnant (control)	85	84	169
Total	204	205	409

the pregnant women group). Participants were paid the equivalent of £5.00 per hour to complete the experiment, which took 10–15 min to complete. As per the study that this conceptual replication is based on, this study followed a 2 (condition: stereotype threat versus control) × 3 (group: pregnant women versus new mothers versus control) between-subjects design.

3.1.2 | Procedure

The experiment procedure was built in the experimental software Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2019). Participants completed one online testing session that lasted approximately 10–15 min. Participants first read an information sheet detailing the study's procedure before answering demographic questions (age, occupation, marital status, education, and ethnicity) and providing informed consent. Participants were randomly assigned to receive either the stereotype threat or a control, before completing three measures of cognitive ability.

3.1.3 | Measures

This study was preregistered on the preregistration server as predicted on 26th February 2020 (Ref: #36324; https://aspredicted.org/zi6ke.pdf). All measures and materials can be accessed via our Gorilla Open Materials link here: https://gorilla.sc/openmaterials/80176. The three tasks were counterbalanced in Gorilla using the counterbalancing "node". All measures are described in sequential order in the study. All participants received materials in the same order.

Stereotype threat manipulation

In the randomly assigned experimental condition, participants were exposed to a negative performance stereotype: "*Research shows that due to a phenomenon known as 'baby brain', pregnant women and women with young children tend to perform worse on these kind of tasks*". Participants in the threat condition were asked to identify whether they were aware of the "baby brain" concept (yes, no). This threat manipulation has been used in previous work (e.g., Lesko & Corpus, 2006). The control condition did not receive this information.

Mathematics ability

To test mathematical ability, participants were shown short mathematics questions, of which were based on either arithmetic (e.g., "10−4 = ?") or knowledge (e.g., "How many weeks in a year?"). These questions were sourced and adapted from Gorilla's bank of experimental materials (https://gorilla.sc/openmaterials/80176) Participants were asked to write their answer using the keyboard and press the enter button when complete. Each question timed out after 5000ms. There was a 200 ms delay between each question. Participants were awarded one mark for every correct item. There were 17 items in total, 11 arithmetic and 6 knowledge based.

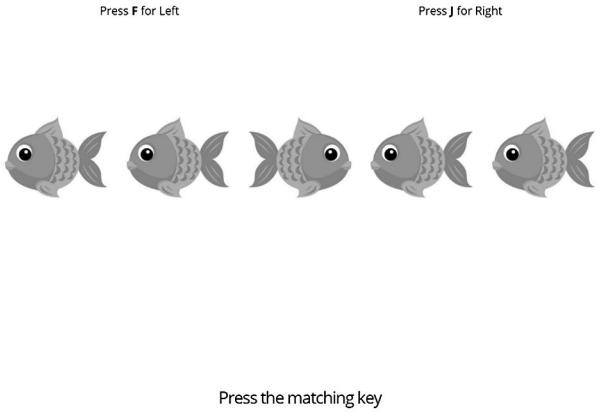

FIGURE 4 Flanker Task participant display in Study 2

All questions were designed to be of equal difficulty. Items to all subsections of the mathematics questions were computed into one mean score.

Executive control

To test executive control, participants were asked to complete a modified version of the Flanker Task (Eriksen & Schultz, 1979). For the purpose of this experiment, the stimuli were animated fish (see Figure 4). Participants were told to press the "*F*" key on their computer keyboard if the target fish, at the center of the screen, was swimming to the Left, and the "*J*" key if the fish was swimming to the Right. There were four practice trials and 17 experimental trials. In this task, participants must ignore peripheral attentional cues (distractor fish) and make decisions about the immediate focal stimuli. Number of correct responses were recorded.

Memory ability

The Levels of Processing task (Craik & Lockhart, 1972) was used to assess memory performance in this study. In the task, participants are shown a word (e.g., "floor") and are asked questions which relate to its structure (e.g., "does the word have four letters?"), phonemic features (e.g., "does the word rhyme with [word]?"), or semantics (e.g., "does the word fit into this [sentence]?"). The words automatically advanced to the next screen after 2000 ms. There were 30 trials, with 10 phonemic questions, 10 structural, and 10 semantic. Participants were awarded one point for each correct answer.

Manipulation check

As a stereotype manipulation check, participants in the stereotype threat condition were told "*Earlier in this experiment you were told some information about pregnant women's cognitive abilities*" and were asked to identify the content of this information from a drop-down list with four options ("pregnant women have [worse/better] cognitive abilities compared with other groups", "there are no differences in the cognitive abilities of pregnant women compared with other groups" or "I don't remember"). This manipulation check was adapted for the purpose of this work.

Stereotype endorsement

To check for participants' endorsement of the stereotype, participants were asked "*Do you think pregnant women do have reduced cognitive abilities (i.e., poorer memory, planning, and attention) compared with other groups of people*" (yes, no, unsure). This was adapted from Jamieson and Harkins (2009).

Cognitive ability identification

Due to evidence that suggests that participants' value of the domain under threat (i.e., in this case, cognitive functioning) affects susceptibility to a stereotype threat that targets the domain, participants answered a 5-item measure of cognitive ability identification. This included items such as "*I think I have good cognitive abilities*" and "*it is important to me that I am perceived to be intelligent*" on a 1 (strongly disagree) to 7 (strongly agree) Likert scale. This was adapted from Spencer et al. (1999) work on mathematics ability identification. Cronbach alpha showed that cognitive identification had a satisfactory internal reliability ($\alpha = 0.73$).

Attention check

As an amendment on the previous stereotype threat study, participants in this study were asked to "*Select 'Agree' for this question*" on one item, to check their attention. This is to ensure that participants included in the final dataset were suitably attentive to the study. Twelve participants failed the attention check (i.e., they did not select "Agree" when asked "choose Agree for this question") and therefore, their data was omitted from further analyses, leaving a final dataset of 128 pregnant women, 109 new mothers, and 160 controls.

Motherhood social identity

As per the previous study, new mothers and pregnant women answered a four-item measure of motherhood social identity, including items such as "*Being a mother is important for my self-image"* with two items reverse-scored (e.g., *Being a mother has very little to do with how I feel about myself*). This involved four statements and participants indicated their agreement on a 1 (strongly disagree) to 7 (strongly agree) Likert scale. Cronbach alpha showed that cognitive identification had a satisfactory internal validity ($\alpha = 0.78$).

3.2 | Results

3.2.1 | Confirmatory results

The number of correct responses to each of the three tasks were summed to create one index per task that represented overall performance. We first analyzed attention and manipulation checks, before conducting three separate 2 × 3 between-subjects ANOVAs for each of the key dependent variables: performance on the Flanker Task, math task, and Levels of Processing task. We followed up any significant main effects of group with post-hoc comparisons with Bonferroni corrections. When there is a significant effect of group on performance, we then conducted a moderated regression with group as the predictor, performance on each task as the outcome, and cognitive identification as the moderator, to test whether identification in the domain of threat affects susceptibility to negative performance effects.

A MANOVA with performance on all three tasks as the outcome variable confirmed that there was no main effect of age in performance on either the Levels of Processing task ($p = .21$), Flanker task ($p = .47$) or math task ($p = .22$).

Responses to the manipulation check were analyzed using a Chi-Squared test to check that participants in the stereotype threat condition could correctly identify the contents of the stereotype threat. There was a significant effect of condition on information identification X^2 (3, $N = 397$) = 210.97, $p < .001$, such that the vast majority participants in the stereotype threat condition were able to correctly identify the contents of the stereotypical information (69.19%, $N = 137$) and participants in the control condition were also able to identify the contents of the control information (83.42%, $N = 155$).

To test the effects of group and stereotype threat condition on performance on the executive control, measured through the Flanker Task, we conducted a 2 (Condition: stereotype threat versus control) × 3 (Group: pregnant women versus new mothers versus never pregnant control) ANOVA, with performance on the Flanker task as the dependent variable. This revealed no significant main effect of group $F(2, 390) = 2.110$, $p = .123$ $\eta p^2 = 0.01$, no significant main effect of condition, $F(1, 390) = 0.43$, $p = .514$, $\eta p^2 = 0.001$, and no interaction between the two variables, $F(2, 390) = 0.66$, $p = .52$, $\eta p^2 = 0.003$.

A second 2 (Condition: stereotype threat versus control) × 3 (Group: pregnant women versus new mothers versus never pregnant control) ANOVA, with performance on the mathematics task as dependent variable revealed a significant main effect of group, $F(2, 390) = 5.84$, $p = .003$ $\eta p^2 = 0.003$. Post-hoc tests with Bonferroni corrections showed that new mothers performed significantly better ($M = 13.02$, $SD = 2.08$) than controls ($M = 12.04$, $SD = 2.84$) $p = .004$ CIs [0.25, 1.71]. Pregnant women performed marginally significantly better ($M = 12.75$, $SD = 2.25$) than controls, $p = .045$ CIs [−1.41, −0.01], but there were no differences between pregnant women and new mothers $p = 1.00$. There was also no main effect of stereotype threat condition, $F(1, 390) = 0.43$, $p = .514$, $\eta p^2 = 0.001$.

A moderation analysis using Haye's PROCESS Macro v3.4 on SPSS was used to assess whether cognitive identification moderated the effects of group on mathematics performance, as this showed some significant effects of group on performance. The moderation model, which included group as the independent variable, cognitive identification as a moderator, and performance on the mathematics task as the outcome, explained 3.72% of overall variance. The overall model was significant, $F(3, 399) = 5.14$, $p = .0017$. The effect of group on performance was not significant ($\beta = -0.55$, $SE = 0.85$, $p = .52$), nor was cognitive identification ($\beta = 0.46$, $SE = 0.53$, $p = .39$. The interaction term in this model was also not statistically significant, $\beta = -0.049$, $SE = 0.23$, $p = .83$, and the addition of the interaction term

did not add a significant amount of variance (R^2 change = 0.0001) $F(1, 399) = 0.045$, $p = .83$). Therefore, moderation did not occur.

Finally, a 2 × 3 ANOVA with performance on the Levels of Processing task revealed no significant main effects for either group, $F(2, 390) = 0.57$, $p = .57$, $\eta p^2 = 0.004$ or condition $F(1, 390) = 0.003$, $p = .96$, $\eta p^2 = 0.00$ and no interaction, $F(2, 390) = 0.25$, $p = .78$, $\eta p^2 = 0.001$.

3.2.2 | Exploratory results

As an exploratory analysis, we investigated whether endorsement of the "baby brain" stereotype was affected by group and condition. This is due to evidence which suggests that endorsement of the stereotype may affect susceptibility to stereotype threat (Schmader et al., 2004). As stereotype endorsement was a categorical outcome (yes, no, unsure), we conducted two Pearson Chi-Squared tests. This revealed a significant effect of group X^2 (4, $N = 397$) = 57.79 $p < .001$, in that pregnant women and new mothers were more likely to endorse the stereotype, whereas never-pregnant controls were more likely to refute it. There was also a significant effect of condition X^2 (2, $N = 397$) = 9.22 $p = .01$, in that those in the stereotype threat condition endorsed the stereotype more than those in the control condition.

We also tested whether pregnant women and new mothers in the stereotype threat condition were generally more aware of the "baby brain" stereotype (by way of answering "yes" to the "are you familiar with the term 'baby brain'?" as part of the threat manipulation). This revealed a significant association, X^2 (2, $N = 203$) = 44.4 $p < .001$, in which pregnant women (90.1%) and new mothers (91.4%) were more likely to answer yes, compared with never pregnant controls (48.8%).

3.3 | Study 2 discussion

To summarize, Study 2 demonstrated a relatively similar pattern of results to Study 1. Activation of the "baby brain" stereotype threat manipulation did not impact women's cognitive performance in the way that we hypothesized. However, given that there is a growing body of literature scrutinizing the social basis to baby brain (e.g., Crawley et al., 2008; Hurt, 2011; Pownall, 2019), coupled with the pervasive research that points to the stereotyping that women face in their pregnancy and into motherhood (Ganong & Coleman, 1995; Halpert et al., 1993; Shields & Cooper, 1983), this should be scrutinized further before any concrete conclusions are drawn.

This study also suggests that pregnant women and new mothers are more aware of the "baby brain" stereotype, which suggests that the stereotype is more widely known to groups that are targeted by it. This is perhaps unlike other stereotypes, such as the "women are poorer at maths" stereotypes, which are more universal and widely recognized. Future work into the baby brain phenomenon may wish to explore this further, testing the extent to which the stereotype is widely endorsed by different groups and contexts.

4 | GENERAL DISCUSSION

This work aimed to empirically assess whether pregnant women's cognitive performance is worsened with a stereotype threat manipulation containing information about the "baby brain" stereotype. We tested this research question across two studies and found that generally, there was not support for the hypothesis that stereotype threat worsens pregnant women's performance in this specific testing context. A number of interesting findings did emerge. For example, Study 1 demonstrated that pregnant women generally underperformed in memory-based tasks compared with new mothers and female controls. However, counter to hypotheses, the presence of a negative performance stereotype did not significantly impact memory performance. There was also no interaction between the stereotype threat condition and group, which was not aligned with our original hypotheses. In Study 2, we aimed to replicate the design and procedure of Study 1, but this time focusing on a more diverse battery of tests that tap into other cognitive constructs under the "baby brain" stereotype. Generally, we found a lack of stereotype threat and group effects across the tasks. Surprisingly, however, our results indicate that new mothers and pregnant women performed better than controls in the mathematics task.

The lack of significant stereotype threat effects in our results may be due to a number of factors, including important moderators (e.g., Shapiro & Neuberg, 2007) and mediators (Pennington et al., 2016). Despite some work that considers how stereotype threat may be an appropriate framework to consider pregnancy-related behaviors (e.g., prevalence of workplace accidents in pregnancy; Lavaysse & Probst, 2020), this is the first study that assesses the capacity of stereotype threat to explain cognitive decline in pregnancy. It is also important to diversify the outcome variable of future stereotype threat work in this area. Immediate memory recall, as used in these two experimental studies here, is only one of the various cognitive areas thought to be associated with cognitive decline in pregnancy and new motherhood (Davies et al., 2018). Other areas for future study include more diverse outcome variables, such as attention and executive functioning, given that Davies et al. (2018) suggested that these capacities are also affected by pregnancy. Thus, more facets of "cognitive ability" must be investigated fully in future studies.

The results of Study 1 result may also be due to the level of executive processing required to perform well in each of the three tasks used in this experimental work. As Onyper et al. (2010) suggest, in a study assessing executive functioning of pregnant women with matched controls, discrepancies in previous cognitive experiments may be because the measures that are used target different executive components. For example, the executive demands of the Flanker Task in this study, which requires information to be retained during a distraction, are higher compared with the mathematics task, which may be answered with pre-potent and learned responding (Davies et al., 2016). As Hutter et al. (2019) explain, tasks which activate dominant and learned responses may lead to performance facilitation effects, despite the existence of a stereotype threat. For example, some "horizontal math problems", such as the ones used

in this work, activate a "method of adjustment" prepotent response (Seitchik & Harkins, 2015). This reliance on a dominant, learned response can be useful for solving the problem effectively (Jamieson & Harkins, 2007). This may be why the mathematics task differed from the other tasks in terms of performance outcome in this study.

There are, however, some emerging group differences in cognitive ability, irrespective of stereotype threat condition (Study 1). To fully test the extent to which stereotype threat plays a role, there may be important situational and contextual factors that must be present in order to elicit stereotype threat effects. As Steele et al. (2002) explain, stereotype threat effects likely differ across testing contexts, due to the presence of different performance-relevant moderators and mediators. One potentially crucial factor, as many previous researchers have proposed, is the perceived *consequences* of test performance. For example, research has demonstrated that stereotype threat effects occur more prominently in "high stakes" testing contexts (Fryer et al., 2008; McFarland et al., 2003; Sackett et al., 2004) Therefore, our lack of stereotype threat effects in the previous two studies may be due to the low stakes, and thus low motivation, nature of the testing contexts. If participants do not exert sufficient effort in the task at hand, or indeed are not sufficiently invested in the outcome of the task, and therefore stressed at the threat of confirming the stereotype (Sherman et al., 2009), stereotype threatening information may likely not affect performance outcomes.

That is, regardless of whether "baby brain" is a stereotype or a product of biological changes, in order for it to elicit stereotype threat effects, the notion of being stereotyped in this way should be sufficiently *threatening*. It is notable, therefore, that the vast majority of new mothers and pregnant participants self-reported cognitive changes throughout their pregnancy (Study 1) and self-reported a significantly higher awareness and endorsement of the "baby brain" stereotype compared with non-pregnant participants (Study 2). This provides some important context to the stereotype threat testing paradigm. If the participants of these studies endorsed the "baby brain" perception, this could mean that the threat of being perceived as conforming to the stereotype was not sufficiently activated, and thus stereotype threat effects did not occur. Similarly, these results may also simply demonstrate that baby brain performance stereotype threat does not impair pregnant women or new mother's cognitive ability, in light of the failed replication concerns of previous studies in this work (Flore & Wicherts, 2015). While there are many situational, contextual and personal factors at play in stereotype threat effects, it should not be overlooked that stereotype threat is not compatible with this groups' performance.

As Jamieson and Harkins (2007) stress, stereotype threat occurs when *motivation* to disprove a negative stereotype outweighs ability to perform in the targeted domain. This "mere effort" approach posits that stereotype threat increases performance motivation, by way of disproving a negative salient stereotype, which increases reliance on dominant *pre-potent* (i.e., dominant and learned; Jamieson & Harkins, 2007, 2009; McFall et al., 2009) responding. In other words, when under stereotype threat, the goal to overcome the performance stereotype potentiates a well-learned and habitual response (Grandjean & Collette, 2011; McFall et al., 2009). Therefore, if the performance measure is not conducive to reliance on pre-potent responding, participants under a stereotype threat underperform; if it is, this leads to performance facilitation effects (Davies et al., 2016; Hutter et al., 2019). This motivational account of stereotype threat also echoes the regulatory focus theory (Higgins, 1998), which suggests that self-regulation is either focused on different individual needs. These motivational principles can be applied to stereotype threat theory, by manipulating the presence of negative outcomes, which changes the emotional relationship with the performance outcome (Keller & Dauenheimer, 2003).

Finally, it is worth acknowledging that the present studies were all conducted online. There is evidence to suggest that testing context (i.e., online versus in the lab) does not impact performance on behavioral tasks. For example, Casler et al. (2013) compared the performance of participants recruited online with in-lab participants on a behavioral task. The online participants conducted an adapted version of the in-lab behavioral task. Overall, the authors concluded that online adaptation was highly effective and responses across the two testing conditions were equivalent. This paper also notes how online testing, such as via MTurk (a US-based recruitment service equivalent to Prolific), allow researchers to recruit more diverse samples with a significantly higher age compared with the typical sample of in-person testing. Therefore, on a logistical level, online data collection is a useful tool, particularly when recruiting niche samples such as those in our study. There are, however, some theoretical considerations of online testing of stereotype threat theory. Online testing means that predominantly blatant stereotype threat manipulations are used, as opposed to more subtle threat cues, for example, gender of experimenter in-lab testing contexts (Stone & McWhinnie, 2008). This allows us to have more control, in an experimental sense, over the types of primes that the study is manipulating.

However, online testing is also inherently less controlled than in-person testing; therefore, factors such as mind wandering or distraction may also affect results in this work (Schuster et al., 2015). Therefore, we suggest that the reappraisal of stereotype threat theory as an explanation for performance deficits in stereotyped groups should also continue to test the effects of threats in varied testing conditions. There is evidence for stereotype threat effects in various field settings, for example in the classroom (Huguet & Régner, 2007; Stricker & Ward, 2004) and the workplace (Neal-Jackson, 2020). Future work should now consider whether different manipulations of stereotype threat can be found in online testing sessions, to contribute towards the reappraisal of stereotype threat as a robust theory in light of publication bias concerns. Also, while online testing is not "the field", this may go some way in addressing the tendency for social psychological theories to be based solely from "decontextualized laboratory data" that does not reflect more "real world" contexts (Berkman & Wilson, 2021, p. 1).

On a final note, one simpler explanation for the null effects in this study may be the robustness of stereotype threat theory as a mechanism of explaining how stereotypes impact performance. These null effects may not be surprising, given the widespread

concerns about stereotype threat theory's replicability and robustness (e.g., Flore & Wicherts, 2015; Stoet & Geary, 2012). Indeed, while there are undoubtedly various situational, theoretical, and contextual factors that should be further examined in the context of a "baby brain" related stereotype threat, the present study also contributes to the ongoing questioning of stereotype threat theory's theoretical utility and methodological robustness. This work has reappraised stereotype threat theory, by investigating how it can provide insights into diverse types of stereotyping to different groups, namely the "baby brain" stereotype. However, given the null effects, this may further question the validity of stereotype threat theory as a theoretical framework. Other scholars interested in providing social psychological explanations to the "baby brain" stereotype may wish to look beyond the stereotype threat paradigm and focus on more replicable theories that relate social expectations to cognitive performance, such as objectification theory (Fredrickson & Roberts, 1997).

This work does not allow us to test fully whether the concept of "baby brain" is a product of stereotyping alone. It does, however, allow us to begin to reassess how stereotyping may contribute to the perception of pregnant women and new mothers having "baby brain" and question 'stereotype threat' as an explanation for this effect. Future work should extend enquiries investigating the utility of stereotype threat as an explanatory theory. For example, future work should manipulate important factors, such as participant's *motivation* to disprove the stereotype, the threatening nature of the stereotype, and the real-world relevance of the testing condition. There may also be scope to experimentally test other explanations for "baby brain" effects, such as the role that objectification plays on cognitive functioning (Winn & Cornelius, 2020) or the qualitative experiences of being subject to a baby brain stereotype. Taken together, this research advances understanding of social mechanisms that underpin seemingly biological effects of cognitive decline in pregnancy and provides another replication attempt of stereotype threat theory as a theoretical framework.

ACKNOWLEDGEMENTS

We would like to thank the following people for their assistance with data collection and data entry: Sheli Algranti, Millie Morgan, Anya Shore, Alice Gilman, Bethan Newton, Ruby Downey-Brown, Harriet Dooley, Jude Paul, and Victoria Morgan.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

ETHICAL APPROVAL

Ethical approval was obtained for both studies from the local School of Psychology Ethics Committee.

DATA AVAILABILITY STATEMENT

Data and materials are available in the Open Science Framework page: https://osf.io/y2m53/.

ORCID

Madeleine Pownall https://orcid.org/0000-0002-3734-8006

REFERENCES

Ambady, N., Paik, S. K., Steele, J., Owen-Smith, A., & Mitchell, J. P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology*, *40*(3), 401–408. https://doi.org/10.1016/j.jesp.2003.08.003

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2019). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Appel, M., Kronberger, N., & Aronson, J. (2011). Stereotype threat impairs ability building: Effects on test preparation among women in science and technology. *European Journal of Social Psychology*, *41*(7), 904–913. https://doi.org/10.1002/ejsp.835

Bargad, A., & Hyde, J. S. (1991). Women's studies: A study of feminist identity development in women. *Psychology of Women Quarterly*, *15*(2), 181–201. https://doi.org/10.1111/j.1471-6402.1991.tb00791.x

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*(2), 256. https://doi.org/10.1037/0096-3445.136.2.256

Berkman, E. T., & Wilson, S. M. (2021). So useful as a good theory? The practicality crisis in (Social) psychological theory. *Perspectives on Psychological Science*, 1–11. https://doi.org/10.1177/1745691620969650

Bleier, R. (1978). Bias in biological and human sciences: Some comments. *Signs: Journal of Women in Culture and Society*, *4*(1), 159–162. https://doi.org/10.1086/493577

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/10.1016/S0022-1031(03)00099-4

Bragger, J. D., Kutcher, E., & Morgan, J. (2002). The Effects of the structured interview on reducing biases against pregnant job applicants. *Sex Roles*, *46*, 215–226. https://doi.org/10.1023/A:1019967231059

Brett, M., & Baxendale, S. (2001). Motherhood and memory a review. *Psychoneuroendocrinology*, *26*(4), 339–362. https://doi.org/10.1016/S0306-4530(01)00003-8

Brindle, P. M., Brown, M. W., Brown, J., Griffith, H. B., & Turner, G. M. (1991). Objective and subjective memory impairment in pregnancy. *Psychological Medicine*, *21*(3), 647–653. https://doi.org/10.1017/S0033291700022285

Brizendine, L. (2007). *The female brain*. Harmony.

Casey, P., Huntsdale, C., Angus, G., & Janes, C. (1999). Memory in pregnancy. II: Implicit, incidental, explicit, semantic, short-term, working and prospective memory in primigravid, multigravid and postpartum women. *Journal of Psychosomatic Obstetrics & Gynecology*, *20*(3), 158–164. https://doi.org/10.3109/01674829909075590

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156–2160. https://doi.org/10.1016/j.chb.2013.05.009

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology*, *112*, 1297–1339. https://doi.org/10.1086/511799

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Crandall, C. S., & Sherman, J. W. (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of*

Experimental Social Psychology, *66*, 93–99. https://doi.org/10.1016/j.jesp.2015.10.002

Crawley, R., Grant, S., & Hinshaw, K. (2008). Cognitive changes in pregnancy: Mild decline or societal stereotype? *Applied Cognitive Psychology*, *22*(8), 1142–1162. https://doi.org/10.1002/acp.1427

Cuttler, C., Graf, P., Pawluski, J. L., & Galea, L. A. (2011). Everyday life memory deficits in pregnant women. *Canadian Journal of Experimental Psychology*, *65*(1), 27. https://doi.org/10.1037/a0022844

Davies, L. C., Conner, M., Sedikides, C., & Hutter, R. R. (2016). Math question type and stereotype threat: Evidence from educational settings. *Social Cognition*, *34*(3), 196–216. https://doi.org/10.1521/soco.2016.34.3.196

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, *28*(12), 1615–1628. https://doi.org/10.1177/014616702237644

Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology*, *88*(2), 276–287. https://doi.org/10.1037/0022-3514.88.2.276

Davies, S. J., Lum, J. A., Skouteris, H., Byrne, L. K., & Hayden, M. J. (2018). Cognitive impairment during pregnancy: A meta-analysis. *Medical Journal of Australia*, *208*(1), 35–40. https://doi.org/10.5694/mja17.00131

de Groot, R. H., Vuurman, E. F., Hornstra, G., & Jolles, J. (2006). Differences in cognitive performance during pregnancy and early motherhood. *Psychological Medicine*, *36*(7), 1023–1032. https://doi.org/10.1017/S0033291706007380

Doyle, R. A., & Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences*, *47*, 103–116. https://doi.org/10.1016/j.lindif.2015.12.018

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, *25*, 249–263. https://doi.org/10.3758/BF03198804

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*(2), 175–191. https://doi.org/10.3758/BF03193146

Fine, C. (2008). Will working mothers' brains explode? The popular new genre of neurosexism. *Neuroethics*, *1*(1), 69–72. https://doi.org/10.1007/s12152-007-9004-2

Finkel, D., Reynolds, C. A., McArdle, J. J., & Pedersen, N. L. (2005). The longitudinal relationship between processing speed and cognitive ability: Genetic and environmental influences. *Behavior Genetics*, *35*(5), 535–549. https://doi.org/10.1007/s10519-005-3281-5

Finnigan, K. M., & Corker, K. S. (2016). Do performance avoidance goals moderate the effect of different types of stereotype threat on women's math performance? *Journal of Research in Personality*, *63*, 36–43. https://doi.org/10.1016/j.jrp.2016.05.009

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878. https://doi.org/10.1037/0022-3514.82.6.878

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. https://doi.org/10.1016/j.jsp.2014.10.002

Forscher, P. S., Taylor, V. J., Cavagnaro, D., Lewis, N. A. Jr, Buchanan, E. M., Moshontz, H., Mark, A. Y., Appleby, S., Batres, C., & Bennett-Day, B., & Chartier, C. R. (2019, July 17). A multi-site examination of stereotype threat in black college students across varying operationalizations. https://doi.org/10.31234/osf.io/6hju9

Fox, A. B., & Quinn, D. M. (2015). Pregnant women at work: The role of stigma in predicting women's intended exit from the workforce. *Psychology of Women Quarterly*, *39*, 226–242. https://doi.org/10.1177/0361684314552653

Fredrickson, B. L., & Roberts, T. A. (1997). Objectification theory: Toward understanding women's lived experiences and mental health risks. *Psychology of Women Quarterly*, *21*(2), 173–206. https://doi.org/10.1111/j.1471-6402.1997.tb00108.x

Fryer, R. G., Jr., Levitt, S. D., & List, J. A. (2008). Exploring the impact of financial incentives on stereotype threat: Evidence from a pilot study. *American Economic Review*, *98*(2), 370–375. https://doi.org/10.1257/aer.98.2.370

Ganley, C. M., Mingle, L. A., Ryan, A. M., Ryan, K., Vasilyeva, M., & Perry, M. (2013). An examination of stereotype threat effects on girls' mathematics performance. *Developmental Psychology*, *49*(10), 1886. https://doi.org/10.1037/a0031412

Ganong, L. H., & Coleman, M. (1995). The content of mother stereotypes. *Sex Roles*, *32*(7–8), 495–512. https://doi.org/10.1007/BF01544185

Glick, P., & Fiske, S. T. (1997). Hostile and benevolent sexism: Measuring ambivalent sexist attitudes toward women. *Psychology of Women Quarterly*, *21*(1), 119–135. https://doi.org/10.1111/j.1471-6402.1997.tb00104.x

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin*, *28*(5), 659–670. https://doi.org/10.1177/0146167202288010

Grandjean, J., & Collette, F. (2011). Influence of response prepotency strength, general working memory resources, and specific working memory load on the ability to inhibit predominant responses: A comparison of young and elderly participants. *Brain and Cognition*, *77*, 237–247. https://doi.org/10.1016/j.bandc.2011.08.004

Green, J. M., Kitzinger, J. V., & Coupland, V. A. (1990). Stereotypes of childbearing women: A look at some evidence. *Midwifery*, *6*(4), 232. https://doi.org/10.1016/s0266-6138(05)80122-6

Hackney, K. J., Daniels, S. R., Paustian-Underdahl, S. C., Perrewé, P. L., Mandeville, A., & Eaton, A. A. (2020). Examining the effects of perceived pregnancy discrimination on mother and baby health. *Journal of Applied Psychology*. Advance Online Publication. https://doi.org/10.1037/apl0000788

Halpert, J. A., Wilson, M. L., & Hickman, J. L. (1993). Pregnancy as a source of bias in performance appraisals. *Journal of Organizational Behavior*, *14*(7), 649–663. https://doi.org/10.1002/job.4030140704

Hebl, M. R., King, E. B., Glick, P., Singletary, S. L., & Kazama, S. (2007). Hostile and benevolent reactions toward pregnant women: Complementary interpersonal punishments and rewards that maintain traditional roles. *Journal of Applied Psychology*, *92*(6), 1499–1511. https://doi.org/10.1037/0021-9010.92.6.1499

Henry, J. D., & Rendell, P. G. (2007). A review of the impact of pregnancy on memory function. *Journal of Clinical and Experimental Neuropsychology*, *29*(8), 793–803. https://doi.org/10.1080/13803390701612209

Henry, J. F., & Sherwin, B. B. (2012). Hormones and cognitive functioning during late pregnancy and postpartum: A longitudinal study. *Behavioral Neuroscience*, *126*(1), 73. https://doi.org/10.1037/a0025540

Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype threat on age differences in memory performance. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58*(1), 3–11. https://doi.org/10.1093/geronb/58.1.P3

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. *Advances in Experimental Social Psychology*, *30*, 1–46.

Hoekzema, E., Barba-Müller, E., Pozzobon, C., Picado, M., Lucco, F., García-García, D., Soliva, J. C., Tobeña, A., Desco, M., Crone, E. A., Ballesteros, A., Carmona, S., & Vilarroya, O. (2017). Pregnancy leads to long-lasting changes in human brain structure. *Nature Neuroscience*, *20*(2), 287–296. https://doi.org/10.1038/nn.4458

Huguet, P., & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology*, *99*(3), 545–560. https://doi.org/10.1037/0022-0663.99.3.545

Hurt, N. E. (2011). Legitimizing "baby brain": Tracing a rhetoric of significance through science and the mass media. *Communication and Critical/Cultural Studies*, *8*(4), 376–398. https://doi.org/10.1080/14791420.2011.619202

Hutter, R. R., Davies, L. C., Sedikides, C., & Conner, M. (2019). Women's stereotype threat-based performance motivation and prepotent inhibitory ability. *British Journal of Social Psychology*, *58*(3), 691–713. https://doi.org/10.1111/bjso.12298

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

Jamieson, J. P., & Harkins, S. G. (2009). The effect of stereotype threat on the solving of quantitative GRE problems: A mere effort interpretation. *Personality and Social Psychology Bulletin*, *35*(10), 1301–1314. https://doi.org/10.1177/0146167209335165

Johnson, T. D. (2008). *Maternity leave and employment: Patterns of first-time mothers 1961–2003*. United States Census Bureau.

Jones, K. P. (2017). To tell or not to tell? Examining the role of discrimination in the pregnancy disclosure process at work. *Journal of Occupational Health Psychology*, *22*, 239–250. https://doi.org/10.1037/ocp0000030

Jurs, B. S., Weber, H. E., Bowen, M. E., Hazelwood, W. H., Cooksey, K. D., & Katzarska-Miller, I. (2019). No longer a threat: A failure to evoke stereotype threat using the race IAT. *Journal of Articles in Support of the Null Hypothesis*, *15*(2), 121–128.

Kaye, L. K., Pennington, C. R., & McCann, J. J. (2018). Do casual gaming environments evoke stereotype threat? Examining the effects of explicit priming and avatar gender. *Computers in Human Behavior*, *78*, 142–150. https://doi.org/10.1016/j.chb.2017.09.031

Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin*, *29*(3), 371–381. https://doi.org/10.1177/0146167202250218

Kitroeff, N., & Silver-Greenberg, J. (2019, February 8). *Pregnancy discrimination is rampant inside America's biggest companies*. The New York Times. https://www.nytimes.com/interactive/2018/06/15/business/pregnancy-discrimination.html

Koenig, A. M., & Eagly, A. H. (2005). Stereotype threat in men on a test of social sensitivity. *Sex Roles*, *52*(7–8), 489–496. https://doi.org/10.1007/s11199-005-3714-x

Lavaysse, L. M., & Probst, T. M. (2020). Pregnancy and workplace accidents: The impact of stereotype threat. *Work & Stress*, *35*(1), 1–17. https://doi.org/10.1080/02678373.2020.1774937

Lesko, A. C., & Corpus, J. H. (2006). Discounting the difficult: How high math-identified women respond to stereotype threat. *Sex Roles*, *54*(1–2), 113–125. https://doi.org/10.1007/s11199-005-8873-2

Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology*, *71*(6), 1092. https://doi.org/10.1037/0022-3514.71.6.1092

Lewis, N. Jr, & Michalak, N. M. (2019). *Has stereotype threat dissipated over time? A cross-temporal meta-analysis*. Pre-Print, PsyArXiv.

Little, L. M., Hinojosa, A. S., Paustian-Underdahl, S., & Zipay, K. P. (2018). Managing the harmful effects of unsupportive organizations during pregnancy. *Journal of Applied Psychology*, *103*(6), 631–643. https://doi.org/10.1037/apl0000285

Little, L., Major, V., Hinojosa, A., & Nelson, D. (2015). Professional image maintenance: How women navigate pregnancy in the workplace. *Academy of Management Journal*, *58*(1), 8–37. https://doi.org/10.5465/amj.2013.0599

Longhurst, R. (1997). 'Going Nuts': Re-presenting pregnant women. *New Zealand Geographer*, *53*(2), 34–39. https://doi.org/10.1111/j.1745-7939.1997.tb00497.x

Longhurst, R. (1999). 'Giving' advice to pregnant women. In E. Teather (Ed.) *Embodied geographies* (pp. 77–89). London: Routledge.

Major, B., Spencer, S., Schmader, T., Wolfe, C., & Crocker, J. (1998). Coping with negative stereotypes about intellectual performance: The role of psychological disengagement. *Personality and Social Psychology Bulletin*, *24*(1), 34–50. https://doi.org/10.1177/0146167298241003

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/10.1016/j.jesp.2005.04.010

Masser, B., Grass, K., & Nesic, M. (2007). 'We like you, but we don't want you'—The impact of pregnancy in the workplace. *Sex Roles*, *57*(9–10), 703–712. https://doi.org/10.1007/s11199-007-9305-2

McDowall, J., & Moriarty, R. (2000). Implicit and explicit memory in pregnant women: An analysis of data-driven and conceptually driven processes. *The Quarterly Journal of Experimental Psychology Section A*, *53*(3), 729–740. https://doi.org/10.1080/713755904

McFall, S. R., Jamieson, J. P., & Harkins, S. G. (2009). Testing the mere effort account of the evaluation-performance relationship. *Journal of Personality and Social Psychology*, *96*(1), 135–154. https://doi.org/10.1037/a0012878

McFarland, L. A., Lev-Arey, D. M., & Ziegert, J. C. (2003). An examination of stereotype threat in a motivational context. *Human Performance*, *16*(3), 181–205. https://doi.org/10.1207/S15327043HUP1603_2

McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology*, *39*(1), 83–90. https://doi.org/10.1016/S0022-1031(02)00513-9

Moè, A., & Pazzaglia, F. (2006). Following the instructions!: Effects of gender beliefs in mental rotation. *Learning and Individual Differences*, *16*(4), 369–377. https://doi.org/10.1016/j.lindif.2007.01.002

Moè, A., & Pazzaglia, F. (2010). Beyond genetics in mental rotation test performance: The power of effort attribution. *Learning and Individual Differences*, *20*(5), 464–468. https://doi.org/10.1016/j.lindif.2010.03.004

Morgan, W. B., Walker, S. S., Hebl, M. M. R., & King, E. B. (2013). A field experiment: Reducing interpersonal discrimination toward pregnant job applicants. *Journal of Applied Psychology*, *98*, 799–809. https://doi.org/10.1037/a0034040

Neal-Jackson, A. (2020). "Well, What Did You Expect?": Black women facing stereotype threat in collaborative academic spaces at a predominantly white institution. *Journal of College Student Development*, *61*(3), 317–332. https://doi.org/10.1353/csd.2020.0030

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*(1), 44–59. https://doi.org/10.1037/0022-3514.83.1.44

Onyper, S. V., Searleman, A., Thacher, P. V., Maine, E. E., & Johnson, A. G. (2010). Executive functioning and general cognitive ability in pregnant women and matched controls. *Journal of Clinical and Experimental Neuropsychology*, *32*(9), 986–995. https://doi.org/10.1080/13803391003662694

Pennington, C. R., Heim, D., Levy, A. R., & Larkin, D. T. (2016). Twenty years of stereotype threat research: A review of psychological mediators. *PLoS One*, *11*(1), e0146487. https://doi.org/10.1371/journal.pone.0146487

Pennington, C. R., Kaye, L. K., & McCann, J. J. (2018). Applying the multi-threat framework of stereotype threat in the context of digital gaming. *PLoS One*, *13*(2), e0192137. https://doi.org/10.1371/journal.pone.0192137

Pennington, C. R., Litchfield, D., McLatchie, N., & Heim, D. (2019). Stereotype threat may not impact women's inhibitory control or mathematical performance: Providing support for the null hypothesis. *European Journal of Social Psychology*, *49*(4), 717–734. https://doi.org/10.1002/ejsp.2540

Pownall, M. (2019). The "baby brain" stereotype and policing of pregnant women's competence. *Feminist Media Studies*, *19*(5), 759–763. https://doi.org/10.1080/14680777.2019.1630917

Pronin, E., Steele, C. M., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology*, *40*(2), 152–168. https://doi.org/10.1016/S0022-1031(03)00088-X

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Raz, S. (2014). Behavioral and neural correlates of cognitive–affective function during late pregnancy: An event-related potentials study. *Behavioural Brain Research*, *267*, 17–25. https://doi.org/10.1016/j.bbr.2014.03.021

Rydell, R. J., Van Loo, K. J., & Boucher, K. L. (2014). Stereotype threat and executive functions: Which functions mediate different threat-related outcomes? *Personality and Social Psychology Bulletin*, *40*, 377–390. https://doi.org/10.1177/0146167213513475

Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On interpreting stereotype threat as accounting for African American-White differences on cognitive tests. *American Psychologist*, *59*(1), 7–13. https://doi.org/10.1037/0003-066X.59.1.7

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*(3), 440–452. https://doi.org/10.1037/0022-3514.85.3.440

Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles*, *50*(11), 835–850. https://doi.org/10.1023/B:SERS.0000029101.74557.a0

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*(2), 336–356. https://doi.org/10.1037/0033-295X.115.2.336

Schneider, Z. (1989). Cognitive performance in pregnancy. *The Australian Journal of Advanced Nursing: A Quarterly Publication of the Royal Australian Nursing Federation*, *6*(3), 40–47.

Schuster, C., Martiny, S. E., & Schmader, T. (2015). Distracted by the unthought–Suppression and reappraisal of mind wandering under stereotype threat. *PLoS One*, *10*(3), e0122207. https://doi.org/10.1371/journal.pone.0122207

Seitchik, A. E., & Harkins, S. G. (2015). Stereotype threat, mental arithmetic, and the mere effort account. *Journal of Experimental Social Psychology*, *61*, 19–30. https://doi.org/10.1016/j.jesp.2015.06.006

Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology*, *39*, 68–74. https://doi.org/10.1016/S0022-1031(02)00508-5

Shahvisi, A. (2020). Nesting behaviours during pregnancy: Biological instinct, or another way of gendering housework? *Women's Studies International Forum*, *78*, 102329. https://doi.org/10.1016/j.wsif.2019.102329

Shapiro, J. R., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, *11*, 107–130. https://doi.org/10.1177/1088868306294790

Sharp, K., Brindle, P. M., Brown, M. W., & Turner, G. M. (1993). Memory loss during pregnancy. *BJOG: An International Journal of Obstetrics & Gynaecology*, *100*(3), 209–215. https://doi.org/10.1111/j.1471-0528.1993.tb15232.x

Sherman, D. K., Bunyan, D. P., Creswell, J. D., & Jaremka, L. M. (2009). Psychological vulnerability and stress: The effects of self-affirmation on sympathetic nervous system responses to naturalistic stressors. *Health Psychology*, *28*(5), 554. https://doi.org/10.1037/a0014663

Shields, S. A., & Cooper, P. E. (1983). Stereotypes of traditional and nontraditional childbearing roles. *Sex Roles*, *9*(3), 363–376. https://doi.org/10.1007/BF00289671

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*(1), 80–83. https://doi.org/10.1111/1467-9280.00111

Smith, J. L., & White, P. H. (2002). An examination of implicitly activated explicitly activated, and nullified stereotypes on mathematical performance: It's not just a woman's issue. *Sex Roles*, *47*(3–4), 179–191. https://doi.org/10.1023/A:1021051223441

Spencer, B., & Castano, E. (2007). Social class is dead. Long live social class! Stereotype threat among low socioeconomic status individuals. *Social Justice Research*, *20*(4), 418–432. https://doi.org/10.1007/s11211-007-0047-7

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. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613–629. https://doi.org/10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797. https://doi.org/10.1037/0022-3514.69.5.797

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, *34*, 379–440. https://doi.org/10.1016/s0065-2601(02)80009-0

Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology*, *16*(1), 93–102. https://doi.org/10.1037/a0026617

Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology*, *44*(2), 445–452. https://doi.org/10.1016/j.jesp.2007.02.006

Stricker, L. J., & Ward, W. C. (2004). Stereotype threat, inquiring about test takers' ethnicity and gender, and standardized test performance 1. *Journal of Applied Social Psychology*, *34*(4), 665–693. https://doi.org/10.1111/j.1559-1816.2004.tb02564.x

Sutton, R. M., Douglas, K. M., & McClellan, L. M. (2011). Benevolent sexism, perceived health risks, and the inclination to restrict pregnant women's freedoms. *Sex Roles*, *65*(7–8), 596–605. https://doi.org/10.1007/s11199-010-9869-0

Wakefield, J. R., Hopkins, N., & Greenwood, R. M. (2012). Thanks, but no thanks: Women's avoidance of help-seeking in the context of a dependency-related stereotype. *Psychology of Women Quarterly*, *36*(4), 423–431. https://doi.org/10.1177/0361684312457659

Winn, L., & Cornelius, R. (2020). Self-objectification and cognitive performance: A systematic review of the literature. *Frontiers in Psychology*, *11*, 20. https://doi.org/10.3389/fpsyg.2020.00020

Woodcock, A., Hernandez, P. R., Estrada, M., & Schultz, P. (2012). The consequences of chronic stereotype threat: Domain disidentification and abandonment. *Journal of Personality and Social Psychology*, *103*(4), 635–646. https://doi.org/10.1037/a0029120

Zigerell, L. J. (2017). Potential publication bias in the stereotype threat literature: Comment on Nguyen and Ryan (2008). *Journal of Applied Psychology*, *102*(8), 1159–1168. https://doi.org/10.1037/apl0000188

How to cite this article: Pownall M, Conner M, Hutter RRC. The effects of activating a "baby brain" stereotype on pregnant women's cognitive functioning. *J Appl Soc Psychol*. 2021;51:809–824. https://doi.org/10.1111/jasp.12802