

Colby



Colby College
Digital Commons @ Colby

Honors Theses

Student Research

2023

Stereotype Threat in Older Adults' Episodic Memory Tests: Susceptibility and Protection

Talia Barrett
Colby College

Follow this and additional works at: <https://digitalcommons.colby.edu/honorstheses>



Part of the [Cognitive Science Commons](#)

Colby College theses are protected by copyright. They may be viewed or downloaded from this site for the purposes of research and scholarship. Reproduction or distribution for commercial purposes is prohibited without written permission of the author.

Recommended Citation

Barrett, Talia, "Stereotype Threat in Older Adults' Episodic Memory Tests: Susceptibility and Protection" (2023). *Honors Theses*. Paper 1402.
<https://digitalcommons.colby.edu/honorstheses/1402>

This Honors Thesis (Open Access) is brought to you for free and open access by the Student Research at Digital Commons @ Colby. It has been accepted for inclusion in Honors Theses by an authorized administrator of Digital Commons @ Colby.

Stereotype Threat in Older Adults' Episodic Memory Tests: Susceptibility and Protection

An Honors Thesis

Presented to

The Faculty of the Department of Psychology

Colby College

In partial fulfillment of the requirements for the

Degree of Bachelor of Arts

By

Talia E. Barrett

Waterville, Maine

May 15, 2023

Abstract

The objective of the present study was to investigate whether prior task success would protect older adults against stereotype threat in an episodic memory test. Previous experiments have established that, whereas stereotype threat negatively impacts older participants' episodic memory performance, prior task success benefits it. However, up until this point, researchers had yet to combine the two manipulations to test their joint effect on episodic memory. Participants were randomly assigned to read a stereotype threat or neutral passage, after which they were placed in the success or no success group. Participants next received a stereotype condition reminder, and they then completed a memory encoding and free recall task. To conclude the experiment, the older adults filled out a manipulation check and responded to an open-ended threat reaction question, and they completed some measures of individual differences. Participants' recall accuracy did not differ as a function of stereotype condition or prior task success group, indicating that success did not buffer threatened participants against the stereotype manipulation. Intrusion rate, however, varied based on stereotype condition: Threatened participants made significantly fewer intrusions than neutral participants, which is consistent with the regulatory fit hypothesis of stereotype threat. The Discussion includes an analysis of these results in the context of previous research. Moreover, interpretation of the nonsignificant findings in light of manipulation check and open-ended threat reaction data points toward a potential avenue of future research to examine the connection between internalized metamemory beliefs and susceptibility to old age-based stereotype threat.

Acknowledgements

First and foremost, thank you so much to my thesis advisor, Professor Jen Coane, for guiding me through this year-long project, answering my many (and frequent) questions, and providing me with moral support during the stressful parts. Your academic and life mentorship – both throughout this project and throughout all of the semesters I have known you – has helped me as a Colby psychology student more than I can say.

Thank you to the Memory and Language Lab Research Assistants for your support and encouragement throughout this entire thesis process. In particular, I am grateful to Yan Xuan and Cara Norris, who helped me significantly with recruitment and testing.

Thank you to the entire Colby Psychology Department for providing me with the opportunity to pursue an honors thesis and for contributing to the rigorous and rewarding experience I have had as a major. The research experience and analytical skills I have gained throughout the psychology courses I have taken prepared me to complete this project.

Thank you to my wonderful participants who generously volunteered their time to contribute to my research project; I would have no honors thesis without them. Interacting with and learning from my older adult participants was such a highlight of this project, and the extremely positive experience I've had has provided me with a newfound interest in working with older adults.

Finally, I extend my gratitude to everyone else in my corner – my parents, my twin brother, and my incredible friends. I am grateful for your love and support always, but I am feeling a bit of extra gratitude right now as I wrap up my senior honors thesis and anticipate being a Colby graduate.

Stereotype Threat in Older Adults' Episodic Memory Tests: Susceptibility and Protection

Ideas about memory loss in older age are widespread. Growing up, children often hear about old age entailing unpleasant cognitive changes, such as increased forgetfulness. The media exacerbates this portrayal of age-related memory impairment. Upon searching “old age memory loss memes” on Google, the search results are endless.

Perceptions of old age-related memory decline have an element of truth to them because certain types of memory worsen over the course of the lifespan. Relative to younger adults, older adults (commonly defined as people ages 60+ in aging research) tend to have greater difficulty with *episodic memory*, which refers to people's memory for contextualized life experiences (e.g., what you ate for breakfast yesterday; Blazer et al., 2015; but see Dodson et al., 2007).

Specifically, older adults generally have stronger item memory than associative memory, meaning they struggle more with binding – integrating distinct pieces of a memory trace (e.g., a piece of information and the context in which it was learned) – than they do with remembering the piece of information itself (Loaiza & Srokova, 2020). Furthermore, older adults have greater difficulty with recall (i.e., active memory retrieval) than recognition (i.e., identifying whether information was previously learned) in part because the former involves an increased likelihood of erroneously combining features of a memory trace (i.e., recall entails more binding than recognition; Dodson et al., 2007). Relative to recognition tasks, recall tasks also provide less external support and require greater self-initiated processing (Craik & Byrd, 1982).

Outside of research assessing older adults' objective memory performance, experiments have revealed that aging affects individuals' perception of their own memory processes. Older adults tend to have many complaints about their memories – complaints that are likely exacerbated by widespread stereotypes regarding age-related memory decline. Ponds and Jolles

(1996) compared high- and low-complaint older adults' memory abilities on one episodic task and two semantic (general knowledge) ones. Both groups performed similarly on the memory tests, but the high-complaint older participants demonstrated low memory self-efficacy, whereas the low-complaint group reported higher memory self-efficacy (Ponds & Jolles, 1996). Overall, this suggests that older adults' preoccupations regarding memory loss are not always reflective of true memory abilities.

Many negative age-based stereotypes exist regarding cognitive decline (e.g., Hess et al., 2003). Specifically, stereotypes about older adults' memory troubles are common, and the activation of these stereotypes can negatively impact older participants' performance on memory tests. This effect is broadly referred to as *stereotype threat* in social psychology research: When individuals are aware of a stereotype about a group to which they belong (generally, a stereotype about one group's expected "worse" performance on something relative to a comparison group), they become focused on avoiding confirmation of the stereotype. Consequently, they underperform on a given task in the stereotyped domain (i.e., it becomes a self-fulfilling prophecy), and this underperformance can occur in a variety of contexts involving older adults (Steele & Aronson, 1995). For example, relative to their non-threatened counterparts, stereotype threat adversely impacts older adults' performance of real-world, everyday tasks (e.g., acquisition of motor skills and driving), and threatened older participants also report worse employment outcomes (e.g., lower occupational self-efficacy and greater intentions to resign or retire; Barber, 2020). It is important to acknowledge that the mechanisms underlying stereotype threat depend on the specific aging context, such that stereotype threat effects in one situation (e.g., older adult memory testing) might not generalize to a second context (e.g., driving in old age; Barber, 2017).

With respect to older adults and memory, stereotype threat worsens older participants' episodic memory performance (e.g., Bouazzaoui et al., 2015; Rossi-Arnaud et al., 2018; Swift et al., 2013). Stereotype threat effects in memory tests have been demonstrated repeatedly; older participants often experience stereotype threat, even when age-related stereotypes are not explicitly mentioned in memory tasks (Bouazzaoui et al., 2015; Chasteen et al., 2005). In other words, stereotype threat can be activated explicitly (e.g., by blatantly mentioning that older adults are expected to perform worse relative to younger adults), subtly (e.g., by simply referencing that the task at hand is assessing memory; Desrichard & Köpetz, 2005), and implicitly (e.g., through priming tasks involving negative old age-related stereotypes; Armstrong et al., 2017). Stereotype threat may drain individuals of cognitive resources by inducing a negative mood, divided attention, and stress, thereby resulting in fewer remaining resources to devote to successful episodic memory retrieval. For instance, stereotype threat may lead participants to think negatively about their performance abilities (i.e., it may induce deficit-based expectations – negative thoughts taxing a participant's limited supply of cognitive resources; Haslam et al., 2012). This explanation is generally known as the *executive control hypothesis*, and it is a potential mechanism underlying old age-related stereotype threat's harmful influence on older participants' episodic memory abilities (Fourquet et al., 2020; Mazerolle et al., 2012; Schmader et al., 2008).

In addition to the executive control hypothesis, researchers have proposed the *regulatory fit (or focus) hypothesis* to explain the effect of old age-based stereotype threat on older adults' memory performance. According to this theory, stereotype threat might lead threatened older participants to respond more conservatively on memory tests (i.e., it may induce a prevention-based focus, causing participants to answer questions more carefully; Barber & Mather, 2013;

Fourquet et al., 2020; but see Brubaker & Naveh-Benjamin, 2018). For example, Barber and Mather (2013) found that – relative to the control group – older participants under stereotype threat recalled fewer veridical memories while simultaneously having lower intrusion rates (i.e., they adopted a prevention-based focus; they made fewer mistakes at the expense of recalling lower numbers of words successfully; see also Smith et al., 2017; Wong & Gallo, 2016; but see Schmader et al., 2008; Thomas & Dubois, 2011). In a more recent memory study investigating stereotype threat effects on older adults' brain activity, older participants in the threat condition showed activity in brain regions associated with self-referential thinking and avoidance of error, which is consistent with the regulatory fit hypothesis (Chen et al., 2022).

Albeit distinct theories, some researchers believe that the executive control and regulatory fit hypotheses are not entirely incompatible; rather, they might be best understood as operating along a single developmental continuum. Specifically, whereas the executive control hypothesis might account for younger adults' underperformance in the face of stereotype threat, the regulatory fit hypothesis might better explain performance changes observed among threatened older adults due to increased emotional regulation abilities over the course of the lifespan (Popham & Hess, 2013). Studies testing the executive control and regulatory fit hypotheses highlight researchers' efforts to make sense of performance changes observed among older adults under stereotype threat.

Various individual factors moderate the effect of stereotype threat on older adults' memory performance. For example, stereotype threat effects vary based on older adults' opinions regarding their own memories. Older individuals with more negative beliefs about their memory abilities (i.e., lower memory self-efficacy, which positively correlates with decreased task persistence; Beaudoin & Desrichard, 2017; Ponds & Jolles, 1996) demonstrate increased

vulnerability to stereotype threat relative to those with more positive beliefs, and likewise, there is a positive correlation between value placed on memory abilities and susceptibility to stereotype threat among older participants (Hess & Hinson, 2006). Furthermore, older adults with more positive perceptions of aging exhibit decreased vulnerability in the face of age-based stereotype threat. Fernández-Ballesteros and colleagues (2015) manipulated passage type to investigate the interaction between perceptions of aging and susceptibility to stereotype threat. Whereas perceptions of aging had little to no effect on participants in the non-threat and positive stereotype conditions, perceptions of aging impacted performance among threatened participants such that individuals with more positive aging perceptions demonstrated lower susceptibility to the threat (Fernández-Ballesteros et al., 2015).

Beyond older adults' beliefs regarding their memory abilities and the aging process in general, individual coping styles and personality traits may influence people's responses to stereotype threat. People high in defensive pessimism – a coping method characterized by rumination and negative thinking in the face of stress – tend to perform best under situations of high stereotype threat (Perry & Skitka, 2009). Personality might play a role, too: Taconnat and colleagues (2022) found that individuals high in neuroticism (a Big Five personality trait characterized by negative emotionality and vulnerability to stress; Soto & John, 2017) exhibited poor episodic memory performance relative to those low in neuroticism, likely related to the positive correlation between neuroticism and low self-efficacy. However, Taconnat et al. (2022) did not directly manipulate stereotype threat in their experiment. Given the connection between vulnerability to stereotype threat and memory self-efficacy, it is plausible that individuals high in neuroticism would be susceptible to performance decrements under stereotype threat (Hess & Hinson, 2006; Ponds & Jolles, 1996).

Just like individual factors moderate the effect of stereotype threat on older adults' episodic memory performance, task-specific features can also be at play. For instance, when the memory task at hand is difficult, older participants are particularly vulnerable to stereotype threat manipulations (Barber & Lui, 2020; Brubaker & Naveh-Benjamin, 2018; Mazerolle et al., 2012). The impact of task difficulty on susceptibility to stereotype threat has been shown to be particularly prominent when older participants feel as though they lack enough resources to succeed given the task demands (Barber et al., 2020). Conversely, high levels of task ease can reduce older participants' stress during memory tests, thereby serving as a protective factor potentially countering the negative impact of stereotype threat on episodic memory performance (Geraci & Miller, 2013).

When older participants experience prior task success (i.e., successfully complete a cognitive task, like a general knowledge test or an easy sentence scramble activity; e.g., Park et al., 2002) before completing an episodic memory test, they tend to feel lower levels of anxiety and perform better relative to control participants (Geraci & Miller, 2013). The beneficial impact of prior task success on older adults' episodic memory performance has also been demonstrated in memory tasks with real-world applications, such as eyewitness testimony (Rossi-Arnaud et al., 2018). Helpful effects of prior task success have only been found in situations in which cognitive tasks are utilized; success in non-cognitive tasks (e.g., a bean bag toss) does not positively impact memory performance (Geraci et al., 2016). Given that stereotype threat negatively impacts older adults' episodic memory performance whereas prior task success benefits it, prior success on a cognitive task may provide older participants with information to disprove stereotype threat (Geraci et al., 2016). Critically, the joint effect of stereotype threat and prior task success on older participants' memory performance has yet to be tested in a single study.

Rossi-Arnaud and colleagues (2018) conducted two experiments in which they first studied the effect of stereotype threat on older participants' eyewitness memory and then tested the influence of prior task success on older participants' eyewitness memory. However, these two studies were separate. In a different experiment, Geraci and colleagues (2016) found that prior task success increased older adults' response speed to stereotypically positive age-related words (e.g., *wise*), suggesting a role of task success in potentially de-emphasizing negative age-related stereotypes. Nevertheless, researchers have yet to explicitly manipulate both stereotype condition and prior task success together. The present experiment investigated both factors' joint effect on older participants' memory performance in a single experiment.

The purpose of the present research was to combine stereotype threat and prior task success in an older adult (ages 60+) memory experiment. We first conducted Experiment 1 to determine whether we could detect an effect of stereotype threat in an online episodic memory test. To preview the results, older adults' recall accuracy did not differ as a function of stereotype condition, so we designed Experiment 2: an in-person episodic memory task with explicit stereotype manipulations and a prior task success component. We investigated whether prior task success would impact older participants' vulnerability to stereotype threat.

Participants read one of two paragraphs at the start of the Experiment 2: a threat paragraph or a neutral one. Then, participants were randomly assigned to a prior task success group (success vs. no success), after which they completed an episodic word recall task. Participants answered questions regarding perceived stereotype threat effects as well as the way in which they personally internalized the threat (if applicable). Finally, participants filled out a questionnaire assessing their perceptions of aging as well as a short personality inventory.

We hypothesized that threatened participants would perform worse relative to neutral participants due to the negative influence of stereotype threat on older adults' episodic memory performance (e.g., Hess et al., 2009). However, we expected that this effect would depend on reactivity, such that only participants who internalized the threat in a negative way (as reported in the end-of-experiment threat reaction question; e.g., as anxiety-inducing) would demonstrate a performance decrement. In accordance with the regulatory fit hypothesis, we hypothesized that the threat group would make fewer intrusions than the neutral group (e.g., Barber & Mather, 2013). Next, we predicted that the prior task success group would – on average – outperform the no success group, consistent with earlier research suggesting that prior task success benefits older participants' episodic memory (Geraci et al., 2016; Geraci & Miller, 2013). We anticipated a stereotype by prior task success condition interaction, such that prior task success would reduce the negative effect of stereotype threat among the threatened success group. In other words, we expected a larger difference between threatened and neutral participants in the no success group than in the success group, because we predicted that prior task success would buffer threatened participants against the threat (Geraci et al., 2016). Finally, we hypothesized that individuals' reactions to the stereotype threat, perceptions of aging, and personality would moderate the influence of stereotype condition and prior task success group on episodic memory performance (e.g., Fernández-Ballesteros et al., 2015; Hess & Hinson, 2006; Taconnat et al., 2022).

Experiment 1

Method

Participants

All participants were recruited through Prolific (<https://prolific.co>) and had United States IP addresses. Sample size was determined based on Hess and colleagues (2009), who

investigated the effect of stereotype threat on older adults' episodic memory performance in a sample of 64 older adults. Data from two individuals were omitted from analyses because the participants reported writing down answers, leaving us with 63 older adults' worth of data to analyze ($M_{age} = 66.68$, $SD = 5.80$, range: 55-84; $M_{education} = 15.76$, $SD = 2.57$, range: 12-22).

Median completion time was 6 minutes and 51 seconds, and following the guidelines of Prolific, participants were compensated at a rate of \$19.27/hour.

Materials

Stereotype manipulation. Participants in the threat condition read a paragraph informing them that the memory task they were about to complete was one on which younger adults tend to outperform older adults. These participants also read that the upcoming test was one that researchers commonly used to investigate the effect of aging on memory, and they were told that poor performance would be predictive of dementia and future cognitive decline. In contrast, the neutral paragraph framed the study as a word processing task rather than a memory test.

Participants in the neutral condition read that the study's objective was to assess individual differences in cognition, and they were assured that the test was free of age-related biases. Both paragraphs were adapted from Barber and Mather (2013); the neutral paragraph was slightly modified. Importantly, analyses comparing the effectiveness of different categories of stereotype threat manipulations (i.e., explicit, subtle, and implicit) have demonstrated that older adults' episodic memory is most vulnerable to explicit stereotype manipulations (e.g., paragraphs, such as the ones that participants in the threat condition read in the present study; Armstrong et al., 2017). For this reason, we manipulated passage type rather than using a more subtle form of manipulation.

Memory task. Participants studied 30 words from Van Overschelde et al.'s (2004) category norms, which was similar to the encoding task used by Hess and colleagues (2003) in their study on stereotype threat and episodic memory. To avoid the confound of semantic relatedness, we chose the seventh category member from the first 30 categories.

Manipulation check. The manipulation check was taken from Chasteen and colleagues (2005) and included three statements about perceived stereotype threat ((i.e., the extent to which: participants believed their memory ability was underestimated; they felt the need to disprove others' inaccurate perceptions; and/or they sensed that the experimenter expected them to perform poorly due to their old age; e.g., *I often feel I have to prove to others that their perceptions of my memory ability are wrong*). Participants indicated their level of agreement by choosing a number along a five-point Likert-type scale, with higher numbers reflecting greater perceived threat.

Procedure

The purpose of Experiment 1 was to determine whether stereotype threat effects can be detected in an online format. The study was programmed on Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2020) and conducted on Prolific (<https://www.prolific.co/>). Participants took the study individually and on the computer, and they all provided informed consent. Then, they were randomly assigned to the threat group or the neutral group, at which point they read the passage from Barber's and Mather's (2013) stereotype manipulation. Participants next were presented with the 30 semantically unrelated words in a random order. Consistent with Barber and Mather (2013), each word was shown on the computer screen for 2 seconds, after which a fixation cross was flashed for 0.5 seconds. The order in which words were presented was randomized across participants. Following the study phase, participants took a

recall test in which they were asked to write down as many words as they could remember, and then they answered three questions assessing their level of perceived threat (Chasteen et al., 2005). Next, participants indicated whether they had cheated (i.e., by writing down words while they were flashed) or completed the study in multiple sittings, and they were asked whether they believed their data should be excluded from analyses. Finally, participants filled out the demographic questionnaire, read the debriefing form, and received monetary compensation through Prolific.

Results and Discussion

We conducted an independent-samples *t*-test to compare threatened and neutral participants' recall accuracy. Overall, there was no difference in performance as a function of stereotype condition, $t(61) = -.88, p = .384$. Furthermore, participants in the threat and neutral conditions responded similarly to the three items in the manipulation check, all $t_s \geq -1.41, p_s \geq .163$. In fact, the nonsignificant pattern of manipulation check data suggests that participants in the neutral group may have perceived more stereotype threat than those in the threatened group, but it is necessary to acknowledge that this was a numerical trend rather than a significant difference. Therefore, we decided to conduct Experiment 2 in person to optimize our chances of detecting differences in participant performance due to stereotype condition. Experiment 2 addressed our primary research question: Whether prior task success buffers older participants against stereotype threat in an episodic memory test.

Experiment 2

Method

Participants

Participants were 108 older adults ($M_{age} = 71.48$, $SE = .62$, range: 60-88; $M_{education} = 16.19$, $SE = .24$, range: 12-24; 85 females, 23 males) from the greater Waterville area, who came into the Memory and Language Lab at Colby College. Most participants were white and native English speakers, and they were recruited and tested throughout February, March, and early April via email, phone calls, flyers, and word of mouth. To determine an appropriate sample size, we conducted an a priori test in G*Power using a medium effect size ($f = .25$), a power of .80, and the standard alpha level of .05. The analysis for the stereotype condition by prior task success group interaction yielded us an estimated sample size of 128 participants (32 per group; Faul et al., 2007). We tried to achieve a sample of as close to 128 participants as possible, but due to the time constraints of Colby's 2023 spring semester, we fell short of that goal. Our original sample included 109 adults, but one participant's intrusion rate was over 2.5 standard deviations from the mean, so we excluded this individual's data (see Table 1 for more condition-specific demographic information). Most adults took under an hour to finish the study. Participants were paid at a rate of \$10 per hour, and they were also gifted a small card stock thank you note upon completion of the experiment.

Materials

Stereotype manipulation. The passages were the same as in Experiment 1.

Prior task success manipulation. Consistent with Geraci and Miller (2013), participants in the success group completed a sentence scramble activity in which they were presented with various series of five words and asked to form logical four-word sentences (e.g., if presented with *she the plays blanket piano*, participants would write "*she plays the piano*"). Geraci and Miller found a 97% success rate among older adult participants in their experiment, meaning these stimuli were intentionally designed to induce the experience of success. There were 30

sentences in total, and participants completed the activity on paper using a pen or pencil. In contrast, adults in the no success condition simply read the 30 four-word sentences. The stimuli were adapted from Geraci and Miller to avoid overlap with words on the memory test, and all words were emotionally neutral and unrelated to aging. Our first 11 participants were exposed to sentences with two words that were included in the episodic memory test (“strawberry” and “swimming”), so following this realization, we edited the sentences to prevent additional word overlap and interference. Analyses on participants’ performance accuracy and intrusion rate as a function of word overlap did not demonstrate any meaningful difference between the first 11 individuals’ performance and the remaining participants’; thus, we did not exclude the first 11 individuals’ data (see Footnote for additional information).

Memory task. The encoding task was identical to that in Experiment 1, with the exception that the fixation cross was removed. Participants studied 30 semantically unrelated words (e.g., *amethyst*, *hill*, and *Spain*; Van Overschelde et al., 2004). Each word was presented for 2 seconds, and the words’ order was randomized across participants (Barber & Mather, 2013). The words were the same as in Experiment 1, excluding the replacement of *grandfather* with *sister* due to a realization regarding the former’s old age connotation – and, thus, its potential to implicitly activate stereotype threat. Overall, the words were emotionally neutral and unrelated to aging.

Manipulation check and threat reaction. The manipulation check was the same as in Experiment 1. To ensure that we were gauging participants’ individual reaction to the stereotype manipulation (e.g., whether they found it motivating, as was explicitly shared by a 2021 online participant in the threat group through Prolific’s internal messaging service) – and in light of the nonsignificant difference between the threatened and neutral groups’ scores on Chasteen et al.’s

(2005) manipulation check in Experiment 1 – we also included an open-ended threat reaction question. The prompt specifically asked participants to indicate whether and how reading the initial passage (threat vs. neutral) affected how they engaged with the memory task. We qualitatively coded the threat reaction by first reading through the majority of participants' responses and identifying preliminary dimensions. Research Assistants in the Memory and Language Lab ultimately agreed on the following eight categories into which we would code participants' responses: Didn't read/doesn't remember; Neutral; Acceptance/lack of surprise; Motivated; Relaxed; Anxiety-inducing; Decreased pressure/defeatism; and Negative self-expectations. Five Research Assistants individually coded the threat reactions, and each Research Assistant was given half of the participants' responses (i.e., three individuals coded the first 55 participants' responses, and the remaining two coded the final 54). Responses referencing a given dimension were coded as a 1 for that particular category, whereas responses not referencing said dimension were coded as a 0. Importantly – to limit coding biases – Research Assistants were naïve to the participants' stereotype conditions while performing the coding. Initial agreement for the first group of coders was 81%, whereas original agreement for the second group was 98%. When there was ambiguity, discrepancies were resolved through conversation between the researcher and the lab's Principal Investigator.

Perceptions of aging. Consistent with Fernández-Ballesteros and colleagues (2015), participants filled out a five-item subscale of the Philadelphia Geriatric Center Morale Scale (PGCMS; Lawton, 1975). The questionnaire assessed the adults' self-perception of aging (i.e., the extent to which they viewed aging in a positive light) and included statements (e.g., *Things keep getting worse as I get older*) to which the participants needed to agree or disagree. Negative responses to the first four items were scored as a 0, positive responses were scored as a 1, and the

first and third items were reverse coded to ensure consistency in the coding scheme. Overall, higher scores on the questionnaire reflected more positive perceptions of aging. Four of the five items required an answer of “yes” (agree) or “no” (disagree); the fifth item, which asked participants to compare their expectations of aging to the reality of aging, required participants to choose among three possible answers (“better,” “worse,” or “the same”). For the purposes of scoring, “better” and “the same” were coded as one. Importantly, two idiosyncratic participants highlighted the ambiguous wording of the fourth statement, which read “I am as happy now as I was when I was younger” (Fernández-Ballesteros et al., 2015). Given that disagreement to said statement could reflect either decreased or increased happiness in older age, there was no way to accurately score this fourth item without additional information from participants. Thus, we excluded the fourth item from analyses.

BFI-2-S. Because we were interested in whether personality influences participants’ susceptibility to stereotype threat, we included the BFI-2-S, a 30-item personality inventory (Soto & John, 2017). Participants read statements about personality (e.g., *I am someone who has difficulty getting started on tasks*), and then they decided the extent to which each statement applied to their personality by choosing along a five-point scale ranging from *strongly disagree* to *strongly agree*. We opted for the 30-item short form rather than the traditional 60-item inventory to limit participant fatigue. Adhering to the coding scheme provided by Soto and John (2017), we coded for the Big Five personality domains and the extraversion facet trait of assertiveness. We chose this specific facet trait because we were interested in whether individuals scoring higher in assertiveness might demonstrate greater resistance to the stereotype threat.

Procedure

The experiment was again programmed on Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2020), and unlike Experiment 1, the older adults completed this study in person. Participants came into Colby College's Memory and Language Lab where they were greeted by a Research Assistant, who gave them a study overview. After providing informed consent, the program randomly assigned participants to the threat group or the neutral group (Barber & Mather, 2013). To increase the extent to which age was made salient for the threat group, threatened participants typed their age into a text box after reading their assigned paragraph. Neutral participants typed the date. Importantly, to prevent biased behavior on the part of the experimenter, the Research Assistant was naïve to the participants' stereotype condition at the start of the study. Participants next completed the prior task success manipulation: They were randomly assigned to the reading (no success) or sentence scramble (success) task, which they completed on a printed out piece of paper (Geraci & Miller, 2013). The experimenter waited in the room to hand each participant their respective paper (reading or sentence scramble), and then they left the room while the participant independently completed the remainder of the experiment at their own pace.

Following the stereotype and prior task success manipulations, participants completed the episodic memory task, which consisted of a study phase and a recall assessment. To ensure that the stereotype manipulation had not dissipated during the prior task success activity, participants received a stereotype condition reminder at the start of the encoding task (e.g., "*As a reminder, this is a memory task on which younger adults generally outperform older adults (60+)*", for the threat group), after which they were shown the 30 words. During the recall test, participants wrote down as many words as they could remember on a piece of paper numbered one through 30. There was no time limit for the recall task, and participants clicked the "next" button on the

screen once they felt they had written down all of the words they remembered. Participants then filled out the manipulation check (Chasteen et al., 2005) and responded to the open-ended threat reaction question, and they also filled out the perceptions of aging (Fernández-Ballesteros et al., 2015) and BFI-2-S (Soto & John, 2017) questionnaires. Finally, participants completed the Shipley vocabulary task (Shipley, 1940), filled out a demographics form, and read detailed text debriefing them. Demographics were deliberately placed after the memory task – not before – to minimize the chances of drawing attention to age differences and subtly activating stereotype threat in the neutral group (Steele & Aronson, 1995). The Research Assistant then met each participant in the testing room and answered any remaining questions about the experiment, after which they administered the Mini-Mental State Exam (MMSE; Folstein et al., 1975). Following the MMSE, the older adults received compensation from the Research Assistant, who walked them out of the lab and thanked them for their participation.

Results

The results section first reports manipulation check and threat reaction data, followed by performance on the memory task (recall accuracy and the number of intrusions). Where appropriate, we report partial η^2 for effect sizes.

Manipulation Check and Responses to Open-Ended Question

We analyzed data from Chasteen's and colleagues' (2005) three-item questionnaire question-by-question. We conducted three 2 (stereotype condition: threat vs. neutral) x 2 (prior task success group: success vs. no success) between-subjects ANOVAs to assess the effect of stereotype condition and prior task success group on older participants' perceived stereotype threat. There was a significant main effect of stereotype threat on responses to the third item, which pertained to negative experimenter expectations. Relative to the neutral group ($M = 2.29$,

$SD = 1.08$), threatened participants ($M = 2.75$, $SD = 1.08$) were more likely to perceive that the experimenter expected them to perform poorly, $F(1, 93) = 4.30$, $p = .041$, partial $\eta^2 = .04$.

Additional effects for the dependent variable of experimenter expectations were nonsignificant, all $F_s \leq 1.01$, $p_s \geq .317$. Moreover, there were no significant effects for the other two dependent measures from Chasteen et al.'s (2005) manipulation check, which measured the extent to which participants felt their memory ability was underestimated as well as the need to disprove others' inaccurate perceptions about their memories, all $F_s \leq 3.10$, $p_s \geq .082$ (see Figure 1). The effect of the stereotype condition on participants' perceptions of the experimenter's expectations reveals – at the very least – some evidence that participants were sensitive to the stereotype manipulation.

Additional evidence for participants' slight sensitivity to the stereotype manipulation comes from responses to the open-ended threat reaction questions. We conducted chi square analyses to determine whether there were different distributions of open-ended threat reaction response dimensions (Didn't read/doesn't remember; Neutral; Acceptance/lack of surprise; Motivated; Relaxed; Anxiety-inducing; Decreased pressure/defeatism; and Negative self-expectations) as a function of stereotype condition. There was a significant difference in references to the neutral dimension as a function of stereotype condition, $\chi^2(1) = 8.97$, $p = .004$. This effect indicated that, whereas 80% (44 out of 55) of neutral participants perceived that the passage had no effect on them, only around 50% (28 out of 53) of participants in the threat group referenced the neutral dimension. In addition, more threatened participants' responses (7/53) fit the acceptance/lack of surprise dimension than those of stereotype neutral participants (1/55), $\chi^2(1) = 5.11$, $p = .030$. Furthermore, there was a significant difference in references to negative expectations as a function of stereotype condition, such that threatened participants (9/53) referenced this dimension more frequently than neutral participants (1/55), $\chi^2(1) = 7.39$, $p = .008$.

The differences in references to the neutral, acceptance, and negative expectations dimensions as a function of stereotype condition highlights that, at least from a metacognitive perspective, threatened and neutral participants perceived the two passages in slightly different ways. There were, however, no significant differences in reported motivation, relaxation, anxiety, or decreased pressure/defeatism due to stereotype group, all χ^2 s $\leq .53$, $ps \geq .523$. Similarly, the number of participants who wrote that they either did not read or could not remember the stereotype passage did not differ meaningfully as a function of stereotype condition, $\chi^2(1) = 2.00$, $p = .271$ (see Table 2 for dimension definitions, examples, and group differences). Importantly, only 8 of the total 108 participants referenced the “Didn’t read/doesn’t remember” dimension. Thus, the majority of participants *did* process the stereotype manipulation to some extent, regardless of whether this affected the way in which they engaged with the episodic memory task.

Our final manipulation check analyses tested whether there were any significant differences between people who perceived an effect of stereotype condition (i.e., those whose open-ended threat reaction fit the neutral dimension) and people who did not (i.e., those who did not reference the neutral dimension). We only analyzed this dimension because the other dimensions did not have enough observations for meaningful comparisons. We wanted to analyze whether any systematic factors might account for the difference between participants who did and did not feel threatened. We first conducted a series of *t*-tests within both stereotype conditions, which revealed non-significant effects on various output measures (i.e., age, education, memory accuracy, intrusions, and Shipley and MMSE scores), all $ts \geq -1.51$, $ps \geq .139$. We then repeated these same *t*-tests for additional output variables (e.g., perceptions of aging and personality), and again, there were no meaningful differences between participants

who reported feeling affected by the stereotype manipulation and those who did not, all $t_s \geq -2.08$, $p_s \geq .043$. In sum, these data reveal that there were no systematic individual factors that can account for the difference between participants who did and did not perceive an effect of the stereotype manipulation.

Proportion of Words Correctly Recalled

To assess the effect of stereotype condition and prior task success group on episodic memory accuracy – our primary research question – we conducted a 2 (stereotype condition: threat vs. neutral) x 2 (prior task success group: success vs. no success) between-subjects ANOVA, with proportion of words recalled as our dependent measure. Recall accuracy did not vary as a function of stereotype condition or prior task success group, and there was also no stereotype condition by prior task success group interaction, all $F_s \leq 1.25$, $p_s \geq .267$ (see Figure 2). Put differently, in contrast to one of our primary hypotheses, the stereotype and prior task success manipulations did not jointly influence the number of words participants were able to recall. Thus, prior task success did not buffer against stereotype threat, and neither stereotype condition nor prior task success group impacted the older adults' memory accuracy.

Number of Intrusions

In light of the regulatory fit hypothesis, we were curious about whether participants' stereotype condition might impact their intrusion rate. To investigate this, we ran the same 2 x 2 between-subjects ANOVA, with the number of intrusions as our dependent variable. We found a significant main effect of stereotype condition, such that the threat group ($M = .70$, $SD = 1.01$) made significantly fewer intrusions than the neutral group ($M = 1.29$, $SD = 1.78$), $F(1, 104) = 4.46$, $p = .037$, partial $\eta^2 = .04$. In other words, although there were no differences in memory accuracy as a function of stereotype condition, the threat group was significantly less likely than

the neutral group to recall words that had not been displayed during the encoding task. No other effects were significant, both $F_s \leq .31$, $p_s \geq .580$.¹

Exploratory Analyses

The sole main effect of stereotype threat that we detected was for the dependent variable of intrusion rate. We were interested in whether the individual factors of perceptions of aging and personality traits would influence participants' susceptibility to the threat manipulation. Our correlational analyses revealed that neither predictor significantly correlated with intrusion rate and, therefore, could not have significantly predicted threat susceptibility. Thus, we did not continue with these exploratory analyses.

Discussion

In the present study, we investigated whether prior task success would buffer against the negative impact of stereotype threat on older adults' episodic memory performance. In light of the empirically-established and harmful impact of stereotype threat on older adults' episodic memory, we were motivated to identify a novel form of environmental support (i.e., success on an easy cognitive task, such as the sentence scramble that we included) that might shield older individuals from this performance decrement.

Data from the manipulation check and open-ended threat reactions demonstrated that participants did – at least to some degree – process the stereotype manipulation. Relative to neutral participants, threatened adults were more likely to agree with the statement suggesting that the experimenter had negative expectations regarding the participants' performance. We included the open-ended threat reaction question to gain insight into the specific ways in which participants internalized the stereotype manipulation, and we found that threatened participants referenced the acceptance and negative self-expectations dimensions significantly more

frequently than neutral participants while also referencing the neutral dimension significantly less frequently. Furthermore, only eight out of 108 participants admitted to not reading or remembering the stereotype manipulation. Thus, many participants *did* perceive the stereotype manipulation, as evidenced by their differential manipulation check and threat reaction responses and the low number of observations in the “didn’t read/doesn’t remember” category. In spite of this, however, participants did not demonstrate observable performance-based changes while under the stereotype manipulation.

In contrast to our primary hypothesis – and in disagreement with prior research – stereotype condition and prior task success did not jointly impact participants’ episodic recall accuracy (but see Geraci et al., 2016). Likewise, stereotype condition alone did not impact recall accuracy (but see Bouazzaoui et al., 2015), and neither did prior task success group (but see Geraci et al., 2016; Geraci & Miller, 2013). Given that the manipulations we included have been empirically-verified – and given some of the differential responses to manipulation check and threat reaction items as a function of stereotype condition – it is unlikely that our manipulations were completely ineffective (Barber & Mather, 2013; Geraci & Miller, 2013). It is unclear why the beneficial effect of prior task success on episodic memory accuracy failed to replicate (Geraci et al., 2016; Geraci & Miller, 2013). One possibility is that our 30 words were semantically unrelated, whereas other researchers have typically used groups of five words from six different semantic categories (Geraci et al., 2016; Geraci & Miller, 2013; Hess et al., 2003). We intentionally chose these word stimuli to avoid the potential confound of semantic relatedness, but it is possible that testing older participants’ memory for 30 isolated words was too challenging. In fact, a previous team of researchers using six semantic categories observed an average recall of .5 (15 words) among their older adult participant pool (Hess & Pullen, 1994,

cited in Hess et al., 2003), whereas the average recall rate among our participants was .22 (between six and seven words). The sheer level of task difficulty may have minimized any potential benefit conferred by the sentence scramble activity. Nevertheless, if this were true, we should have detected a main effect of stereotype condition, given the positive correlation between task difficulty and stereotype threat susceptibility (e.g., Brubaker & Naveh-Benjamin, 2018).

It could be argued that the null effects of stereotype condition and prior task success group stemmed from problems with the pairing of the two manipulations. For instance, the prior task success manipulation may have lessened the efficacy of the stereotype manipulation, such that the experience of success counteracted the sense of threat that some participants were supposed to feel. In fact, this was the primary hypothesis motivating this research. Similarly, the stereotype manipulation reminder may have reduced the effect of the task success in that threatened participants' feeling of success dissipated once they were reminded of the threat. Nevertheless, neither explanation seems particularly plausible considering the lack of main effects of either stereotype condition or prior task success group on recall accuracy. Specifically, if prior task success eliminated or reduced the effect of threat, participants in the no success group should have still shown effects of threat relative to those in the neutral condition. Furthermore, if the threat reminder counteracted the effect of success, participants in the neutral condition should have still demonstrated benefits of prior task success.

It is possible that self-selection bias was at play in our experiment, meaning that the type of participant who volunteered to take part in this research may have been particularly resilient in the face of age-based stereotype threat. However, considering that other researchers have detected stereotype threat effects conducting in-person research, self-selection bias is not an

entirely logical explanation for the null results in the present study. Nonetheless, Colby College's Memory and Language Lab is a very friendly environment that has an established rapport with many members of its older adult participant pool. The warm atmosphere of the lab, coupled with the self-selection bias among participants, may have reduced threatened participants' vulnerability to the stereotype manipulation. This potential explanation for the nonsignificant effect of stereotype threat on recall accuracy is particularly plausible in light of research on the *positivity effect*, which has demonstrated that – relative to younger adults – older adults' attention and memory are biased toward positive stimuli over negative stimuli (Mather & Carstensen, 2005). Given older adults' cognitive bias for positively-valenced stimuli, it is possible that the older participants in our study were too focused on the pleasant aspects of the research experience (e.g., the friendly greeting they received upon arrival and their interactions with the Research Assistant) to feel sensitive to the threat.

It is intriguing to compare the responses to manipulation check items and threat reaction dimensions that did and did not differ as a function of stereotype condition because they might help clarify our null effects. In Chasteen's and colleagues' (2005) three-item questionnaire, the first two items pertain to the participants' feeling of being underestimated by others due to their age and needing to disprove others' inaccurate beliefs about their memory performance. In contrast, the third item pertains to the experimenter's expectation that the older participants would perform badly due to their age. Whereas the first two items addressed the older participants' general beliefs about others' age-based judgments, the third item was unique in that it was more task-specific; it pertained to the participants' perceptions regarding the experimenter's expectations during the study at hand. It is possible that the participants do perceive a sense of stereotype threat in their daily lives (e.g., when interacting with younger

people in their life; Barber & Lui, 2020) yet did not feel particularly stereotype threatened during the study – an artificial series of tasks that may have felt less impactful than real-life experiences. In this way, threatened participants' greater agreement with the experimenter expectations item relative to that of neutral participants may have reflected their awareness of the manipulation intentions rather than their own feelings of stereotype threat during the study.

This speculation is particularly plausible given the threat reaction response data. Threatened participants were significantly more likely than neutral participants to reference the acceptance/lack of surprise and negative self-expectations dimensions, yet the frequency of references to the dimensions of anxiety-inducing, motivated, relaxed, and decreased pressure/defeatism did not differ as a function of stereotype condition. The dimensions of acceptance/lack of surprise and negative self-expectations pertained more to participants' general beliefs about ageist stereotypes and their overall memory abilities, whereas the other dimensions were more task-specific; they assessed participants' metacognitive state during the free recall task at hand. It is possible that the disconnect between the items/dimensions that were task-specific relative to more general assessments of perceived stereotype threat in daily life might clarify why stereotype threat impacted participants' manipulation check scores and dimension references for certain items but not others.

In light of the potential distinction between task-specific and more general perceptions of stereotype threat, participants' preexisting memory beliefs may have limited the extent to which they demonstrated performance susceptibility to the stereotype manipulation. Whereas other types of stereotype threat – such as (completely inaccurate and unfounded) stereotypes about racial differences in intellect (Steele & Aronson, 1995) – have been characterized as threats to group identity, old age-related stereotype threat is considered a self-concept threat (Barber,

2017). Self-concept stereotype threats are those characterized by individuals' fear that their own behavior will confirm to them that the activated negative stereotype is true of themselves (e.g., that they genuinely are suffering from old age-related memory loss; Barber, 2017). The conviction that age-based stereotype threat is a form of self-concept threat corroborates multiple threatened participants' insight into a potential reason why the stereotype threat passage had no observable effect on their performance. Numerous participants made a comment suggesting that they had entered the experiment with the knowledge that their memories were getting worse and were thus not impacted by the passage; rather, the text reiterated general information about their own memory abilities of which they were already aware, and they thus felt no effect of the stereotype threat. Of course, participants' metacognitive judgments should be taken with a grain of salt, given the limits to people's metacognitive accuracy (e.g., Morson et al., 2015). That said, participants may have participated with pre-existing strong beliefs about their own memories, whether based on a lifetime of internalized stereotypes, real-life encounters with ageism, or personal observations of their own memory-related changes.

Albeit regularly oversimplified and overgeneralized by the media, aspects of episodic memory (such as binding) do decline with older age (e.g., Blazer et al., 2015; Loaiza & Srokova, 2020). Most participants likely had ideas about age-related memory changes, and it is also plausible that participants with strongly internalized beliefs about their own memories – regardless of whether they were positive or negative – demonstrated a lack of susceptibility to the stereotype manipulation. Researchers have attempted to implement self-efficacy training to improve older participants' memory performance, but these trainings are often ineffective, which sheds light on the extent to which these beliefs are ingrained in people (e.g., Geraci et al., 2016; Geraci & Miller, 2013). Moreover, memory self-efficacy might partially determine the extent to

which a participant would demonstrate vulnerability to a stereotype threat manipulation (see Barber & Lui, 2020; Bouazzaoui et al., 2015). Desrichard and Köpetz (2005) found that participants with lower memory self-efficacy showed impaired performance in the face of stereotype threat, whereas individuals with high memory self-efficacy did not. Future experiments should investigate the way in which the strength of people's metamemory beliefs (e.g., memory self-efficacy beliefs) – no matter their valence – may influence the degree to which stereotype threat might induce measurable changes in older adults' episodic memory accuracy. Such research might shed light on individual factors that may or may not leave an older individual vulnerable to age-based stereotype threat.

Whereas neither stereotype condition nor prior task success group impacted participants' recall accuracy, we did detect a significant effect of stereotype condition on intrusion rates, such that threatened participants made significantly fewer memory intrusions than neutral participants. These data are consistent with the regulatory fit hypothesis, which posits that stereotype threat reduces memory intrusions (e.g., Fourquet et al., 2020). Stereotype threat can induce a prevention-based focus in participants, such that threatened older adults respond more carefully and engage in greater task monitoring relative to neutral participants (Barber & Lui, 2020; Barber & Mather, 2013). Barber and Mather (2013) have explained that, in cognitive assessments, participants typically have either a promotion- (characterized by the objective of accomplishment and advancement) or prevention-based focus (characterized by the minimization of error; Liu et al., 2021). Considering the differential effects of stereotype condition on recall accuracy and intrusion rate in Experiment 2, it is possible that stereotype threat did not impact recall accuracy because such an outcome would be more consistent with a promotion-based focus (Barber & Mather, 2013). Rather, stereotype threat may have induced error avoidance in

our participants – regardless of whether this was fully conscious or more nonconscious on the part of the participants – such that we only detected a significant effect of stereotype condition when measuring memory intrusions.

In sum, the present study did not support our hypothesis that prior task success would counter the negative effect of stereotype threat on participants' episodic memory recall, nor that prior task success would benefit participants' recall. Likewise, stereotype condition did not affect participants' recall accuracy. Stereotype condition did, however, significantly impact intrusion rates, with threatened participants making significantly fewer memory intrusions than neutral participants. Although these experimental findings do not support many of the conclusions we had anticipated, our study contributes to a growing – and imperative – understanding of task-based factors that may protect older adults from experiencing the negative effects of stereotype threat in memory tests, both inside and outside of a laboratory setting. Further research into the phenomenological experiences associated with stereotype threat is necessary for gaining insight into the mechanisms underlying observable performance changes – or lack thereof – in threatened individuals. Experiments like these might also shed light on the role of individual factors in vulnerability to, as well as protection against, old age-related stereotype threat.

Funding: This research was funded by Colby Students' Special Projects Fund as well as the *Understanding Human Cognition* grant # 220020426 from the James S. McDonnell Foundation awarded to Jennifer Coane.

Institutional Review Board Statement: The experiments were approved by the Institutional Review Board of Colby College (protocol code 2021-082, 11 November 2022).

Open Science Statement: Experiment 2 was preregistered on the Open Science Framework (<https://osf.io>).

References

- Anwyl-Irvine, A.L., Massoné J., Flitton, A., Kirkham, N.Z., & Evershed, J.K. (2020). Gorilla in our midst: An online behavioural experiment builder. *Behavior Research Methods*, *52*(1), 388-407. <https://doi.org/10.3758/s13428-019-01237-x>
- Armstrong, B., Gallant, S. N., Li, L., Patel, K., & Wong, B. I. (2017). Stereotype threat effects on older adults' episodic and working memory: A meta-analysis. *Gerontologist*, *57*(S2), S193-S205. <https://doi.org/10.1093/geront/gnx056>
- Barber, S. J. (2017). An examination of age-based stereotype threat about cognitive decline: Implications for stereotype-threat research and theory development. *Perspectives on Psychological Science*, *12*(1), 62-90. <https://doi.org/10.1177/1745691616656345>
- Barber, S. J. (2020). The applied implications of age-based stereotype threat for older adults. *Journal of Applied Research in Memory and Cognition*, *9*(3), 274–285. <https://doi.org/10.1016/j.jarmac.2020.05.002>
- Barber, S. J., & Lui, K. (2020). Stereotype threat and the cognitive performance of older adults. In A. K. Thomas & A. Gutchess (Eds.), *The Cambridge handbook of cognitive aging: A life course perspective* (pp. 400-416). Cambridge University Press. <https://doi.org/10.1017/9781108552684.025>
- Barber, S. J., & Mather, M. (2013). Stereotype threat can reduce older adults' memory errors. *Quarterly Journal of Experimental Psychology*, *66*(10), 1888-1895. <https://doi.org/10.1080/17470218.2013.840656>
- Beaudoin, M., & Desrichard, O. (2017). Memory self-efficacy and memory performance in older adults: The mediating role of task persistence. *Swiss Journal of Psychology*, *76*(1), 23-33. <https://doi.org/10.1024/1421-0185/a000188>

- Blazer, D. G., Yaffe, K., & Liverman, C. T. (2015). Characterizing and assessing cognitive aging. In D. G. Blazer, K. Yaffe, & C. T. Liverman (Eds.), *Cognitive aging: Progress in understanding and opportunities for action* (pp. 37-88). The National Academies Press. <https://doi.org/10.17226/21693>
- Bouazzaoui, B., Follenfant, A., Ric, F., Fay, S., Croizet, J. C., Atzeni, T., & Taconnat, L. (2015). Ageing-related stereotypes in memory: When the beliefs come true. *Memory*, *24*(5), 659–668. <https://doi.org/10.1080/09658211.2015.1040802>
- Brubaker, M. S., & Naveh-Benjamin, M. (2018). The effects of stereotype threat on the associative memory deficit of older adults. *Psychology and Aging*, *33*(1), 17-29. <https://doi.org/10.1037/pag0000194>
- Chasteen, A. L., Bhattacharyya, S., Horhota, M., Tam, R., & Hasher, L. (2005). How feelings of stereotype threat influence older adults' memory performance. *Experimental Aging Research*, *31*(3), 235-260. <https://doi.org/10.1080/03610730590948177>
- Chen, Y.-T., McDonough, I. M., Faig, K. E., Norman, G. J., & Gallo, D. A. (2022). Impact of stereotype threat on brain activity during memory tasks in older adults. *NeuroImage*, *260*, 119413. <https://doi.org/10.1016/j.neuroimage.2022.119413>
- Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik, & E. Trehub (Eds.). *Aging and cognitive processes* (pp. 191–211). New York: Plenum Press.
- Desrichard, O., & Köpetz, C. (2005). A threat in the elder: The impact of task-instructions, self-efficacy and performance expectations on memory performance in the elderly. *European Journal of Social Psychology*, *35*, 537-552. <https://doi.org/10.1002/ejsp.249>

Dodson, C. S., Bawa, S., & Krueger, L. E. (2007). Aging, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Aging, 22*(1), 122-133.

<https://doi.org/10.1037/0882-7974.22.1.122>

Faul, F., Erdfelder, E., Lang, A., & 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>

Fernández-Ballesteros, R., Bustillos, A., & Huici, C. (2015). Positive perception of aging and performance in a memory task: Compensating for stereotype threat? *Experimental Aging Research, 41*, 410-425. <https://doi.org/10.1080/0361073X.2015.1053757>

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State: a Practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198.

Fourquet, N. Y., Patterson, T. K., Li, C., Castel, A. D., & Knowlton, B. J. (2020). Effects of age-related stereotype threat on metacognition. *Frontiers in Psychology, 11*.

<https://doi.org/10.3389/fpsyg.2020.604978>

Geraci, L., Hughes, M. L., Miller, T. M., & De Forrest, R. L. (2016). The effect of prior task success on older adults' memory performance: Examining the influence of different types of task success. *Experimental Aging Research, 42*(4), 365-381.

<https://doi.org/10.1080/0361073X.2016.1191860>

Geraci, L., & Miller, T. M. (2013). Improving older adults' memory performance using prior task success. *Psychology and Aging, 28*(2), 340-345. <https://doi.org/10.1037/a0030332>

Haslam, C., Morton, T. A., Haslam, S. A., Varnes, L., Graham, R., & Gamaz, L. (2012). "When the age is in, the wit is out": Age-related self-categorization and deficit expectations

- reduce performance on clinical tests used in dementia assessment. *Psychology and Aging*, 27(3), 778-784. <https://doi.org/10.1037/a0027754>
- Hess, T. M., Auman, C., & Colcombe, S. J. (2003). The impact of stereotype threat on age differences in memory performance. *The Journals of Gerontology*, 58(1), P3-P11. <https://doi.org/10.1093/geronb/58.1.P3>
- Hess, T. M., Emery, L., & Queen, T. L. (2009). Task demands moderate stereotype threat effects on memory performance. *The Journals of Gerontology: Series B*, 64B(4), 482-486. <https://doi.org/10.1093/geronb/gbp044>
- Hess, T. M., & Hinson, J. T. (2006). Age-related variation in the influences of aging stereotypes on memory in adulthood. *Psychology and Aging*, 21(3), 621-625. <https://doi.org/10.1037/0882-7974.21.3.621>
- Hess, T. M., & Pullen, S. M. (1994). Adult age differences in informational biases during impression formation. *Psychology and Aging*, 9, 237-250.
- Lawton, M. P. (1975). The Philadelphia Geriatric Center Morale Scale: A revision. *Journal of Gerontology*, 30(1), 85-89. <https://doi.org/10.1093/geronj/30.1.85>
- Liu, S., Liu, P., Wang, M., & Zhang, B. (2021). Effectiveness of stereotype threat interventions: A meta-analytic review. *Journal of Applied Psychology*, 106(6), 921-949. <https://doi.org/10.1037/apl0000770>
- Loaiza, V. M., & Srokova, S. (2020). Semantic relatedness corrects the age-related binding deficit in working memory and episodic memory. *J Gerontol B Psychol Sci Soc Sci*, 75(9), 1841-1849. <https://doi.org/10.1093/geronb/gbz055>

- Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences*, 9(10), 496-502.
<https://doi.org/10.1016/j.tics.2005.08.005>
- Mazerolle, M., Régner, I., Morisset, P., Rigalleau, F., & Huguet, P. (2012). Stereotype threat strengthens automatic recall and undermines controlled processes in older adults. *Psychological Science*, 23(7), 723-727. <https://doi.org/10.1177/0956797612437607>
- Morson, S. M., Moulin, C. J.A., & Souchay, C. (2015). Selective deficits in episodic feeling of knowing in ageing: A novel use of the general knowledge task. *Acta Psychologica*, 157, 85-92. <https://doi.org/10.1016/j.actpsy.2015.02.014>
- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, 17, 299-320. <https://doi.org/10.1037//0882-7974.17.2.299>
- Perry, S. P., & Skitka, L. J. (2009). Making lemonade? Defensive coping style moderates the effect of stereotype threat on women's math test performance. *Journal of Research in Personality*, 43, 918-920. <https://doi.org/10.1016/j.jrp.2009.05.013>
- Ponds, R. W. H. M., & Jolles, J. (1996). Memory complaints in elderly people: The role of memory abilities, metamemory, depression, and personality. *Educational Gerontology*, 22(4), 341-357. <https://doi.org/10.1080/0360127960220404>
- Popham, L. E., & Hess, T. M. (2013). Age differences in the underlying mechanisms of stereotype threat effects. *The Journals of Gerontology: Series B*, 70(2), 223-232.
<https://doi.org/10.1093/geronb/gbt093>

- Rossi-Arnaud, C., Spataroa, P., & Geraci, L. (2018). Effects of stereotype threat and prior task success on older adults' eyewitness memory. *Journal of Applied Research in Memory and Cognition*, 7(3), 422-431. <https://doi.org/10.1037/h0101822>
- 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>
- Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *Journal of Psychology*, 9, 371-377.
- Smith, A. M., Gallo, D. A., Barber, S. J., Maddox, K. B., & Thomas, A. K. (2017). Stereotypes, warnings, and identity-related variables influence older adults' susceptibility to associative false memory errors. *Gerontologist*, 57(S2), S206-S215. <https://doi.org/10.1093/geront/gnx057>
- Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, 68, 69-81.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797-811. <https://doi.org/10.1037/0022-3514.69.5.797>
- Swift, H.J., Abrams, D., & Marques, S. (2013). Threat or boost? Social comparison affects older people's performance differently depending on task domain. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 68(1), 23–30. <https://doi.org/10.1093/geronb/gbs044>
- Taconnat, L., Pinard, F., Vanneste, S., Bouazzaoui, B., Fay, S., Martinez, L., Alibrán, E., & Geraci, L. (2022). Personality traits affect older adults' memory differently depending on

the environmental support provided at encoding. *Personality and Individual Differences*, *191*, 111572. <https://doi.org/10.1016/j.paid.2022.111572>

Thomas, A. K., & Dubois, S. J. (2011). Reducing the burden of stereotype threat eliminates age differences in memory distortion. *Psychological Science*, *22*, 1515–1517. <https://doi.org/10.1177/0956797611425932>

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*(3), 289-335. <https://doi.org/10.1016/j.jml.2003.10.003>

Wong, J. T., & Gallo, D. A. (2016). Stereotype threat reduces false recognition when older adults are forewarned. *Memory*, *24*(5), 650-658. <https://doi.org/10.1080/09658211.2015.1036885>

Footnote

¹The first 11 participants were exposed to two words that showed up on the episodic memory encoding task, so we determined whether this word overlap caused interference and impacted recall accuracy and intrusion rate. We analyzed whether the first 11 participants' accuracy and intrusion rates differed from the other participants' by excluding these 11 individuals and re-running our two 2 x 2 between-subjects ANOVAs. For correct recall, accuracy did not differ as a function of stereotype condition or prior task success group, $F_s \leq 1.96$, $p_s \geq .165$. For intrusions, there was again a significant main effect of stereotype condition, $F(1, 93) = 5.27$, $p = .024$, partial $\eta^2 = .05$; the other effects were not significant, both $F_s \leq .08$, $p_s \geq .781$. We also conducted *t*-tests to examine whether the initial 11 participants' recall accuracy and intrusion rate differed significantly from the other 97 participants' data. There were no significant differences between the two groups, $t_s \leq .80$, $p_s \geq .426$. These analyses verified that there was no significant effect of exposure to the two words on accuracy or intrusion rate.

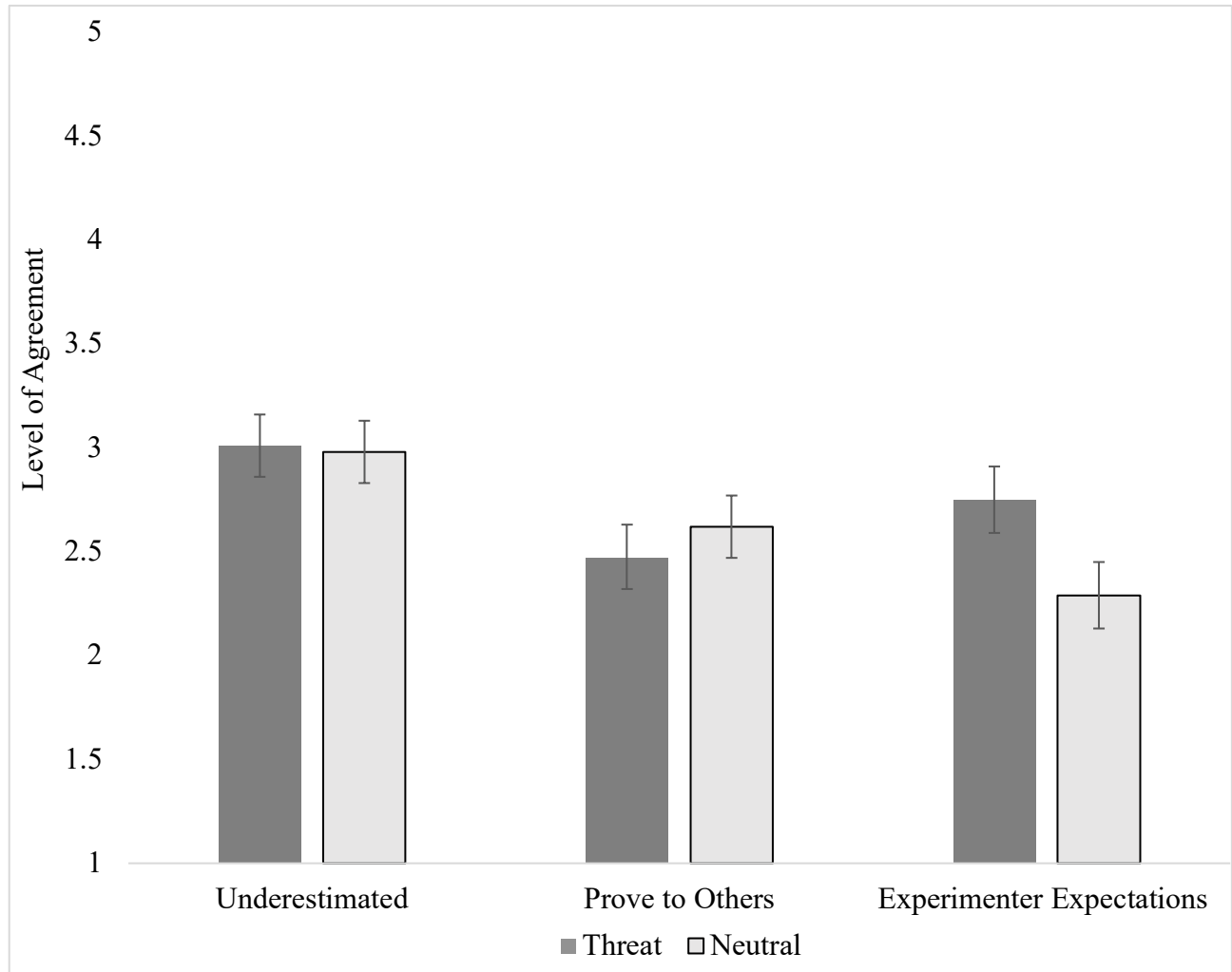
Table 1*Participant Demographic Information for Experiment 2 Based on Condition*

	Threat		Neutral	
	Success (N = 26)	No Success (N = 27)	Success (N = 27)	No Success (N = 28)
Age	72.50 (5.80)	70.04 (6.99)	72.19 (5.84)	71.25 (7.15)
Education	16.28 (2.48)	16.44 (2.66)	16.08 (2.57)	16.00 (2.00)
Gender	19 Female; 7 Male	23 Female; 4 Male	19 Female; 8 Male	24 Female; 4 Male
Shipley	35.38 (4.29)	36.74 (2.41)	35.11 (3.40)	35.79 (3.50)
MMSE	28.88 (1.24)	29.26 (.86)	29.11 (1.63)	29.14 (.97)
Accuracy	.23 (.10)	.23 (.08)	.24 (.10)	.20 (.08)
Intrusions	.69 (1.09)	.70 (.95)	1.44 (1.93)	1.14 (1.65)

Note. Parentheses represent the means' standard deviations, and accuracy reflects the proportion of words that participants successfully recalled. It is necessary to note that the means for education and MMSE score do not account for all participants. Five participants' data are missing responses to the question about education level (one participant is missing from the threat/success cell's education mean; two are missing from that of the neutral/success cell; and two are missing from the threat/no success cell's education mean). Moreover, one threatened participant in the success group did not take the MMSE due to being hard of hearing.

Figure 1

Threatened and Neutral Participants' Responses to the Three Manipulation Check Items



Note. Error bars represent standard error of the mean.

Table 2

Open-Ended Threat Reaction Qualitative Coding Dimensions, Definitions, Example Responses, and Frequency of References as a Function of Stereotype Condition

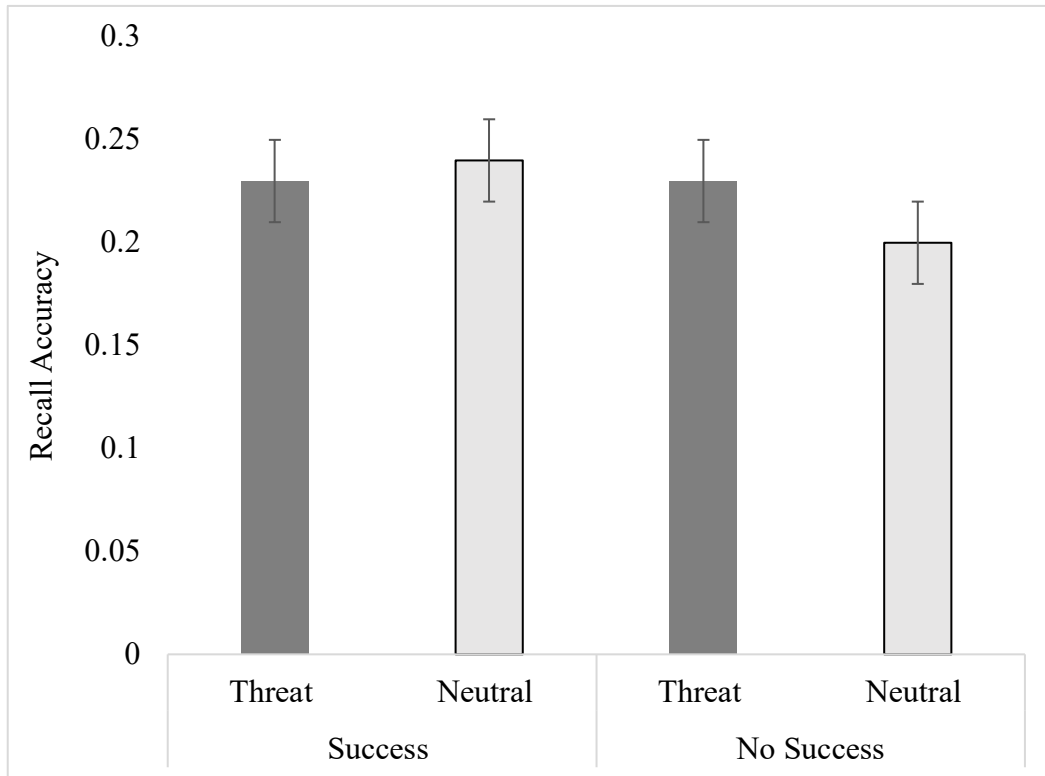
Dimension	Definition	Example	Group Differences
Didn't read/doesn't remember	Indicates having not read the passage or a lack of memory for the passage	"I don't remember reading it."	Threat = Neutral
Neutral	Response reveals no perceived effect of the passage	"No, it didn't impact me."	Threat < Neutral
Acceptance/lack of surprise	Participant indicates acceptance of age-related performance differences and/or mentions that the passage was consistent with their preconceived notions regarding aging	"No—it was no surprise that younger brains have faster recall."	Threat > Neutral
Motivated	Response reveals a sense of motivation to perform well and/or disprove the stereotype and experimenter expectations	"Yes I tried to remember more than I was actually able to remember so that I	Threat = Neutral

		could prove the assumption invalid.”	
Relaxed	Participant indicates feeling calm, relaxed, and/or reassured	“I felt relaxed and reassured.”	Threat = Neutral
Anxiety-inducing	Response reveals increased anxiety	“Yes. I knew it was timed so became nervous.”	Threat = Neutral
Decreased pressure/defeatism	Participant indicates a lack of desire to try hard/succeed	“Maybe. It gave me permission or an expectation to perform poorly. I did not bother to take time to think about a strategy that might enhance my performance, since it doesn’t matter.”	Threat = Neutral
Negative self- expectations	Response reveals that the participant expected they would perform poorly	“It set expectations that, as an older person, I would likely not do well.”	Threat > Neutral

Note. Differences between threatened and neutral participants’ frequency of references to each dimension were significant at a *p*-value lower than .05.

Figure 2

Recall Accuracy as a Function of Stereotype Condition and Prior Task Success Group



Note. Error bars represent standard error of the mean.