Psychology and Aging
2018, Vol. 33, No. 1, 17–29

0882-7974/18/$12.00 http://dx.doi.org/10.1037/pag0000194

The Effects of Stereotype Threat on the Associative Memory Deficit of Older Adults

Matthew S. Brubaker
Springfield College

Moshe Naveh-Benjamin
University of Missouri

The purpose of the current study was to assess whether the age-related associative memory deficit is affected by stereotype threat, which has been shown to negatively affect performance on a wide variety of cognitive tasks, including memory performance of older adults. To date, the effects of stereotype threat on older adults' memory performance have only been shown using tests of item memory, and using only between-subjects manipulations. The first experiment manipulated stereotype threat by providing younger and older adults with either stereotype consistent (threat condition) or stereotype inconsistent (nonthreat condition) information before studying lists of item pairs and being tested for both item and associative memory. The results revealed a triple interaction of Age × Test × Condition (a strong associative memory deficit in the threat condition that disappeared in the nonthreat condition), and this pattern of results was found both between- and within-subjects. A follow-up control experiment also yielded an age-related associative deficit, suggesting that stereotype threat may be activated even under normal testing conditions. These results imply that stereotype threat could affect older adults' associative memory, and that one effective strategy to reduce the associative deficit of older adults is to reverse negative stereotypes that exist.

Keywords: memory, older adults, associative deficit, stereotype threat, item-associative memory

Research indicates that older adults' episodic memory is poorer compared with that of younger adults' (for a review, see Old & Naveh-Benjamin, 2008; Zacks, Hasher, & Li, 2000), suggesting an inevitable decline in memory performance as people age—even among otherwise healthy individuals. Evidence suggests several possible reasons for this decline, including sensory deficits (Lindenberger & Baltes, 1994; Naveh-Benjamin & Kilb, 2014; Schneider, Daneman, & Pichora-Fuller, 2002; Wingfield, Tun, & McCoy, 2005), neuroanatomical changes (Raz, 2000; Raz et al., 2005; Rugg & Vilberg, 2013), and general cognitive deficits such as slower processing rates (Salthouse, 1996), reduced storage capacity of working memory (Foos, 1989), and reductions in attentional resources (Craik & Byrd, 1982). However, these are all degenerative factors that inevitably worsen with age, and because of this it can be difficult to counteract their effects. More recently researchers have also begun looking at the possible role that psychosocial factors, such as stereotype threat, may play in age-related memory deficits.

Stereotype threat occurs whenever individuals feel at risk of confirming negative stereotypes about a group they are a part of and subsequently perform worse on tasks related to that domain. This term was first used by Steele and Aronson (1995) to explain racial differences on tests of intellectual performance, but has been shown to affect performance on a variety of tasks for a variety of different populations. For example, it has also been shown to impair women's math (Spencer, Steele, & Quinn, 1999) and driving (Yeung & von Hippel, 2008) performance compared with men's, older adults' driving performance compared with younger adults' (Lambert et al., 2016; Joanisse, Gagnon, & Voloaca, 2013), science students' intellectual performance compared with psychology students' (Croizet et al., 2004), Hispanics' intellectual performance compared with Caucasians' (Schmader & Johns, 2003), Caucasian men's math performance compared with Asian men's (Aronson et al., 1999), and Caucasian men's athletic performance compared with African American men's (Stone, Lynch, Sjomeling, & Darley, 1999).

In the last two decades, research has also started to look at the effect of stereotype threat on older adults' memory performance. Negative views on aging are prevalent (see the review by Hummert, 1999)—as people tend to associate aging with memory impairments (Hummert, Garstka, Shaner, & Strahm, 1994)—and it

Editor's Note. Ulrich Mayr served as the action editor for this article.—UM

Matthew S. Brubaker, Psychology Department, Springfield College; Moshe Naveh-Benjamin, Department of Psychological Sciences, University of Missouri.

Some of the ideas and data appearing in this article have been previously disseminated at the following professional conferences: the annual meeting of the Psychonomic Society in Long Beach, CA (November 2014) and the Cognitive Aging Conference in Atlanta, GA (April 2016).

We thank Nelson Cowan, Bruce Bartholow, and Linda Day for their helpful comments and suggestions. We also thank members of the Memory and Cognitive Aging Laboratory, including past members Angie Kilb, Andrea Smyth, Dwight Peterson, and Hope Fine, and current member Sanchita Gargya, for offering advice and feedback.

Correspondence concerning this article should be addressed to Matthew S. Brubaker, Psychology Department, Springfield College, 263 Alden Street, Springfield, MA 01109, or to Moshe Naveh-Benjamin, Department of Psychological Sciences, University of Missouri, 106 McAlester Hall, Columbia, MO 65211. E-mail: mbrubaker@springfield.edu or navehbenjaminm@missouri.edu

is possible that these views may be helping to exacerbate age-related memory declines. To test this, Hess, Auman, Colcombe, and Rahhal (2003) had older and younger adults read fake newspaper reports either confirming age-related declines in memory (negative view of aging) or disconfirming them (positive view of aging), before being given a free-recall memory task. Their results showed no effect of condition (negative vs. positive) on the younger adults, but the older adults recalled significantly more items in the positive condition than they did in the negative condition—suggesting that stereotype threat does negatively affect older adults' memory performance. However, a control condition in which participants did not read any newspaper reports showed memory results more in line with the positive group as opposed to the negative group—suggesting that stereotype threat may only affect memory performance when explicitly induced.

The way in which stereotype threat is manipulated in older adults has been varied. Some studies, such as the one by Hess et al. (2003), utilize fact-based manipulations to induce stereotype threat by explicitly mentioning age-related differences in memory (also see Eich, Murayama, Castel, & Knowlton, 2014; Hess, Hinson, & Hodges, 2009). Other studies utilize stereotype-based manipulations, which can be more subtle, such as having participants report their age before beginning the experiment (Kang & Chasteen, 2009) or merely stating that there are both younger and older adults participating in the experiment (Mazerolle, Regner, Morisset, Rigalleau, & Huguet, 2012). In all of these cases participants know that they are being tested on their memory. However, another way to induce stereotype threat in older adults is by simply stating that the test measures memory performance, as opposed to deemphasizing the memory component of the task by explaining that it is measuring something entirely different (e.g., Rahhal, Colcombe, & Hasher, 2001).

These subtle manipulations, in particular, highlight the notion that stereotype threat for older adults may be occurring even under normal testing situations, similar to how Steele (1997) explained that threat was operating in younger adults. For example, when older adults come into a lab to participate in a memory experiment, stereotype threat may already be activated simply by knowing that they are being tested on their memory performance. This automatic activation of stereotype threat would cause the age differences seen in most research findings to be exacerbated. The logical conclusion from this line of reasoning would be that one way to improve older adults' memory performance would be to get rid of the stereotype threat that is automatically occurring. Since stereotype threat is not based on age-related biological or cognitive deficits, it may be easier to manipulate and create lasting change in older adults' memory performance than by some of the other factors mentioned earlier. And being able to reduce stereotype threat in older adults may be of great practical significance. In a recent study, 70% of older adults scored below the cutoff for dementia when assessed under stereotype threat, as opposed to only 14% when assessed under a no-threat condition (Haslam et al., 2012).

While many studies have shown effects of stereotype threat on older adults' memory performance (see meta-analysis by Lamont, Swift, & Abrams, 2015), several contradictions still exist within the literature. For instance, while stereotype threat has been shown to increase older adults' false alarms to related lures when using a classic Deese-Roediger-McDermott (DRM) procedure (Thomas & Dubois, 2011), it has also been shown more recently to *decrease* these same critical errors (Wong & Gallo, 2016). The only major difference between these two studies was that Wong and Gallo warned participants not to fall victim to the DRM illusion before the test phase began. In addition, while some research shows a decrease in working memory capacity under stereotype threat (Jordano & Touron, 2017; Mazerolle et al., 2012; Schmader & Johns, 2003), other studies do not (Barber & Mather, 2013a; Hess, Hinson, et al., 2009).

These contradictions have led to several competing theories to explain the underlying mechanism(s) responsible for this effect—including the regulatory focus theory (see the review by Barber, 2017) and the executive control interference hypothesis (Schmader & Johns, 2003)—which will be the focus of the current study. The executive control interference hypothesis states that stereotype threat focuses one's attention on the stereotype, using up working memory resources that would otherwise be applied to the primary task, thus resulting in lowered performance due to lowered working memory capacity. Schmader and Johns (2003) tested this in younger adults by showing that stereotype threat did indeed lead to lower working memory performance (using an operation span task) for women when the task was described as diagnostic of math ability (Experiment 1) and Hispanics when the task was described as diagnostic of intelligence (Experiment 2). A third experiment showed that stereotype threat still lowered working memory performance even when the working memory task was separate from the stereotyped task and not linked to the same domain.

Other evidence also fits with the executive control interference hypothesis. For instance, Croizet et al. (2004) found that stereotype threat decreased heart rate variability in participants and lowered task performance. They argued that the decreased heart rate variability implies increased cognitive load, which interferes with performance for the stereotyped individuals. Mazerolle et al. (2012) found that older adults under stereotype threat had lower working memory scores than those not under threat. Also, Hess et al. (2003), as mentioned earlier, found that stereotype threatened older adults not only recalled fewer words from the initial study list, but also had lower clustering scores for the related words in the list. The lower clustering scores suggest that stereotype threat undermines one's ability to use effective strategies, which would make sense if one's working memory resources were impaired. It should be noted, though, that this clustering effect has not been replicated (e.g., Hess & Hinson, 2006; Hess, Hinson, & Statham, 2004).

To date, all episodic memory-based stereotype threat studies have only tested variations of item memory performance, and never in comparison to associative memory performance.[1] While there is both anecdotal and empirical evidence that episodic memory declines with age, not all episodic memory tasks show a similar age-related decline. For example, Chalfonte and Johnson

[1] It should be noted, however, that Rahhal et al. (2001) studied the effect of stereotype threat on memory for trivia statements that were either labeled as true or false or not labeled. Correct recall required not just remembering whether a statement was old or new, but also what label—if any—it had been given at study (which was in essence, an associative memory test). While they did not have a true measure of item memory, their results did show that stereotype threat reduced older adults' memory for which statements had been labeled true or false while actually increasing memory for which statements were unlabeled.

(1996) found a larger age-related memory difference for bound features than for the individual features themselves. In order to account for these task performance differences, Naveh-Benjamin (2000) suggested the associative deficit hypothesis which states that a major reason for older adults' episodic memory decline is a problem they have in binding separate items of information into cohesive units and later retrieving these bound associations. He tested this hypothesis by giving older and younger adults a series of word/nonword pairs (Experiment 1), unrelated word pairs (Experiment 2), and word/font pairs (Experiment 3) to study, and then tested them over their memory for the individual items (item memory) and for the pairings (associative memory). Results supported the hypothesis, showing that there were larger age differences in memory for the associations than for the items. These findings have since been found to be consistent across different types of stimuli and associations (see the meta-analysis by Old & Naveh-Benjamin, 2008).

One underlying reason that could be causing the associative memory deficit is older adults' impairment in the utilization of encoding and retrieval strategies which could impair their ability to bind items together in memory and retrieve these bound episodes. Such a suggestion is supported by several lines of empirical evidence. For instance, whereas older adults show no differences in associative memory performance whether they are told about the test ahead of time (intentional learning) or not (incidental learning), younger adults show a marked boost in associative memory performance under intentional learning instructions (e.g., Naveh-Benjamin, 2000, Exp. 2). This suggests that younger adults may be using strategies more effectively than older adults since their performance increases when they are made aware of the associative test ahead of time. More direct evidence, though, comes from Naveh-Benjamin, Brav, and Levy (2007) who were able to successfully reduce older adults' associative deficit of unrelated word-pairs by encouraging all participants to use an elaborative strategy (e.g., to form a sentence using both words) during both study and test. The minimal performance gain seen in younger adults in this case suggests that they were already utilizing an appropriate strategy even when not told to do so. In contrast, the significant boost in older adults' associative memory performance when they were told to utilize the strategy suggests that they were not using it as much when not told to do so.

Furthermore, older adults' decline in associative memory may also be due to an age-related decline in recollection (Hay & Jacoby, 1999). It has been shown that younger adults use recollection during memory tests more than older adults do (Jennings & Jacoby, 1993, 1997). In a typical item memory test, one can rely on either familiarity or recollection to determine whether the item is old or new. However, in the associative memory test, in order to get a correct response one must inhibit item familiarity and rely on recollection (recalling whether the two old items now shown at test were originally paired together at study or whether they were each paired with a different item). Continuing to rely on familiarity of the components in the associative test would lead to increased false alarms since all of the individual items are old. This is exactly the pattern of results that is typically seen when analyzing the data—older adults show more false alarms in the associative than in the individual item test compared with younger adults, leading to the overall associative deficit (e.g., Kilb & Naveh-Benjamin, 2011). Other measures, such as the remember/know procedure, have also been used to support the notion that older adults rely more on familiarity than recollection (e.g., Bastin & Van der Linden, 2003). Overall, the executive control interference hypothesis, mentioned earlier, would predict that stereotype threat would especially affect older adults' memory for associations, compared with items, as the encoding of associations and their retrieval via recollection-based processes require more executive control strategies than those involved in item memory.

The current experiments will assess the effects of stereotype threat on associative memory compared with item memory in both younger and older adults, which has not been looked at previously, to see whether stereotype threat affects associative memory more than item memory and if so, whether this effect is helping cause or augment the age-related associative deficit.

Experiment 1

The goal of the first experiment was to manipulate stereotype threat in order to assess its effects on the age-related associative deficit. Also, in addition to using a between-subject manipulation of stereotype threat, this is the first study to use a within-subject manipulation, as well. No studies have yet looked at the effects of stereotype threat on older adults' memory performance using a within-subject design (although Spencer et al., 1999, Experiment 2, used a within-subject design of stereotype threat on women's math ability). Previous studies using between-subject designs have demonstrated that stereotype threat effects on memory and aging do exist. But a within-subject design would show whether stereotype threat can be manipulated within the same individual, simply by changing one aspect of the task. It is also a stronger manipulation as it controls for effects of individual differences that exist in between-subject design. If stereotype threat truly is "situational" in nature, as Steele (1997) claimed, meaning that it is only activated under salient circumstances, then stereotype threat could be manipulated (activated and deactivated) within the same participant simply by changing the context of the situation.

In order to form a convincing argument for the participants that the context of the situation has indeed changed, the current experiment used two different types of stimuli—verbal and visual, one for the stereotype threat condition, and the other for the nonstereotype threat condition, with the type of stimuli counterbalanced across stereotype threat conditions. Participants were told that the two different stimuli have been shown to have different effects on age-related memory performance. This allowed us to use the same item-associative test paradigm for a direct comparison of both conditions within subjects.

Method

Participants. Sixty younger adults (ages 18–25) and 66 older adults (ages 65–87) participated in Experiment 1 (however, six older adults were dropped from the final analysis due to not remembering the instructions used for creating the stereotype threat manipulation on a subsequent posttest questionnaire). The younger adults were recruited from psychology classes at the University of Missouri and received course credit for participating. The older adults were recruited from the local community in central Missouri and received $15 for participating. All participants reported being in good physical and mental health, with no

Table 1
Demographic Information for Both Experiments

Age condition	N	Proportion male	Age (SD)	Education (SD)
Experiment 1				
Young	60	.27	19.50 (1.54)	13.37 (1.30)
Old	60	.23	72.52 (6.29)	14.70 (1.95)
Experiment 2				
Young	40	.28	18.84 (1.26)	12.34 (.94)
Old	40	.30	73.95 (6.01)	15.60 (2.23)

major vision, hearing, or cognitive impairments. As in many age-related studies, older adults had significantly more years of formal education than did the younger adults, $t(118) = 4.40$, $p < .001$, (Table 1 displays additional demographic information).

Design. Experiment 1 used a 2 (age: young vs. old) × 2 (test: item vs. associative) × 2 (stereotype threat: threat vs. nonthreat) mixed factorial design. Age was a between subjects factor, while test and stereotype threat were manipulated within subjects.

Stimuli. Four study lists were used, each consisting of 26 pairs of stimuli (the first and last pair in each list were buffers). Two of the study lists used word pairs as stimuli, and the other two study lists used face-scene picture pairs as stimuli (Figure 1 shows an example of each). The word pairs were all made up of two-syllable, high frequency nouns. Words within a given pair were unrelated. The faces all had neutral expressions, consisting of an equal number of both males and females and older and younger adults, and were taken from Minear and Park's (2004) life span database. The scenes were all outdoor scenic photos consisting of various landscape and urban settings, and were taken from Brubaker and Naveh-Benjamin (2014). The reason for choosing these two different sets of stimuli is that both unrelated word pairs (e.g., Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2007) and face/scene picture pairs (e.g., Kilb & Naveh-Benjamin, 2011) have been used in the past to show age-related associative deficits. In addition, using two sets of stimuli allows for a within-subject manipulation of stereotype threat and should also add to the generalizability of the results.

Item tests contained 32 items (16 targets, 16 distractors), and associative tests contained 16 pairs (8 intact, 8 recombined). Targets were items that were shown during the study list, while distractors were new items that were not shown before. Intact pairs were made up of items that had been paired together during the study phase, while recombined pairs were made up of previously studied items that had not been paired together during the study phase. Each studied stimulus appeared either in the item or the associative test.

Procedure. This research received Institutional Review Board approval from the University of Missouri, and all participants were tested under the ethical guidelines set by the American Psychological Association. Participants received four study-test blocks, each consisting of a study list (26 pairs) followed by a recognition test of the individual items (16 old items, 16 new items) and a recognition test of the pairs (8 intact pairs, 8 recombined pairs). Type of test was counterbalanced so that half of the participants always received the item test first and the other half always received the associative test first. The presentation rate at study was 4,000 ms per pair with a 500 ms interstimulus interval. All test events were experimenter paced with a response deadline of 4,000 ms (shown in the past to provide ample time for younger and older adults to respond to both tests). As an interpolated activity between the study and the test phases, participants were instructed to count backward by threes from a random three-digit number for 30 s. Two of the blocks consisted of the face/scene picture pairs, while the other two consisted of the unrelated word pairs (this order was counterbalanced so that half of the participants received the word pairs for the first two blocks, and the other half received the face/scene picture pairs for the first two blocks). This allowed the manipulation of stereotype threat within-subject with half of the participants receiving the threat condition first and the nonthreat condition second, and the other half receiving the nonthreat condition first and threat condition second (this was completely counterbalanced with the order of stimulus type). This design allowed the results to be analyzed both within subjects (by looking at all four blocks) and between subjects (by looking at only the first two blocks for each participant).

Participants were told that they would be completing two separate experiments (each including 2 study blocks), one involving word-pairs and the other involving face/scene pictures pairs. Intentional learning instructions were given before the first experiment, and again before the second experiment. Similar to the stereotype threat manipulation used by Hess, Emery, and Queen (2009), participants in the threat condition were read a paragraph explaining that the purpose of the current experiment is to examine age differences in memory ability, that the next two blocks will be examining memory, and that age-related declines in memory performance have been shown in the past when using the current set of stimuli (either words or faces/scenes). For the nonthreat condition, participants were read a paragraph (see the Appendix for exact wording of the instructions) explaining that the purpose of the experiment is to examine individual differences in performance, that the next two blocks will be examining verbal (word pairs) or spatial (face/scene pairs) processing (each to half of the participants), and that no age differences have been shown in the past when using the current set of stimuli (either words or faces/scenes—again, each to half of the participants). Also, in order to further separate the first two blocks from the last two blocks, a picture-rating task was given in between the two experiments in which participants were shown a series of line drawings and asked to rate how pleasant the drawings were on a scale of 1–9.

In order to evaluate how participants' performance expectations change as a function of the threat manipulation, a performance ex-

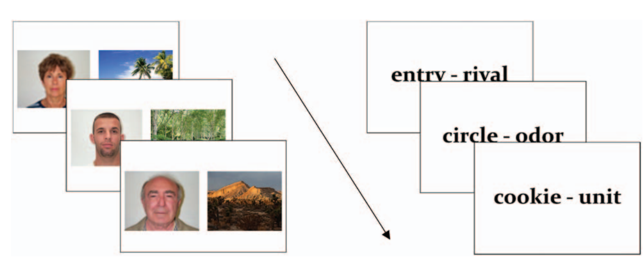

Figure 1. Example of stimuli used in both experiments. Two of the blocks used face-scene picture pairs, and the other two blocks used unrelated word pairs. Photos of faces are from Minear and Park's (2004) database. See the online article for the color version of this figure.

pectation question was included for both sets of stimuli. Derived from Desrichard and Kopetz (2005), the question was given immediately following the instructions and practice phase, and asked how participants expected to perform on both the item and associative memory tests using a 9-point response scale (1 = *very poorly*, 9 = *very good*).

Results

Within-subject analyses. Memory accuracy was measured in proportion hits minus proportion false alarms. Because each of the two types of stimuli showed a similar pattern of results when analyzed separately (including an Age × Test interaction and an Age × Test × Threat interaction), for the purpose of the main analysis, the two different types of stimuli were combined to form an overall item memory score and an overall associative memory score (Figure 2).

A three-way analysis of variance (ANOVA) was run, with type of test (item vs. associative) and stereotype threat condition (threat vs. nonthreat) as within-subject factors, and age (old vs. young) as a between-subject factor. A significant main effect of test showed that item memory performance (M = .63, SD = .14) was better than associative memory performance (M = .50, SD = .22); and a significant main effect of age, $F(1, 118) = 7.73$, $p < .01$, $\eta_p^2 =$ .06, showed that younger adults (M = .61, SD = .25) performed better than older adults (M = .52, SD = .25). Overall, the age-related associative deficit was replicated with an Age × Test interaction, $F(1, 118) = 8.73$, $p < .01$, $\eta_p^2 = .07$, showing significantly better performance in the associative test for younger adults (M = .56, SD = .24) than for older adults (M = .43, SD = .21), $t(118) = 3.14$, $p < .01$, $d = .57$, yet only marginal age differences in the item test (M = .65, SD = .16 for younger adults, and M = .60, SD = .14 for older adults), $t(118) = 1.82$, $p = .07$, $d = .34$.

To the main question of whether or not stereotype threat affects the age-related associative deficit, a significant Age × Test × Threat interaction was also found, $F(1, 118) = 5.68$, $p = .02$, $\eta_p^2 =$ .05, suggesting that at different levels of stereotype threat there is a difference in the age-related associative deficit.[2] This triple interaction was broken down further by looking separately at performance in the threat and nonthreat conditions. There was an Age × Test interaction in the threat condition, $F(1, 118) = 13.81$, $p < .001$, $\eta_p^2 = .11$, which was driven mainly by the associative test. Younger adults had significantly higher performance in the associative test (M = .58, SD = .27) than older adults (M = .39, SD = .26), $t(118) = 3.87$, $p < .001$, $d = .71$, yet there were no age differences in the item test (M = .64, SD = .19 for younger adults, and M = .61, SD = .17 for older adults), $t(118) = .99$, $p = .32$, $d = .18$. Conversely, in the nonthreat condition no Age × Test interaction was found, $F(1, 118) = 0.00$, $p = .997$, $\eta_p^2 = .00$.

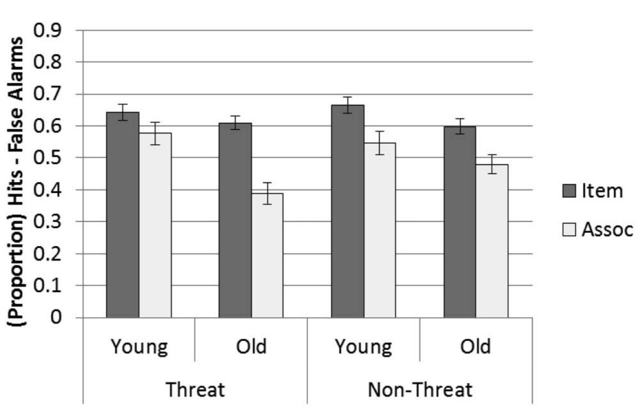

Figure 2. Proportion hits minus proportion false alarms as a function of age, test, and stereotype threat condition in Experiment 1 (error bars represent standard errors around the mean).

As further evidence that stereotype threat affected associative memory but not item memory, older adults' associative memory was significantly worse in the threat condition (M = .39, SD = .26) than the nonthreat condition (M = .48, SD = .22), $t(59) =$ 2.74, $p < .01$, $d = .35$, whereas older adults' item memory performance (M = .61, SD = .19 for threat; M = .60, SD = .17 for nonthreat) did not change as a function of stereotype threat condition, $t(59) = .40$, $p = .69$, $d = .05$. Younger adults showed no difference in either condition.

Hit and false alarm rates were also analyzed separately in the same manner as the hits minus false alarms. The hit rates showed no significant main effects or interactions. However, the false alarm rates produced results in line with the overall performance, including significant main effects of test, $F(1, 118) = 164.96$, $p <$ .001, $\eta_p^2 = .58$, age, $F(1, 118) = 11.32$, $p < .001$, $\eta_p^2 = .09$, and an Age × Test interaction, showing a higher rate of false alarms in the associative test for older adults (M = .34, SD = .21) than for younger adults (M = .23, SD = .21), $t(118) = 4.29$, $p < .001$, $d =$ .79, yet no age differences in the item test (M = .17, SD = .14 for older adults, and M = .14, SD = .14 for younger adults), $t(118) =$ 1.30, $p = .20$, $d = .24$.

Most importantly the Age × Test × Threat interaction was significant, $F(1, 118) = 5.76$, $p = .02$, $\eta_p^2 = .05$. Post hoc analysis revealed that older adults' associative false alarms were significantly higher in the threat condition (M = .38, SD = .18) than the nonthreat condition (M = .31, SD = .19), $t(59) = 3.08$, $p < .01$, $d = .40$, whereas older adults' item false alarms (M = .17, SD = .11 for threat; M = .17, SD = .12 for nonthreat) did not change as a function of stereotype threat condition, $t(59) = .28$, $p = .78$, $d =$ .04. Younger adults showed no difference in either condition.

Between-subject analyses. Because of the way Experiment 1 was designed, between-subject effects of stereotype threat can also be looked at by only analyzing the first two blocks (the first stereotype threat condition) that each participant received (Figure 3). A three-way ANOVA using the hits minus false alarm measure was run on only the first condition (i.e., the first 2 blocks), with type of test (item vs. associative) as a within-subject factor, and age (old vs. young) and stereotype threat condition (threat vs. nonthreat) as between-subject factors. In line with the within-subject results, significant main effects of test, $F(1, 116) = 54.23$, $p < .001$, $\eta_p^2 = .32$, and age, $F(1, 116) = 7.32$, $p < .01$, $\eta_p^2 = .06$, were found, as well as a significant Age × Test interaction, $F(1, 116) = 3.95$, $p = .049$, $\eta_p^2 = .03$ indicating an age-related associative deficit. To the main question of whether or not stereotype threat affects the age-related associative deficit, a significant

[2] A signal detection analysis confirmed that the significant triple interaction of age, test, and threat was shown for memory sensitivity (d') but not for response bias (C). This analysis also indicated a slightly more liberal bias for older adults in the stereotype threat condition relative to the nonthreat condition, though this was not significantly different in comparison to younger adults.

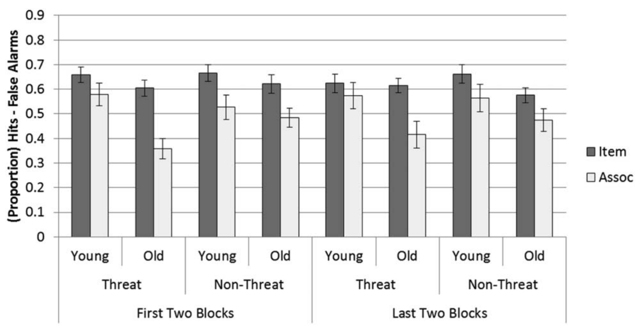

Figure 3. Separate proportion hits minus proportion false alarms for the first two blocks and last two blocks (first and second tested conditions) as a function of age, test, and stereotype threat condition in Experiment 1 (error bars represent standard errors around the mean).

Age × Test × Threat interaction was found, $F(1, 116) = 4.20$, $p = .043$, $\eta_p^2 = .04$.[3] This triple interaction was broken down further by looking separately at performance in the threat and nonthreat conditions. There was a significant Age × Test interaction in the threat condition, $F(1, 58) = 8.26$, $p < .01$, $\eta_p^2 = .13$, which was driven mainly by the associative test. Younger adults had significantly higher performance in the associative test ($M = .58$, $SD = .26$) than older adults ($M = .36$, $SD = .22$), $t(58) = 3.55$, $p < .001$, $d = .93$, yet there were no age differences in the item test ($M = .66$, $SD = .17$ for younger adults, and $M = .60$, $SD = .18$ for older adults), $t(58) = 1.21$, $p = .23$, $d = .32$. Conversely, in the nonthreat condition no Age × Test interaction was found, $F(1, 58) = 0.002$, $p = .96$, $\eta_p^2 = .00$.

As further evidence that stereotype threat affected associative memory rather than item memory, older adults' associative memory was significantly worse in the threat condition ($M = .36$, $SD = .22$) than the nonthreat condition ($M = .48$, $SD = .21$), $t(58) = 2.23$, $p = .03$, $d = .29$, whereas older adults' item memory performance ($M = .62$, $SD = .20$ for threat; $M = .60$, $SD = .18$ for nonthreat) did not change as a function of stereotype threat condition, $t(58) = .34$, $p = .74$, $d = .04$. Younger adults showed no difference in either condition.

Hit and false alarm rates were also analyzed separately in the same manner as the hits minus false alarms. The hit rates showed no significant main effects or interactions. However, the false alarm rates produced results in line with the overall performance, including significant main effects of test, $F(1, 116) = 91.83$, $p < .001$, $\eta_p^2 = .44$, age, $F(1, 116) = 11.95$, $p < .001$, $\eta_p^2 = .09$, and an Age × Test interaction, $F(1, 116) = 10.42$, $p < .01$, $\eta_p^2 = .08$, showing a higher rate of false alarms in the associative test for older adults ($M = .35$, $SD = .18$) than for younger adults ($M = .23$, $SD = .15$), $t(118) = 4.10$, $p < .001$, $d = .76$, yet no age differences in the item test ($M = .16$, $SD = .12$ for older adults, and $M = .14$, $SD = .12$ for younger adults), $t(118) = 1.25$, $p = .22$, $d = .23$. Most importantly, there was also a significant Age × Test × Threat interaction, $F(1, 116) = 4.07$, $p = .046$, $\eta_p^2 = .03$. Post hoc analysis revealed that older adults' associative false alarms were significantly higher in the threat condition ($M = .40$, $SD = .15$) than the nonthreat condition ($M = .30$, $SD = .19$), $t(58) = 2.14$, $p = .037$, $d = .28$, whereas older adults' item false alarms ($M = .17$, $SD = .12$ for threat; $M = .16$, $SD = .13$ for nonthreat) did not change as a function of stereotype threat condition, $t(58) = .30$, $p = .77$, $d = .04$. Younger adults showed no difference in either condition (see Table 2 for separate hit and false alarm rates).

Performance expectation ratings. Expectation ratings (Table 3) assessing how well participants expected to perform on both the item and the associative memory tests were taken after participants were given instructions and the stereotype threat manipulation but before beginning the first and third blocks. To determine whether expectation ratings were in line with the overall memory performance and could be a mediating factor, a similar three-way ANOVA was run with type of test (item vs. associative) and stereotype threat condition (threat vs. nonthreat) as within-subject factors, and age (old vs. young) as a between-subject factor. A significant main effect of test, $F(1, 118) = 79.00$, $p < .001$, $\eta_p^2 = .40$, showed that the expectations regarding item memory performance ($M = 6.33$, $SD = 1.17$) were higher than those for associative memory performance ($M = 5.39$, $SD = 1.42$). Importantly, no Age × Test × Threat interaction was found, $F(1, 118) = 2.11$, $p = .15$, $\eta_p^2 = .02$.

However, when analyzing the stereotype threat condition as a between-subject factor, a somewhat different pattern of results emerged. When only looking at the first condition (the first 2 blocks), a significant Age × Test × Threat interaction was found, $F(1, 116) = 4.06$, $p = .046$, $\eta_p^2 = .03$. This triple interaction was broken down further by looking separately at the expectation ratings of younger adults and older adults. There was a Test × Threat interaction in older adults, $F(1, 58) = 6.16$, $p = .016$, $\eta_p^2 = .10$, which was driven mainly by the associative test. Older adults in the threat condition ($M = 4.10$, $SD = 1.56$) expected their associative memory performance to be significantly lower than those in the nonthreat condition ($M = 5.43$, $SD = 1.72$), $t(58) = 3.15$, $p < .01$, $d = .41$. However, there were no differences in older adults' expectation ratings for the item test as a function of stereotype threat ($M = 5.47$, $SD = 1.41$ for the threat condition, and $M = 6.00$, $SD = 1.46$ for the nonthreat condition), $t(58) = 1.44$, $p = .16$, $d = .19$. Conversely, younger adults showed no differences in performance expectations as a function of stereotype threat, $F(1, 58) = .28$, $p = .60$, $\eta_p^2 = .005$. Interestingly, when only looking at the second condition (the last two blocks), no Age × Test × Threat interaction was found, $F(1, 116) = .22$, $p = .64$, $\eta_p^2 = .002$.

Discussion

The results of this experiment show an age-related associative deficit when participants are given information consistent with memory-based stereotypes. However, when participants are given information inconsistent with memory-based stereotypes, the associative deficit disappears. Similar results were found within subjects (where each participant participated in both the threat and nonthreat conditions) and between subjects (where only the first condition for each participant was used—half being threat and half being nonthreat conditions). This is the first study to show an

[3] In contrast, when analyzing the last two blocks—the second condition only—the triple interaction of Age × Test × Threat was only marginally significant, $F(1, 116) = 2.75$, $p = .10$, $\eta_p^2 = .02$, suggesting a slightly reduced effect of stereotype threat most likely due to carry-over effects of the first condition.

Table 2
Proportion Hits, and Proportion False Alarms Separately, as Well as Proportion of Hits Minus False Alarms as a Function of Age, Test, and the First Stereotype Threat Condition for Both Experiments

Test and age conditions	Experiment 1						Experiment 2		
	Threat			Nonthreat			Control		
	H	FA	H-FA	H	FA	H-FA	H	FA	H-FA
Item test									
Young									
Mean	.79	.13	.66	.81	.14	.67	.79	.14	.65
SD	.12	.12	.17	.12	.13	.19	.14	.08	.19
Old									
Mean	.77	.17	.61	.78	.16	.62	.81	.17	.64
SD	.12	.12	.18	.13	.13	.20	.08	.10	.14
Associative test									
Young									
Mean	.80	.22	.58	.77	.25	.53	.76	.22	.54
SD	.16	.12	.26	.15	.17	.27	.16	.13	.24
Old									
Mean	.76	.40	.36	.79	.30	.48	.78	.33	.44
SD	.16	.15	.22	.12	.19	.21	.14	.18	.26

Note. H = proportion hits; FA = proportion false alarms.

effect of stereotype threat on older adults' associative memory in comparison to item memory, and the first to show any effect of stereotype threat on memory within subjects. The fact that stereotype threat had an effect on associative memory but not item memory in Experiment 1 is in line with the executive control interference hypothesis and the theory that stereotype threat operates by undermining strategic binding and/or recollection as opposed to automatic encoding and/or familiarity processes.

Interestingly, Experiment 1 showed no effect of stereotype threat on item memory performance, neither for older nor younger adults. However, most age-based stereotype threat studies have not found changes in item recognition, and previous studies that have shown effects of stereotype threat on older adults' item memory performance have typically used either free-recall measures (e.g., Hess et al., 2003; Hess, Hinson, et al., 2009) or recognition tests with high performance constraints such as a limited response deadline (e.g., Hess, Emery, et al., 2009). Free-recall memory, as opposed to recognition memory, involves greater effort and strategic demands by requiring recollection of specific items from memory. Item recognition memory, on the contrary, allows for familiarity to be utilized in assessing whether or not an item was shown before. Also, in line with the executive control interference hypothesis (Schmader & Johns, 2003) that states that stereotype threat negatively affects cognitive load, tasks with higher task demands should be more adversely affected by stereotype threat than tasks with lower task demands. And Hess, Emery, et al. (2009) were only able to show an effect of stereotype threat on older adults' item recognition memory when a strict response deadline of 2,000 ms was used at test—suggesting that item recognition performance is only affected by stereotype threat when the task is perceived as being difficult. Previous studies in other domains (e.g., math performance) have also shown that stereotype threat has a greater effect on more difficult tasks (Kimball, 1989; Spencer et al., 1999). Due to the fact that the current experiment used a recognition memory paradigm with a more lenient response deadline of 4,000 ms, the null effect of stereotype threat on older adults' item memory in Experiment 1 can be considered in line with previous results. One could argue that if stereotype threat affects recollection and since a shorter response deadline mostly eliminates recollection to begin with (e.g., Sauvage, Beer, & Eichenbaum, 2010), any effect of stereotype threat should be stronger when using a longer response deadline as opposed to a shorter response deadline. However, an item recognition task does not necessarily require the use of recollection (whereas an associative recognition task does), so the response deadline affecting recollection is inconsequential to item memory performance—as evidenced by the fact that Hess, Emery, et al. (2009) found no memory differences between their 2-s deadline and unlimited deadline conditions when stereotype threat was not used.

The overall effect in Experiment 1 was driven by older adults' increased associative false alarms in the threat condition. In fact, the same triple interaction of Age × Test × Threat appeared when analyzing just the false alarms rates, similar to the overall hits-

Table 3
Mean Expectation Ratings as a Function of Age, Test, and Stereotype Threat Condition (With Standard Deviations in Parentheses) in Experiment 1

Age and test conditions	First two blocks		Last two blocks	
	Threat	Nonthreat	Threat	Nonthreat
Old				
Item (*SD*)	5.47 (1.41)	6.00 (1.46)	5.70 (1.62)	5.70 (1.47)
Assoc (*SD*)	4.10 (1.56)	5.43 (1.72)	4.67 (1.71)	4.60 (1.71)
Young				
Item (*SD*)	6.70 (1.18)	7.10 (.85)	7.43 (1.22)	6.53 (1.17)
Assoc (*SD*)	6.03 (1.35)	6.23 (1.52)	6.40 (1.85)	5.67 (1.21)

Note. Assoc = associative.

minus-false alarms analysis. The fact that the effect of stereotype threat was driven exclusively by an increase in associative false alarms for older adults is consistent with the results reported by Thomas and Dubois (2011), who showed an increase in false alarms—particularly for related lures—in a DRM paradigm when under stereotype threat. It also supports the notion that stereotype threat reduces strategic binding at encoding and strategic recollection. For example, poorer binding of associations at encoding may not necessarily result in lower hit rates in older adults' associative memory since intact pairs from the encoding phase may be endorsed based on either the recollection of the pair or on the familiarity of the two item components, with older adults' poorer recollection possibly compensated by strong effects of the components familiarity. However, poorer binding at encoding along with poorer recollection, in line with the executive control interference hypothesis, could make older adults' overreliance on the two components' familiarity in the absence of recollection leading to an increased false recognition of recombined pairs—yet it would not negatively affect older adults' item memory. This increased rate of associative false alarms for associations has also been implicated as a cause for the age-related associative deficit (e.g., Kilb & Naveh-Benjamin, 2011), providing further evidence that stereotype threat may partially be driving the associative deficit even under normal testing situations.

Experiment 1 showed an effect of the stereotype manipulation on age and type of test, and this effect appears to be mediated by personal expectations. Whereas younger adults' expectation ratings for both the item and associative tests did not change as a function of the threat condition, older adults' expectation ratings for the associative test were significantly reduced in the threat condition compared with the nonthreat condition whereas older adults' expectation ratings for the item test did not significantly change as a function of the threat condition (at least based on the between-subject analyses). These expectation ratings are in line with the overall memory performance, suggesting a potential mediating effect of expectations. Other studies investigating stereotype threat in memory and aging have shown similar results (Desrichard & Kopetz, 2005; Hess, Hinson, et al., 2009).

Experiment 2

The results from Experiment 1 indicate that a manipulation of stereotype threat can result in a large age-related associative deficit, and it raises the possibility that stereotype threat is one of the underlying factors driving the associative deficit typically seen in older adults (e.g., Old & Naveh-Benjamin, 2008). In fact, older adults may be operating under stereotype threat in most laboratory settings—even if differences in age-related memory performance are not explicitly mentioned. Perhaps simply by being in a memory lab and knowing that they are being compared with younger adults, stereotype threat is already activated in older adult participants. However, it is also possible that stereotype threat may not be activated unless it is explicitly induced (like in the threat condition of Experiment 1 where stereotype consistent information was given to participants ahead of time), and that the presence of stereotype threat augments the age-related associative deficit, but is not responsible for it under normal testing situations. This second suggestion seems somewhat unlikely, however, considering how closely the results from the threat condition in Experiment 1 mirror those typically seen when comparing item and associative memory (Old & Naveh-Benjamin, 2008).

Due to the lack of a true control group in Experiment 1, though, no direct conclusions can be made as to whether or not stereotype threat is activated in normal testing situations when no information about age-related stereotypes is explicitly mentioned. Hess et al. (2003) included a control group that was not given any information to read about memory and aging prior to being tested (unlike the positive and negative groups in the same study). While the control group's free recall memory performance was similar to that of the positive group—suggesting that stereotype threat was not activated unless explicitly induced—there was other evidence to indicate that stereotype threat was activated in the control group. For instance, after the stereotype manipulation but before the recall task, all participants were given a stereotype activation task (derived from Banaji & Hardin, 1996) in which they classified positive and negative trait words, and a prime word of "old" or "young" was presented immediately before some of the trials. As expected, participants in the negative group were faster at identifying negative traits and slower at identifying positive traits that followed the old prime than were participants in the positive group. However, control group participants showed similar results on this task to the negative group, not the positive group, providing evidence that the normal lab setting might be enough to activate stereotype threat. Interestingly, the Hess et al. (2003) study is one of the few memory-based stereotype threat studies to include a true control group.

Therefore, the purpose of Experiment 2 was to run a control condition where no information about age-related differences in memory performance is explicitly stated. If stereotype threat is naturally induced in regular laboratory settings, then the results of Experiment 2 should be similar to the threat condition in Experiment 1. However, if stereotype threat is only activated by explicitly mentioning age-related differences in memory performance, then the results of Experiment 2 should be similar to the nonthreat condition in Experiment 1.

Method

Participants. Forty younger adults (ages 18–23) and 40 older adults (ages 65–86) participated in Experiment 2. The younger adults were recruited from psychology classes at the University of Missouri and received course credit for participating. The older adults were recruited from the local community in central Missouri and received \$15 for participating. All participants reported being in good physical and mental health, with no major vision, hearing, or cognitive impairments. As in Experiment 1, older adults had significantly more years of formal education than did the younger adults, $t(78) = 8.60$, $p < .001$ (see Table 1 for additional demographic information). None of the participants had taken part in Experiment 1.

Design. Experiment 2 used a 2 (age: young vs. old) $\times$ 2 (test: item vs. associative) mixed factorial design. Age was a between subjects factor, while test was manipulated within subjects.

Stimuli. The materials were the same as those used in Experiment 1 (see Figure 1).

Procedure. The procedure was identical to Experiment 1, with four study-test blocks, two with unrelated word pairs, and two with face-scene pairs, except for one key aspect: participants were not told anything about age-related memory performance prior to beginning the study. In this way, there was no explicit manipulation of stereotype threat (either stereotype consistent- or stereotype-inconsistent information). Also, in keeping with the procedure used in past experiments testing the associative deficit, participants were not asked to give performance expectation ratings prior to beginning the first and third blocks. The one downside of this being that we therefore could not compare the mediating role of expectations in the control condition as we could for both conditions in Experiment 1.

Results

Memory accuracy was measured in proportion hits minus proportion false alarms. For the purpose of the main analysis, the two different types of stimuli were combined to form an overall item memory score and an overall associative memory score (see Table 2 for the separate hit and false alarm rates, and Figure 4 for the overall performance).

A two-way ANOVA was run, with type of test (item vs. associative) as a within-subject factor, and age (old vs. young) as a between-subject factor. A main effect of test, $F(1, 78) = 69.91$, $p < .001$, $\eta_p^2 = .47$, showed that item memory performance ($M = .65$, $SD = .14$) was better than associative memory performance ($M = .49$, $SD = .22$). Most importantly, the age-related associative deficit was replicated with a significant Age $\times$ Test interaction, $F(1, 78) = 5.41$, $p = .023$, $\eta_p^2 = .07$.

Hit and false alarm rates were also analyzed separately in the same manner as the hits minus false alarms. The hit rates showed no significant main effects or interactions. However, the false alarm rates produced results in line with the overall performance, including a significant main effect of test, $F(1, 78) = 90.05$, $p < .001$, $\eta_p^2 = .54$, and a significant Age $\times$ Test interaction, $F(1, 78) = 9.29$, $p < .01$, $\eta_p^2 = .11$, showing a higher rate of false alarms in the associative test for older adults ($M = .33$, $SD = .18$) than for younger adults ($M = .22$, $SD = .13$), $t(78) = 3.19$, $p < .01$, $d = .72$, yet no age differences in the item test ($M = .17$, $SD = .09$ for older adults, and $M = .14$, $SD = .08$ for younger adults), $t(78) = 1.59$, $p = .12$, $d = .36$.

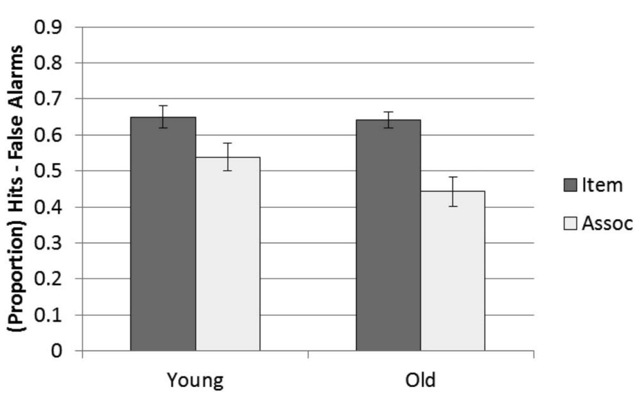

Figure 4. Proportion hits minus proportion false alarms as a function of age and test in Experiment 2 (error bars represent standard errors around the mean).

Discussion

The results from Experiment 2 replicated the age-related associative deficit (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008), as expected. As such, these results seem somewhat similar to the results from the threat condition in Experiment 1, supporting the notion that older adults are experiencing stereotype threat even under normal testing conditions in which nothing is explicitly mentioned regarding age-related memory differences. However, a follow-up analysis did not reveal an Age $\times$ Test $\times$ Condition interaction, though, $F(1, 136) = 1.54$, $p = .22$, when comparing the control condition of Experiment 2 with the nonthreat condition of Experiment 1. So while the control condition and the threat condition are similar and both show an age-related associative deficit (while the nonthreat condition does not), it cannot be said that the control condition induces the same level of stereotype threat as when stereotype-consistent information is explicitly mentioned (like the threat condition from Experiment 1). Also, since there were no expectation ratings prior to Experiment 2, it cannot be concluded whether participants' memory performance in the control condition was mediated by expectations, as in Experiment 1. On balance, the pattern of the results obtained seems to indicate that the control condition falls somewhat in the middle between the stereotype threat and the nonthreat conditions, indicating some potential stereotype threat effects involved in the associative deficit under control conditions.

General Discussion

Overall, through two experiments we have demonstrated that stereotype threat can affect the age-related associative deficit. In the first experiment, the associative deficit was eliminated whenever stereotype-inconsistent information was provided to participants (i.e., the nonthreat condition), and the associative deficit was present whenever stereotype-consistent information was provided to participants (i.e., the threat condition). Experiment 1 showed this effect even when using a within-subject design, and even though a fact-based manipulation was used (which recently was shown in a meta-analysis by Lamont, Swift, & Abrams, 2015, to produce smaller effects than stereotype-based manipulations). Experiment 2 was important in establishing a baseline level of memory performance when participants were not told anything ahead of time about age-related memory differences. Finding a somewhat similar age-related associative deficit in the control condition of Experiment 2, as was found in the threat condition of Experiment 1, suggests that older adults may be operating under some stereotype threat even under normal testing conditions, and that this is one factor driving the associative deficit typically seen in laboratory settings. However, it should be noted that the associative deficit was not as strong in the control condition as it was in the threat condition, evidenced by the lack of a triple interaction when comparing the control condition with the nonthreat condition. Taken together, these results suggest that providing stereotype-consistent information can enhance the effects of stereotype threat that are already seen in normal testing situations, and that in order to get rid of the effects of stereotype threat on older adults' memory performance, it may be necessary to provide information to individuals that directly counteracts these negative age-related stereotypes.

The fact that the overall effect was driven almost exclusively by an increase in false alarm rates as opposed to a decrease in hit rates when older adults were under stereotype threat (both in the threat condition and the control condition) is also in line with previous research on the associative deficit (Kilb & Naveh-Benjamin, 2011; Naveh-Benjamin et al., 2007) and provides further evidence that stereotype threat is affecting the associative deficit by disrupting strategic executive binding processes at encoding and recollection-based processes at retrieval.

Another important finding from the current study is that the effects of stereotype threat were shown using both a within- and a between-subject design. While other studies have shown effects of stereotype threat using between-subject designs, this is the first study to show an effect of stereotype threat on memory and aging using a within-subject design. This within-subject effect provides further evidence that stereotype threat is situational in nature (Steele, 1997). Stereotype threat is not a global effect that is always present and affecting behavior, but rather only manifests itself at the individual level under salient conditions. For instance, stereotype threat is not always active in older adults, but especially when one is actively engaged in the stereotyped domain (e.g., a memory task) and that stereotype has been made evident to the individual (e.g., explicitly mentioning expected age-related differences). In the current study, older adults' memory performance was affected in comparison with younger adults' when an age-based stereotype was explicitly mentioned before the study phase—even though the procedure and materials were otherwise similar.

Stereotype threat appears to operate by calling up negative expectations for individual performance based on group stereotypes, which then increases executive control interference and reduces elaborative strategies, resulting in lower task performance. In Experiment 1, participants' performance expectations were in line with their overall memory performance, suggesting that stereotype threat may affect memory performance by lowering performance expectations when participants are given stereotype-consistent information, and by raising performance expectations when participants are given stereotype-inconsistent information. Within the stereotype threat literature, most studies have not investigated the role of performance expectations. But as noted earlier, the present results are consistent with those that have. For instance, Desrichard and Kopetz (2005) found that presenting a task as memory based (as opposed to nonmemory based) lowered older adults' performance expectations, which, in turn, led to lower memory performance. Also, Hess, Hinson, et al. (2009) found that explicitly activating stereotype threat in older adults (similar to the procedure used in the current set of experiments) led to lower performance expectations and lower memory recall performance. Other stereotype threat domains have also showed a link between performance expectations and stereotype threat. Cadinu, Maass, Frigerio, Impagliazzo, and Latinotti (2003) found that women's expectations and performance on a difficult math test decreased when given negative gender stereotypes, as opposed to positive gender stereotypes. And Stone et al. (1999) found that participants under stereotype threat expected that they would need more strokes to complete a golf round than those not under stereotype threat. These results are in line with the expectancy-value theory (Eccles, 1983), which predicts that people's beliefs about how they will perform and the value they place on a task directly influences performance outcomes. What sets stereotype threat apart from the expectancy-value theory, however, is the fact that one's personal expectations can be derived from group-level stereotypes even when nothing is explicitly mentioned about individual expectations.

While stereotype threat reduced older adults' performance expectations for the associative memory task, performance expectations for the item memory task were not significantly affected. This is interesting because the threat manipulation did not mention anything about item versus associative memory, but only memory in general. One explanation as to why the item recognition task may have been less susceptible to stereotype threat is because participants perceived it as an easier task. Experiment 1 revealed a main effect of test when analyzing expectation ratings, showing that participants expected to perform significantly better on the item test than the associative test—regardless of the threat condition or participants' age. This would fit with the notion that stereotype threat has a larger effect on tasks that are perceived as being more difficult (Spencer et al., 1999). One limitation of the current study was that since the expectation rating questionnaire was not used in Experiment 2, nothing can be said about the mediating effects of expectations in the control condition.

Another potential limitation of the current study is that Experiment 2 cannot be considered a true control condition since it was not included as a separate condition in Experiment 1, due, among other things, to potentially lengthening the experiment to the point of becoming too fatiguing, especially for the older adults. Because of this, participants were not randomly assigned across all three conditions (stereotype threat, nonstereotype threat, and control) and performance expectations were not measured during Experiment 2. However, in light of these limitations, Experiment 2 still adds significant value in replicating the associative memory deficit (Naveh-Benjamin, 2000) while providing a baseline level of performance using the exact same stimuli as Experiment 1. Furthermore, the results of Experiment 2 suggest that the stereotype threat manipulation alone may not be the sole cause of the associative deficit, but that stereotype threat may be present even under normal testing conditions and that its effects can increase or decrease the associative memory deficit via the presentation of stereotype-consistent or -inconsistent information to participants, respectively.

One area of future research would be to clarify the underlying mechanisms responsible for stereotype threat, and in this case, its effect on older adults' associative memory performance. Some tentative conclusions can be drawn from the current study in favor of the executive control interference hypothesis (Schmader & Johns, 2003), most notably that stereotype threat had an effect on older adults' associative memory but not item memory by increasing false alarm rates (similar to Thomas & Dubois, 2011)—in line with the notion that stereotype threat impairs recollection while preserving, or perhaps even increasing, reliance on familiarity (Mazerolle et al., 2012). In contrast to this, though, others have recently argued that stereotype threat naturally causes a prevention focus where people attempt to limit errors, and this mindset runs counter to most promotion-focused memory-based tasks in which participants are told to try to remember as many words as possible (see the review by Barber, 2017). In support of this regulatory focus theory, negative memory effects of age-based stereotype threat can be eliminated and even reversed when memory task demands are manipulated in such a way as to make them prevention focused (Barber & Mather, 2013a, 2013b). While these results are intriguing, the current study cannot directly address the regulatory focus theory because we did not manipulate promotion and prevention task goals, and we used a recognition based memory test in which participants focus on *both* getting as many hits as possible (promotion focus) *and* avoiding as many false alarms as possible (prevention focus). Therefore, any effect of regulatory focus should be nullified in the current study. The fact that we were still able to see an effect of stereotype threat on memory, and that our threat condition did not cause a more conservative response bias, though, does not support the regulatory focus theory. However, the main purpose of this study was not to answer the question of underlying mechanisms, but instead to see if stereotype threat affects the age-related associative deficit—which it does. More research is needed to tease apart the true underlying mechanisms responsible for this effect.

In conclusion, the current study adds to the literature on the age-related associative memory deficit by showing that it can be reduced (and possibly amplified) by manipulating the presence of stereotype threat within participants. This research has significant practical applications, as it suggests that in addition to reducing general cognitive, sensory, and neural deficits, one way to overcome age-related memory decline is to work at reducing negative age-related stereotypes about memory that exist in society. Reducing ageism in general could even have significant health benefits for older adults (Nelson, 2016). And ultimately this may prove to be one of the most feasible approaches toward finding a solution to the problem of aging and memory.

References

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35,* 29–46. http://dx.doi.org/10.1006/jesp.1998.1371

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science, 7,* 136–141. http://dx.doi.org/10.1111/j.1467-9280.1996.tb00346.x

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,* 62–90.

Barber, S. J., & Mather, M. (2013a). Stereotype threat can both enhance and impair older adults' memory. *Psychological Science, 24,* 2522–2529. http://dx.doi.org/10.1177/0956797613497023

Barber, S. J., & Mather, M. (2013b). Stereotype threat can reduce older adults' memory errors. *The Quarterly Journal of Experimental Psychology, 66,* 1888–1895. http://dx.doi.org/10.1080/17470218.2013.840656

Bastin, C., & Van der Linden, M. (2003). The contribution of recollection and familiarity to recognition memory: A study of the effects of test format and aging. *Neuropsychology, 17,* 14–24. http://dx.doi.org/10.1037/0894-4105.17.1.14

Brubaker, M. S., & Naveh-Benjamin, M. (2014). The effects of presentation rate and retention interval on memory for items and associations in younger adults: A simulation of older adults' associative memory deficit. *Aging, Neuropsychology, and Cognition, 21,* 1–26. http://dx.doi.org/10.1080/13825585.2013.772558

Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology, 33,* 267–285. http://dx.doi.org/10.1002/ejsp.145

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition, 24,* 403–416. http://dx.doi.org/10.3758/BF03200930

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 191–211). New York, NY: Plenum Press. http://dx.doi.org/10.1007/978-1-4684-4178-9_11

Croizet, J. C., Després, G., Gauzins, M. E., Huguet, P., Leyens, J. P., & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin, 30,* 721–731. http://dx.doi.org/10.1177/0146167204263961

Desrichard, O., & Kopetz, 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. http://dx.doi.org/10.1002/ejsp.249

Eccles, J. S. (1983). Expectancies, values, and academic behavior. In J. T. Spencer (Ed.), *Achievement and achievement motivation* (pp. 75–146). San Francisco, CA: Freeman.

Eich, T. S., Murayama, K., Castel, A. D., & Knowlton, B. J. (2014). The dynamic effects of age-related stereotype threat on explicit and implicit memory performance in older adults. *Social Cognition, 32,* 559–570. http://dx.doi.org/10.1521/soco.2014.32.6.559

Foos, P. W. (1989). Adult age differences in working memory. *Psychology and Aging, 4,* 269–275. http://dx.doi.org/10.1037/0882-7974.4.3.269

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,* 778–784. http://dx.doi.org/10.1037/a0027754

Hay, J. F., & Jacoby, L. L. (1999). Separating habit and recollection in young and older adults: Effects of elaborative processing and distinctiveness. *Psychology and Aging, 14,* 122–134. http://dx.doi.org/10.1037/0882-7974.14.1.122

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,* P3–P11. http://dx.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: Psychological Sciences and Social Sciences, 64B,* 482–486. http://dx.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,* 621–625. http://dx.doi.org/10.1037/0882-7974.21.3.621

Hess, T. M., Hinson, J. T., & Hodges, E. A. (2009). Moderators of and mechanisms underlying stereotype threat effects on older adults' memory performance. *Experimental Aging Research, 35,* 153–177. http://dx.doi.org/10.1080/03610730802716413

Hess, T. M., Hinson, J. T., & Statham, J. A. (2004). Explicit and implicit stereotype activation effects on memory: Do age and awareness moderate the impact of priming? *Psychology and Aging, 19,* 495–505. http://dx.doi.org/10.1037/0882-7974.19.3.495

Hummert, M. L. (1999). A social cognitive perspective on age stereotypes. In T. M. Hess & F. Blanchard-Fields (Eds.), *Social cognition and aging* (pp. 175–196). San Diego, CA: Academic Press. http://dx.doi.org/10.1016/B978-012345260-3/50009-4

Hummert, M. L., Garstka, T. A., Shaner, J. L., & Strahm, S. (1994). Stereotypes of the elderly held by young, middle-aged, and elderly adults. *Journal of Gerontology, 49,* P240–P249. http://dx.doi.org/10.1093/geronj/49.5.P240

Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging, 8,* 283–293.

Jennings, J. M., & Jacoby, L. L. (1997). An opposition procedure for detecting age-related deficits in recollection: Telling effects of repetition. *Psychology and Aging, 12,* 352–361.

Joanisse, M., Gagnon, S., & Voloaca, M. (2013). The impact of stereotype threat on the simulated driving performance of older drivers. *Accident Analysis and Prevention, 50,* 530–538. http://dx.doi.org/10.1016/j.aap.2012.05.032

Jordano, M. L., & Touron, D. R. (2017). Stereotype threat as a trigger of mind-wandering in older adults. *Psychology and Aging, 32,* 307–313. http://dx.doi.org/10.1037/pag0000167

Kang, S. K., & Chasteen, A. L. (2009). The moderating role of age-group identification and perceived threat on stereotype threat among older adults. *The International Journal of Aging & Human Development, 69,* 201–220. http://dx.doi.org/10.2190/AG.69.3.c

Kilb, A., & Naveh-Benjamin, M. (2011). The effects of pure pair repetition on younger and older adults' associative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 706–719. http://dx.doi.org/10.1037/a0022525

Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin, 105,* 198–214. http://dx.doi.org/10.1037/0033-2909.105.2.198

Lambert, A. E., Watson, J. M., Stefanucci, J. K., Ward, N., Bakdash, J. Z., & Strayer, D. L. (2016). Stereotype threat impairs older adult driving. *Applied Cognitive Psychology, 30,* 22–28. http://dx.doi.org/10.1002/acp.3162

Lamont, R. A., Swift, H. J., & Abrams, D. (2015). A review and meta-analysis of age-based stereotype threat: Negative stereotypes, not facts, do the damage. *Psychology and Aging, 30,* 180–193. http://dx.doi.org/10.1037/a0038586

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9,* 339–355. http://dx.doi.org/10.1037/0882-7974.9.3.339

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,* 723–727. http://dx.doi.org/10.1177/0956797612437607

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods: Instruments & Computers, 36,* 630–633. http://dx.doi.org/10.3758/BF03206543

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1170–1187. http://dx.doi.org/10.1037/0278-7393.26.5.1170

Naveh-Benjamin, M., Brav, T. K., & Levy, O. (2007). The associative memory deficit of older adults: The role of strategy utilization. *Psychology and Aging, 22,* 202–208. http://dx.doi.org/10.1037/0882-7974.22.1.202

Naveh-Benjamin, M., & Kilb, A. (2014). Age-related differences in associative memory: The role of sensory decline. *Psychology and Aging, 29,* 672–683. http://dx.doi.org/10.1037/a0037138

Nelson, T. D. (2016). Promoting healthy aging by confronting ageism. *American Psychologist, 71,* 276–282. http://dx.doi.org/10.1037/a0040221

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23,* 104–118. http://dx.doi.org/10.1037/0882-7974.23.1.104

Rahhal, T. A., Colcombe, S. J., & Hasher, L. (2001). Instructional manipulations and age differences in memory: Now you see them, now you don't. *Psychology and Aging, 16,* 697–706. http://dx.doi.org/10.1037/0882-7974.16.4.697

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. Craik & T. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 1–90). Mahwah, NJ: Erlbaum.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., . . . Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15,* 1676–1689. http://dx.doi.org/10.1093/cercor/bhi044

Rugg, M. D., & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Current Opinion in Neurobiology, 23,* 255–260. http://dx.doi.org/10.1016/j.conb.2012.11.005

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428. http://dx.doi.org/10.1037/0033-295X.103.3.403

Sauvage, M. M., Beer, Z., & Eichenbaum, H. (2010). Recognition memory: Adding a response deadline eliminates recollection but spares familiarity. *Learning & Memory, 17,* 104–108. http://dx.doi.org/10.1101/lm.1647710

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

Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K. (2002). Listening in aging adults: From discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 56,* 139–152. http://dx.doi.org/10.1037/h0087392

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4–28. http://dx.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. http://dx.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,* 797–811. http://dx.doi.org/10.1037/0022-3514.69.5.797

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology, 77,* 1213–1227. http://dx.doi.org/10.1037/0022-3514.77.6.1213

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

Wingfield, A., Tun, P. A., & McCoy, S. L. (2005). Hearing loss in older adulthood: What it is and how it interacts with cognitive performance. *Current Directions in Psychological Science, 14,* 144–148. http://dx.doi.org/10.1111/j.0963-7214.2005.00356.x

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

Yeung, N. C. J., & von Hippel, C. (2008). Stereotype threat increases the likelihood that female drivers in a simulator run over jaywalkers. *Accident Analysis and Prevention, 40,* 667–674. http://dx.doi.org/10.1016/j.aap.2007.09.003

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In F. Craik & T. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 293–357). Mahwah, NJ: Erlbaum.

Appendix

Stereotype Threat Instructions Used in Experiment 1

Threat Condition

A primary goal of this experiment is to examine age differences in memory ability.

For the next two lists, I am going to examine your memory ability using verbal stimuli (unrelated word pairs) that have been used extensively by researchers to study aging effects on memory. Large age differences in memory performance are usually seen on this task when using word pairs. Specifically, older adults typically do poorer than younger adults.

Nonthreat Condition

A primary goal of this experiment is to examine individual differences in performance and the factors that account for those differences.

For the next two lists, I am going to examine your ability to process visual-spatial information. In an effort to reduce potential age biases, we will be using stimuli (faces and scenes) that have been shown to be appropriate for individuals of all ages. Specifically, older adults have been shown to do quite well on this task when using faces and scenes.

Note. Order of conditions was counterbalanced between subjects, as was stimuli type to threat condition.

Received May 11, 2017
Revision received July 31, 2017
Accepted August 4, 2017 ■