Journals of Gerontology: Psychological Sciences
cite as: J Gerontol B Psychol Sci Soc Sci, 2020, Vol. 75, No. 2, 241–250
doi:10.1093/geronb/gby031
Advance Access publication March 28, 2018

Original Research Report

The Unexpected Relationship Between Retrieval Demands and Memory Performance When Older Adults Are Faced With Age-Related Stereotypes

Ayanna K. Thomas, PhD,[1] Amy M. Smith, MS,[1] and Marie Mazerolle, PhD[2]

[1]Department of Psychology, Tufts University, Medford, Massachusetts. [2]Département de Psychologie, University of Bourgogne Franche Comté, France.

Address correspondence to: Ayanna K. Thomas, PhD, Psychology Department, Tufts University, 490 Boston Avenue, Medford, MA 02155. E-mail: ayanna.thomas@tufts.edu

Received: November 7, 2017; Editorial Decision Date: March 4, 2018

Decision Editor: Nicole Anderson, PhD, CPsych

Abstract

Objectives: In two studies, we examined the effects of age-related stereotype threat on eyewitness memory using the misinformation paradigm to (a) examine stereotype threat in the context of a more ecologically valid memory task and (b) to determine the relationship between task difficulty and susceptibility to stereotype threat.
Methods: After watching a video that depicted a crime, older and younger adult participants were presented with a written synopsis in which information consistent or inconsistent with the original event was presented. Half of the participants were then presented with information designed to activate negative stereotypes about aging. Finally, participants completed a memory test.
Results: In Study 1, when participants were instructed to report information from either the video or the synopsis to complete the final memory test, older adults under high stereotype threat were less accurate than those under low threat. In Study 2, when participants were required to engage in more controlled processes at retrieval and respond with only video information, older adults under stereotype threat performed as well or better than those under low threat.
Discussion: The results are consistent with the Regulatory Focus Model of Stereotype Threat.

Keywords: Eyewitness memory, Memory and aging, Stereotype threat in older adults

Research has consistently demonstrated that as we age, we perform less well on tests of episodic memory (Craik & Salthouse, 2011). Further, age-related deficits in memory are sometimes exacerbated when negative stereotypes about aging are activated (for a review see Barber, 2017). The present study examined the influence of negative stereotype activation in an ecologically valid eyewitness memory paradigm. We tested whether stereotype activation would impact memory when older adults were required to play the role of an eyewitness and remember information from a cohesive episodic narrative. We further examined whether the difficulty of the eyewitness memory task would influence the occurrence of stereotype threat effects.

We employed the robust misinformation paradigm to investigate older and younger adult eyewitness memory. In two studies, participants watched a video of a crime and then read a written synopsis that included information that was consistent or inconsistent with the video. Following, participants were either presented with an article highlighting negative aging stereotypes (high threat) or positive age-related changes (low threat). They then completed a cued recall test. Memory was assessed in situations where the same information was presented in the video and synopsis (consistent trials), information was presented only in the video (neutral trials), or information differed between the video and synopsis (misleading trials).

In a typical misinformation experiment, researchers are concerned with the retroactive influence of the written synopsis on memory for the original event. Therefore, when participants are given the final cued recall test, they are asked to only respond with the original video information, and ignore information presented in the synopsis. Importantly, participants are required to differentiate between information from two separate sources (video and synopsis) on certain trials. Across several studies older adults have demonstrated increased susceptibility to postevent misleading information as compared to younger adults (Bulevich & Thomas, 2012; Cohen & Faulkner, 1989; Coxon & Valentine, 1997; Karpel, Hoyer, & Toglia, 2001; Mitchell, Johnson, & Mather, 2003). Research suggests that age-related increased misinformation susceptibility may result, because older adults are less likely than younger adults to rely on effortful search processes and postretrieval monitoring necessary for effective source monitoring of retrieved details (see Thomas, Gordon, & Bulevich, 2014 for discussion). Negative stereotypes associated with aging may create added pressure on older adults that could further interfere with cognitive performance (see Barber, 2017; Lamont, Swift, & Abrams, 2015). In fact, recent studies have focused on the relationship between the activation of threat and retrieval processes and have generally reported detrimental effects of threat on memory (e.g., Krendl, Ambady, & Kensinger, 2015; Smith, Barber, Gallo, Maddox, & Thomas, 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016).

In the majority of stereotype threat studies with older adults, stereotype activation has been manipulated prior to encoding. In these studies, older adults under stereotype threat have demonstrated substandard working memory performance (Mazerolle, Régner, Morisset, Rigalleau, & Huguet, 2012), photo recall (Stein, Blanchard-Fields, & Hertzog, 2002), recall of prose passages (Kang & Chasteen, 2009), word recall (e.g., Hess, Auman, Colcombe, & Rahhal, 2003), word recognition (Hess, Emery, & Queen, 2009), and even prospective memory (Zuber, Ilhle, Blum, Desrichard, & Kliegel, 2017). In the present study, we were motivated to examine the impact of stereotype threat at retrieval because the activation of threat could interfere with the effortful search for previously stored memories, as well as the use of postretrieval monitoring and evaluation processes used to regulate memory accuracy in the misinformation paradigm (e.g., Johnson & Raye, 1981).

Guided by prior research examining stereotype activation at retrieval, we predicted that older adults would demonstrate stereotype threat related decrements in performance, and that decrement would be more apparent when the demands of the retrieval task were high. Our predictions regarding the influence of age-related stereotyping on memory were also informed by the *Executive Control Interference Integrated Process Model (EC)* of stereotype threat (Schmader, Johns, & Forbes, 2008). This model posits that stereotype threat consumes cognitive resources by increasing self-monitoring, self-regulation, and physiological arousal during the stereotyped task. Therefore, it follows that a challenging cognitive task, such as differentiating between information presented in the video and synopsis, would be vulnerable to stereotype threat-related impairment. In Study 1, after the high or low stereotype threat induction procedure, we gave participants cued recall instructions that encouraged responding with information that came from either the video or the synopsis. In Study 2, participants were required to respond with video information only. The unconventional instructions employed in Study 1 have been used in previous eyewitness memory research to investigate whether answers to questions on a final test would differ based on the type of instructions given (Roediger, Jacoby, & McDermott, 1996). Roediger et al. (1996) found that when participants could respond with either video or synopsis information, they performed better than when they were forced to respond with only video information.

Conceptually, the cued recall instructions given in Studies 1 and 2 are like the inclusion/exclusion paradigm used to investigate automatic and controlled processing in memory (for a review, see Yonelinas & Jacoby, 2012). Inclusion task instructions require participants to identify items that were previously studied, regardless of the way they were studied, thereby allowing for the influences of automatic and controlled processes on responding (Jacoby, 1991). Exclusion task instructions require participants to only identify the previously studied items that were studied in a specific way or from a specific source (e.g., list 1 vs list 2, Jacoby, 1991, Experiment 2). Research has consistently demonstrated that when older adults are required to engage in controlled memory tasks requiring exclusion of information, they perform less well than younger adults because age-related reductions in controlled processes allow automatic influences to have a greater influence on responding (e.g., Jacoby, 1999).

Successful performance on the cued recall task in Study 1 could result from the reliance on both automatic and controlled processes, because participants were encouraged to respond with information from the video or synopsis. Under these low-demand conditions, we predicted that stereotype threat activation would not affect older adult performance. Stereotype threat activation was also predicted to have no influence on younger adult performance. In Study 2, a reduction in the reliance on control processes was predicted to result in age-related susceptibility to the misinformation effect. Further, we expected that stereotype threat activation would increase misinformation susceptibility in older adults, due to the high cognitive demand of the final memory test. Importantly, this effect was expected only on misleading trials where two different pieces of information were presented across two sources.

Study 1

In Study 1, we used inclusion-based test instructions, to establish a testing scenario with reduced cognitive demands. Older and younger adults were expected to perform similarly. Further, based on EC, we predicted that stereotype threat activation would not affect older adult performance, because the demands of the final memory test were low. Because stereotype threat was predicted to have little or no impact on the eyewitness memory test of primary interest, we also included the Operation Span task (OSPAN; Unsworth, Heitz, Schrock, & Engle, 2005). Prior research has demonstrated that stereotype threat activation does influence working memory abilities (Abrams, Eller, & Bryant, 2006; Desrichard & Kopetz, 2005; Mazerolle et al., 2012; but see Hess, Hinson, & Hodges, 2009). Therefore, the present study used the OSPAN task to determine whether stereotype threat activation did in fact influence older adults' working memory performance as predicted. OSPAN was completed prior to and following stereotype threat induction. In both cases, participants were told they would be completing simple math problems and recalling letters presented between problems.

Method

Design

We employed a 2 (Age: Younger, Older) × 2 (Threat: High, Low) × 3 (Item Type: Consistent, Neutral, Misleading) mixed factorial design. Age and Threat were manipulated between-subjects, and Item Type was manipulated within-subjects.

Participants

Assuming an effect size of Cohen's d = .79 derived from the average of previously published studies (Abrams et al., 2006; Barber & Mather, 2013; Desrichard & Kopetz, 2005; Grimm, Markman, Maddox, & Baldwin, 2009; Hess, Auman, Colcombe, & Rahhal, 2003; Hess, Emery, et al., 2009; Hess, Hinson, et al., 2009; Kang & Chasteen, 2009; Mazerolle et al., 2012; Popham & Hess, 2015; Seibt & Förster, 2004; Smith et al., 2017; Thomas & Dubois, 2011), a significance level of $\alpha = .05$, and four between-subjects groups, we determined that a total sample size of participants (n = 30 per group) would provide 95% power to detect effects. The sample size used in these studies was based on an a priori power analysis conducted in G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007).

One hundred twenty-three adults participated in the study. Younger adults (n = 61) were undergraduate students recruited from Tufts University who received course credit (M_{age} = 19.41, Female = 31). Older adults (n = 62) were selected from a participation pool maintained by the Cognitive Aging and Memory Laboratory at Tufts University and were paid $15.00 for their participation (M_{age} = 73.51, Female = 38). Older adults were more highly educated than younger adults, $t(121) = 7.86$, $p < .001$. To be a member of the participant pool, older adults were prescreened for psychological and neurological health issues and could not be taking medications that might affect cognition (e.g., antidepressants, drugs with anticholinergic properties, benzodiazepines, opiates, and/or anticonvulsants).

One younger adult was not included in subsequent analyses for failing to comply with task instructions. Twenty-nine younger and 31 older adults were randomly assigned to the high threat group and 32 younger and 31 older adults were randomly assigned to the low threat group. Younger adults were tested in groups ranging from one to four and older adults were tested individually. Independent samples t tests revealed that group versus individual testing did not affect younger adult performance on any measures of memory (all t's < 1). Additional demographic and sample information can be found in Table 1.

Materials and Procedure

After providing informed consent, participants completed an automated version of the OSPAN test (Unsworth et al., 2005). In this task, participants were asked to read and verify a simple math problem and then read a letter after the operation. After participants were presented with a series of alternating math verification problems and letters, they were prompted to recall in order the letters that followed each math operation. The number of math/letter strings in a sequence was increased and decreased to measure each participant's operation span. Participants completed a brief practice phase to become familiar with the task before taking the OSPAN test. After the OSPAN test, participants viewed a 22-min video of the film *Rififi* (Bezard, Bérard, Cabaud, & Dassin, 1955) depicting four men committing a burglary. Participants were simply instructed to watch the video and were not told that their memory for the video would

Table 1. Demographic Information

		Age	Vocabulary	Years of education
Study 1				
Older adults				
	High Threat	74.7 (6.8)	14.97 (2.8)	15.87 (2.1)
	Low Threat	72.3 (7.2)	14.25 (2.5)	14.61 (3.0)
Younger Adults				
	High Threat	19.5 (1.2)	13.67 (1.3)	12.4 (1.0)
	Low Threat	19.3 (1.7)	13.55 (1.8)	12.3 (1.2)
Study 2				
Older adults				
	High Threat	71.8 (5.8)	13.4 (2.4)	16.60 (2.3)
	Low Threat	72.6 (8.1)	13.5 (2.5)	15.94 (2.5)
Younger Adults				
	High Threat	20.9 (2.3)	15.3 (2.4)	13.0 (2.3)
	Low Threat	20.5 (2.5)	15.3 (2.3)	13.3 (1.8)

Note: Means and *SD* are displayed for age, vocabulary scores, and years of education.

be tested. Afterward, participants completed the Vocabulary Subtest of the Shipley Institute of Living Scale 2 (see Table 1; Shipley, 1946) on which older adults (M = 15.39) outperformed younger adults (M = 13.40), t (121) = 4.60, $p < .01$.

Participants then listened to a 6-min synopsis of the video. They were simply instructed to listen to the synopsis. The synopsis consisted of 115 sentences, eight of which contained details that were consistent with the video (e.g., The man was wearing *a watch* around his wrist), eight of which contained nonspecific "neutral" details (e.g., The man was wearing *something* around his wrist), and eight of which introduced details that were inconsistent with the video (e.g., The man was wearing *a bracelet* around his wrist). All other sentences were used for filler and contained information that was never assessed on the subsequent cued recall test. The 24 sentences that presented consistent, neutral, and inconsistent details were counterbalanced across participants.

Immediately after the synopsis, participants read one of the two stereotype passages (previously used in Thomas & Dubois, 2011). The *high threat* manipulation was a fabricated scientific article that presented evidence that memory declines with age (307 words). The *low threat* article described research showing that some types of memory do not decline with age (344 words). Participants were then given standard instructions for the cued recall test, which made no mention of examining age differences in memory (borrowed from Gordon & Thomas, 2014). They were told to answer questions based on their memory of the video and/or synopsis. Participants were also instructed that they could choose not to provide answers to questions, further reducing the demands of the test. Eight questions were associated with consistent trials in which the same information was presented in the video and the synopsis. Eight questions were associated with neutral trials in which non-specific information had been presented in the synopsis, and thus critical details could only be retrieved from memory for the video. Eight questions were associated with misleading trials in which different information was presented between the video and the synopsis. All cued recall questions required a one- or two-word response. Questions were always presented in the same order, and no feedback regarding correctness was provided. Participants were given up to 15 s to answer each question and were not forced to respond. An analysis of response onset time suggests that older adults began responding to a question on average at 2.5 s. Response onset times for older adults ranged from 500 ms to 4.52 s. Therefore, we are confident that the 15 s window did not put any undue pressure on older adults. The cued recall test took 6 min to complete.

Finally, participants completed the second OSPAN test, which was again preceded by a practice phase. The OSPAN tests, video, synopsis, and cued recall test were presented using E-Prime software (Version 2.1; Schneider, Eschman, & Zuccolotto, 2002). All procedures occurred in one testing session.

Results

Accurate Recall

A cued recall response was considered correct if it was consistent with what was presented in the video, or if it was presented as inconsistent information in the synopsis. A 2 (Age: Younger, Older) × 2 (Threat: High, Low) × 3 (Item Type: Consistent, Neutral, Misleading) analysis of variance (ANOVA) on mean proportion of accurate recall found main effects of Age, $F(1, 119) = 10.47$, $p < .005$, $\eta_p^2 = .08$, Threat, $F(1, 119) = 8.74$, $p < .005$, $\eta_p^2 = .07$, and Item Type, $F(2, 238) = 66.97$, $p < .001$, $\eta_p^2 = .36$. These main effects were considered within the context of the significant interaction among the three variables, $F(2, 238) = 6.10$, $p < .005$, $\eta_p^2 = .05$.

We decomposed this three-way interaction by conducting separate 2 (Threat: High, Low) × 3 (Item Type: Consistent, Neutral, Misleading) ANOVAs for each age group. We chose to examine the three-way interaction in analyses that separately compared older and younger adults, because the stereotype threat manipulation was targeted to older and not younger adults. Further, the age-related increase in misinformation susceptibility was captured by the main effect of age in the omnibus test. For older adults, we found a main effect of Threat, $F(1, 60) = 8.54$, $p < .005$, $\eta_p^2 = .13$. As Table 2 demonstrates, older adults under high threat (M = 0.54) produced on average fewer correct responses than older adults under low threat (M = 0.67). We also found a main effect of Item Type, $F(2, 120) = 31.39$, $p < .001$, $\eta_p^2 = .34$. Older adults produced on average more correct responses on consistent (M = 0.68) and misleading (M = 0.68) trials than on neutral trials (M = 0.45). However, these main effects should be considered in the context of the significant interaction between Threat and Item Type, $F(2, 120) = 9.45$, $p < .001$, $\eta_p^2 = .14$. As Table 2 demonstrates, older adults under high threat produced fewer correct answers on consistent (M = 0.59) and misleading (M = 0.56) trials than older adults under low threat (M = 0.78 and M = 0.80, respectively). Planned comparisons using a Bonferroni correction confirmed that the differences on consistent ($t(60) = 3.32$, $p < .005$, $d = .88$) and misleading trials ($t(60) = 3.62$, $p < .005$, $d = .90$) were significant. Threat condition did not influence performance on neutral trials. For younger adults, we only found a main effect of Item Type, $F(2, 118) = 36.42$, $p < .001$, $\eta_p^2 = .38$, as they demonstrated better accuracy on consistent (M = 0.78) and misleading (M = 0.77) as compared to neutral (M = 0.57) trials.

Errors of Omission and Commission

Because participants were given the opportunity to leave questions blank on the cued recall test, incorrect responses could be categorized as either errors of omission or commission. We next examined whether stereotype threat influenced the occurrence of these two types of errors. Because threat did not impact younger adults' accuracy, we did not examine omission and commission error patterns for the younger adult group.

Table 2. Average Accuracy, Errors of Omission, and Errors of Commission on the Cued Recall Test in Study 1

		Consistent	Neutral	Misleading
Older Adults				
High Threat				
	Accuracy	0.59 (0.23)	0.47 (0.20)	0.57 (0.32)
	Omission	0.22 (0.22)	0.27 (0.22)	0.26 (0.27)
	Commission	0.19 (0.16)	0.26 (0.03)	0.17 (0.11)
Low Threat				
	Accuracy	0.78 (0.23)	0.44 (0.22)	0.80 (0.17)
	Omission	0.06 (0.12)	0.11 (0.13)	0.07 (0.12)
	Commission	0.17 (0.16)	0.45 (0.23)	0.13 (0.12)
Younger Adults				
High Threat				
	Accuracy	0.76 (0.24)	0.53 (0.25)	0.75 (0.17)
	Omission	0.05 (0.11)	0.06 (0.11)	0.07 (0.10)
	Commission	0.19 (0.18)	0.41 (0.23)	0.18 (0.15)
Low Threat				
	Accuracy	0.80 (0.17)	0.59 (0.19)	0.79 (0.19)
	Omission	0.03 (0.02)	0.07 (0.10)	0.04 (0.08)
	Commission	0.17 (0.16)	0.34 (0.20)	0.17 (0.16)

Note: Means and *SD* are presented.

Errors of omission occurred when participants left answers blank. Mean proportions of omission errors were subjected to a 2 (Threat: High, Low) × 3 (Item Type: Consistent, Neutral, Misleading) ANOVA. We found a main effect of Threat, $F(1, 60) = 17.35$, $p < .001$, $\eta_p^2 = .22$. As shown in Table 2, older adults under high threat were significantly more likely to leave answers blank ($M = 0.26$) as compared to older adults under low threat ($M = 0.08$). No other effects were significant.

Commission errors were defined as cued recall responses that did not meet the criteria for accurate recall. Mean proportions of commission errors were examined using a 2 (Threat: High, Low) × 3 (Item Type: Consistent, Neutral, Misleading) ANOVA. We found a marginal main effect of Threat, $F(1, 60) = 3.22$, $p = .078$, and a main effect of Item Type, $F(2, 120) = 29.23$, $p < .001$, $\eta_p^2 = .33$. We examined these effects in the context of the significant two-way interaction, $F(2, 120) = 8.07$, $p < .001$, $\eta_p^2 = .12$. As demonstrated in Table 2, older adults who received high threat induction were significantly less likely to make commission errors on neutral trials ($M = 0.27$) than older adults who received low threat induction ($M = 0.46$). Planned comparisons using a Bonferroni correction confirmed that this difference was significant, $t(60) = 3.38$, $p < .001$, $d = .88$. There were no other significant comparisons.

OSPAN

We examined whether stereotype threat induction had generalized effects on working memory. In previous research, relative to older adults in a low threat group, older adults under high threat demonstrated poorer letter recall on the OSPAN task but no differences in performance on the math verification problems (Jordano & Touron, 2017). Therefore, we separately examined the effects of age and stereotype threat on (a) the partial span score, which is the total number of letters accurately recalled in the correct order across all trials, and (b) the total number of errors made on the math verification problems. Means and standard deviations for these measures are presented in Table 3.

As Table 3 demonstrates, whereas age group influenced pre- and post-threat OSPAN performance, threat had no impact on the OSPAN task administered prior to induction. Therefore, we conducted a 2 (Age: Younger, Older) × 2 (Threat: High, Low) ANCOVA, with average post-threat partial span scores as the dependent measure and average pre-threat partial span scores as a covariate. We found a main effect of Age, $F(1, 109) = 9.85$, $p < .002$, $\eta_p^2 = .08$. On average, older adults demonstrated lower partial span scores ($M = 30.25$) than younger adults ($M = 62.75$). We also found a marginal interaction between Age and Threat, $F(1, 109) = 3.50$, $p = .06$. Older adults in the high threat group ($M = 28.8$) demonstrated lower partial span scores than older adults in the low threat group ($M = 36.6$). Younger adults in the high ($M = 63.9$) and low ($M = 65.3$) threat groups did not demonstrate differences.

We next examined the mean number of errors made during math verification on the post-threat OSPAN task in a 2 (Age: Younger, Older) × 2 (Threat: High, Low) ANCOVA where average pre-threat math errors served as the covariate. We found a main effect of Age, $F(1, 110) = 4.66$, $p = .033$, $\eta_p^2 = .04$, as younger adults made fewer math verification errors ($M = 4.8$) than older adults ($M = 9.5$). No other effects were significant.

Discussion of Study 1

Older adults who were exposed to negative age-related stereotypes recalled fewer items than those who were exposed to a positive passage. Further, poorer recall performance for these older adults was driven by an increase in errors of omission. These results are counter the predictions made by the EC. In the context of a less cognitively effortful task, EC would predict little to no effect of stereotype threat on memory performance. That said, consistent with EC, we did find a marginal effect of stereotype threat on working memory performance.

Although inconsistent with EC, these eyewitness memory results are broadly consistent with the *Regulatory Focus Model of Stereotype Threat* (RF) (Barber, 2017). The RF model predicts stereotype-related cognitive impairment when there is a conflict between the short term regulatory focus instantiated by the activation of negative stereotypes (e.g., Seibt & Förster, 2004), and the regulatory focus engendered by traditional cognitive tests (e.g., Grimm et al., 2009), regardless of the cognitive demands of

Table 3. Average OSPAN Partial Span Scores (i.e., total number of letters recalled in the correct order) and Average Math Verification Errors Made During the Prethreat and post-threat OSPAN Tasks in Study 1

	Prethreat span score	Post-threat span score	Prethreat math errors	Post-threat math errors
Older Adults				
High Threat	26.84 (17.76)	28.80 (16.53)	7.94 (4.27)	9.70 (5.61)
Low Threat	28.96 (18.74)	36.63 (19.96)	7.41 (4.71)	9.33 (11.93)
Younger Adults				
High Threat	60.39 (11.09)	63.93 (8.82)	4.14 (2.61)	4.76 (2.90)
Low Threat	61.90 (11.89)	65.26 (9.95)	4.72 (2.90)	4.81 (2.53)

Note: Means and *SD* are presented.

the test. The implicit gains-based frame of the memory test in Study 1 may have led to regulatory conflict for prevention-focused, threat-activated older adults. This conflict in regulatory focus may have resulted in older participants under high threat choosing to withhold responses, rather than risking making errors of commission.

Study 2

In Study 2, we examined whether older adults under high and low stereotype threat would continue to demonstrate differences in eyewitness memory when the test required a more controlled, effortful memory search. We increased the demands of the final cued recall test by instructing participants to base their answers solely on their memory for the video, by forcing participants to provide a response to each question, and by introducing a source monitoring question that accompanied each item on the cued recall test. Finally, in Study 2, we removed the OSPAN task because our primary interest was performance on the eyewitness memory tasks.

Method

Design

The study employed the same design as Study 1.

Participants

Based on the a priori power analysis discussed in Study 1, we recruited 132 adults to participate in the study. Participant recruitment and compensation were identical to Study 1. Sixty-six younger adults (M_{age} = 20.73, Female = 39) and 66 older adults (M_{age} = 72.59, Female = 51) participated. As in Experiment 1, younger adults were tested in groups ranging from 1 to 4. Performance for younger participants tested individually as compared to within a group did not differ. Older participants were more highly educated than younger adults, $t(131) = 7.91$, $p < .001$. As in Study 1, older adults also performed better than younger adults on the Shipley vocabulary test, $t(131) = 2.70$, $p < .01$. One younger adult was not included in subsequent analyses because of failure to comply with instructions.

Materials and Procedure

The materials and procedures used in Study 2 were identical to those used in Study 1 with the following exceptions. Participants did not complete the OSPAN tasks before and after threat induction. In addition, the cued recall test instructed participants to report only information that they remembered from the video, and participants were required to provide an answer to every question. After the cued recall test, participants completed a source monitoring test. They were represented with each cued recall question and the response they provided, one question at a time. Upon viewing each question and their answer, they were asked to indicate whether their response was associated with information presented in the video, the synopsis, both, or neither. Both the cued recall test and subsequent source monitoring test were self-paced.

Results

Accurate Recall

A 2 (Age: Younger, Older) × 2 (Threat: High, Low) × 3 (Item Type: Consistent, Neutral, Misleading) ANOVA on mean proportion of accurate recall found main effects of Age, $F(1, 127) = 13.40$, $p < .001$, $\eta_p^2 = .10$, and Item Type, $F(2, 254) = 68.12$, $p < .001$, $\eta_p^2 = .35$. As shown in Table 4, on average, older adults (M = 0.49) were less accurate than younger adults (M = 0.59). In addition, average performance was better on consistent (M = 0.68) as compared to neutral (M = 0.51) trials, and performance on neutral trials was better than on misleading trials (M = 0.43). The latter comparison demonstrates the robust misinformation effect. Both comparisons were statistically significant after correcting for alpha inflation using the Bonferroni method (consistent-neutral: $t(130) = 7.89$, $p < .001$, $d = .74$; neutral-misleading: $t(130) = 4.00$, $p < .001$, $d = .36$).

Misleading Errors of Commission

A 2 (Age: Younger, Older) × 2 (Threat: High, Low) × 3 (Item Type: Consistent, Neutral, Misleading) ANOVA on mean proportions of misinformation commission errors

Table 4. Average Accuracy and Misleading Errors of Commission on the Forced Cued Recall Test in Study 2

			Consistent	Neutral	Misleading
Older Adults					
	High Threat				
		Accuracy	0.61 (0.24)	0.52 (0.20)	0.34 (0.14)
		Commission Errors	0.09 (0.10)	0.09 (0.09)	0.34 (0.21)
	Low Threat				
		Accuracy	0.65 (0.23)	0.44 (0.22)	0.38 (0.25)
		Commission Errors	0.09 (0.09)	0.13 (0.08)	0.37 (0.23)
Younger Adults					
	High Threat				
		Accuracy	0.73 (0.22)	0.55 (0.22)	0.50 (0.21)
		Commission Errors	0.08 (0.11)	0.09 (0.07)	0.28 (0.19)
	Low Threat				
		Accuracy	0.74 (0.22)	0.55 (0.25)	0.49 (0.23)
		Commission Errors	0.06 (0.08)	0.12 (0.11)	0.32 (0.22)

Note: Means and *SD* are presented.

revealed the classic misinformation effect as participants were more likely to produce misleading details on misleading trials (M = 0.32) than on neutral (M = 0.11) or consistent (M = 0.08) trials, $F(2, 254) = 125.11$, $p < .001$, $\eta_p^2 = .50$ (see Table 4). We also found a marginal main effect of Age, $F(1, 127) = 3.26$, $p = .07$, as older adults produced numerically more misleading details than younger adults. No other effects were significant.

Source Attributions for Correctly Recalled Items

We next examined source attributions made for items that were correctly recalled on the cued recall test. Across all groups and all item types, average correct cued recall was .65. A source attribution was scored as correct for a consistent item if it was attributed to the video *and* synopsis. A source attribution for a correctly recalled neutral or misleading item was scored as correct if that response was attributed to the video. Analysis of correct source attributions was conducted on these correctly recalled items using a 2 (Age: Younger, Older) × 2 (Threat, High, Low) × 3 (Item Type: Consistent, Neutral, Misleading) ANOVA on average proportions of correct source attributions. We found a three way interaction, $F(2, 254) = 3.02$, $p < .05$, $\eta_p^2 = .03$.

This interaction was deconstructed by separate 2 (Threat: High, Low) × 3 (Item Type: Consistent, Neutral, Misleading) ANOVAs for each age group. For younger adults, we found a main effect of Item Type, $F(2, 126) = 30.63$, $p < .001$, $\eta_p^2 = .32$. As Table 5 demonstrates, younger adults were less accurate at attributing source on consistent (M = 0.43) as compared to neutral (M = 0.74) or misleading (M = 0.74) trials, (consistent – neutral: $t(64) = 6.45$, $p < .001$, $d = 1.22$; consistent – misleading: $t(64) = 6.16$, $p < .001$, $d = 1.13$; neutral – misleading: $t < 1$). No other effects were significant, *F's* < 1. For older adults, we found a main effect of item type, $F(2, 128) = 36.04$, $p < .001$, $\eta_p^2 = .36$. As with younger adults, older adults demonstrated poorer source accuracy on consistent (M = 0.34) as compared to neutral (M = 0.72) and misleading (M = 0.76) trials, (consistent – neutral: $t(65) = 6.69$, $p < .001$, $d = 1.24$; consistent – misleading: $t(65) = 6.61$, $p < .001$, $d = 1.30$; neutral – misleading: $t < 1$). In addition, for older adults, we found a main effect of threat, $F(1, 64) = 4.24$, $p < .05$, $\eta_p^2 = .06$. Older adults in the low threat group (M = 0.56) were less accurate at attributing the source to correctly recalled items than older adults in the high threat group (M = 0.65). The interaction between Threat and Item Type was not significant, $F = 1.02$.

Discussion of Study 2

On the cued recall test in Study 2, we instructed participants to only respond with information from one source, thereby establishing exclusion-based task demands. Such task demands have been shown to be more challenging than those used in Study 1 for both older and younger adults. Further, participants were required to respond to every question, thereby reducing their ability to behave in a risk-averse manner. Finally, participants engaged in a source monitoring test for information produced on the cued recall test. Each of these task constraints were put in place to test predictions made by the EC account. This model would predict stereotype threat-related impairment on both the cued-recall and source monitoring tests. However, stereotype activation did not result in memory impairment in older adults.

It is possible that we did not detect threat effects in this study because the study was underpowered. A recent meta-analysis of stereotype threat suggested that threat effects are only small to medium in magnitude (see Lamont et al., 2015). That said, the power calculation used in Study 1,

Table 5. Source Attribution Accuracy for Correctly Recalled Items in Study 2

	Consistent	Neutral	Misleading
Older Adult			
High Threat	0.43 (0.35)	0.75 (0.25)	0.78 (0.35)
Low Threat	0.24 (0.25)	0.69 (0.34)	0.75 (0.33)
Younger Adult			
High Threat	0.39 (0.28)	0.77 (0.19)	0.78 (0.25)
Low Threat	0.47 (0.29)	0.70 (0.29)	0.70 (0.29)

Note: Means and *SD* are presented.

in which stereotype threat effects were found, were again used in Study 2. Additionally, in Study 2, we found that older adults under high threat demonstrated better source accuracy for correctly remembered details than older adults under low threat. Although these results may seem counterintuitive, they are also broadly consistent with the RF model of stereotype threat. Although the source monitoring task in the present study did not directly manipulate incentives as is common practice in direct tests of the RF model, the instructions to review one's answers from the cued recall test and attribute those answers to the appropriate source may have engendered a more cautious and analytic approach to the memory test (see Thomas & Bulevich, 2006). A regulatory fit may thus have been achieved when older adults' regulatory focus under high threat (oriented to prevention) matched the task orientation (also oriented to prevention). This fit may have resulted in the more deliberate consideration of memorial cues that foster effective source monitoring, thus resulting in our observed benefit of stereotype threat on source attributions.

General Discussion

Across two studies, we investigated the impact of stereotype threat on older adults' eyewitness memory within the misinformation paradigm. Although research has consistently demonstrated that older adults perform more poorly across a variety of memory tests when under the pressure of stereotype threat (for reviews see Barber, 2017; Lamont et al., 2015), only one other study has investigated the impact of threat on memory for a complex event (Henkel, 2014). Further, the present study is one of only a handful to investigate the impact of stereotype threat on memory retrieval. Thus, the present investigation is the first to provide evidence of how negative age-related stereotyping influences retrieval of complex information. Further, through our use of different test instructions in Studies 1 and 2, we gained insight on how stereotype threat affects older adults' approach to a test of eyewitness memory. The combined results of the two studies suggest that when negative stereotypes are activated prior to retrieval, older adults may approach a memory test in a more cautious manner than when threat is not activated. These results are consistent with a growing body of research that suggests that older adult cognitive function should be understood within the context of contributing social interaction factors (e.g., Kensinger & Gutchess, 2017).

In Study 1, in which participants were presented with a memory test that allowed them to respond with information from either of two studied sources and could leave answers blank, older participants under high threat were less accurate on the final memory test than those under low threat. However, this difference was due to greater withholding of responses for older adults in the high threat group. These data suggest that negative stereotype activation may result in older adults being more cautious in how they approach a final test of memory, which is consistent with research showing that stereotype threat engenders a prevention focus (Seibt & Förster, 2004).

By changing the task instructions of the cued recall test in Study 2, eliminating the option to withhold responses, and adding a source monitoring test, we were better able to examine whether the EC model could account for stereotype threat effects in the misinformation paradigm. Under these constraints, stereotype threat had null or positive effects on memory. Although standard misinformation effects were demonstrated in both older and younger adults, when comparing accuracy for neutral and misleading items, stereotype threat did not further exacerbate this effect. Further, when asked to make source attributions for correctly retrieved items, older adults under high threat demonstrated better accuracy than those under low threat.

We suggest that the pattern of results across the two studies indicates that the activation of negative age-related stereotypes in older adults did not consume the cognitive resources necessary to complete an eyewitness memory retrieval task. Under relatively low demands in Study 1, stereotype threat still resulted in substandard performance. However, the performance difference between the older adult high and low threat groups seemed to be driven by cautious withholding of answers as opposed to memory inaccessibility. The results of Study 2 corroborated the evidence against a resource depletion explanation, since the memory impairment observed in Study 1 was eliminated in the context of an even more demanding memory test.

Though we find that the RF model most suitably accounts for the observed interactions between stereotype threat and older adult memory performance, the role of executive resource depletion cannot be completely dismissed. Performance on the OSPAN task after stereotype threat induction suggests that stereotype threat may have negatively impacted available cognitive resources, consistent with previous research (Abrams et al., 2006; Desrichard & Kopetz, 2005; Mazerolle et al., 2012). These results suggest that executive resource depletion may remain a consequence of stereotype threat activation in older adults. Further, in Study 1, the OSPAN tasks themselves may have impacted resource consumption resulting in the observed pattern of results. While possible, we argue that this explanation is unlikely, as the inclusion of OSPAN would have equally impacted both groups of older adults.

As a final note, the pattern of threat-related caution that we observed in the present set of studies may be advantageous in the context of older adult witness memory. As opposed to focusing on completeness, older adults under stereotype threat may be more concerned with accuracy. Consider the older witness to a crime. It is likely that the witness will be motivated to give an accurate report of the witnessed event. Our results suggest that the activation of age-related stereotypes in this scenario would not affect memory accuracy. On the contrary, the stereotyped witness may actually be more inclined to withhold information of which the witness is less certain, providing a more accurate testimony. Thus, withholding may result in the omission of correct details, but such a price may be worth paying to encourage scrutiny of eyewitness self-reports.

Funding

None declared.

Ethics Approval and Consent to Participate

This research was approved by the Tufts University Social, Behavioral, and Educational Research Institutional Review Board. All participants provided written consent to participate in this study.

Consent for Publication

Not applicable.

Availability of Data and Material

Data and materials are available upon request. All requests should be made directly to the first/corresponding author.

Author Contributions

All authors contributed to the conceptual development of the hypotheses and experiments. Amy M. Smith was primarily responsible for data collection. Data analysis was conducted by the first author, Ayanna K. Thomas. Primary drafting of the manuscript was done by Amy M. Smith and Ayanna K. Thomas. The first and second authors contributed equally to the development of this manuscript.

Conflict of Interest

None reported.

References

Abrams, D., Eller, A., & Bryant, J. (2006). An age apart: The effects of intergenerational contact and stereotype threat on performance and intergroup bias. *Psychology and Aging*, **21**, 691–702. doi:10.1037/0882-7974.21.4.691

Barber, S. J. (2017). An examination of age-based stereotype threat about cognitive decline. *Perspectives on Psychological Science*, **12**, 62–90. doi:10.1177/1745691616656345

Barber, S. J., & Mather, M. (2013). Stereotype threat can reduce older adults' memory errors. *The Quarterly Journal of Experimental Psychology*, **10**, 1888–1895. doi:10.1080/17470218.2013.840656

Bezard, R., Bérard, H., Cabaud, P., & Dassin, J. (1955). *Rififi* [Motion Picture]. France: Pathé Consortium Cinéma.

Bulevich J. B., & Thomas A. K. (2012) Retrieval effort improves memory and metamemory in the face of misinformation *Journal of Memory and Language*, **67**, 45–58. doi:10.1016/j.jml.2011.12.012

Cohen, G., & Faulkner, D. (1989). Age differences in source forgetting: Effects on reality monitoring and on eyewitness testimony. *Psychology and Aging*, **4**, 10–17. doi:10.1037/0882-7974.4.1.10

Coxon, P., & Valentine, T. (1997). The effects of the age of eyewitnesses on the accuracy and suggestibility of their testimony. *Applied Cognitive Psychology*, **11**, 415–430. doi:10.1002/(SICI)1099-0720(199710)11:5%3C415::AID-ACP469%3E3.0.CO;2-A

Craik, F. I., & Salthouse, T. A. (2011). *The handbook of aging and cognition*. New York, NY: Psychology 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. doi:10.1002/ejsp.249

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, **39**, 175–191. doi:10.3758/BF03193146

Gordon, L. T., & Thomas, A. K. (2014) Testing potentiates new learning in the misinformation paradigm. *Memory & Cognition*, **42**, 186–97. doi:10.3758/s13421-013-0361-2

Grimm, L. R., Markman, A. B., Maddox, W. T., & Baldwin, G. C. (2009). Stereotype threat reinterpreted as a regulatory mismatch. *Journal of Personality and Social Psychology*, **96**, 288–304. doi:10.1037/a0013463

Henkel, L. A. (2014). Do older adults change their eyewitness reports when re-questioned? *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, **69**, 356–365. doi:10.1093/geronb/gbt071

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**, 3–11. doi: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. doi: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. doi: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. doi:10.1080/03610730802716413

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, **30**, 513–541. doi:10.1016/0749-596X(91)90025-F

Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25**, 3–22. doi:10.1037/0278-7393.25.1.3

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, **88**, 67–85. doi:10.1037/0033-295X.88.1.67

Jordano, M. L., & Touron, D. R. (2017). Stereotype threat as a trigger of mind-wandering in older adults. *Psychology and Aging*, **32**, 307–313. doi: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. doi:10.2190/AG.69.3.c

Karpel, M. E., Hoyer, W. J., & Toglia, M. P. (2001). Accuracy and qualities of real and suggested memories: Nonspecific age differences. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, **56**, 103–110. doi:10.1093/geronb/56.2.P103

Kensinger, E. A., & Gutchess, A. H. (2017). Cognitive aging in a social and affective context: Advances over the past 50 years. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, **72**, 61–70. doi:10.1093/geronb/gbw056

Krendl, A. C., Ambady, N., & Kensinger, E. A. (2015). The dissociable effects of stereotype threat on older adults' memory encoding and retrieval. *Journal of Applied Research in Memory and Cognition*, **4**, 103–109. doi:10.1016/j.jarmac.2015.02.001

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. doi:10.1037/a0038586

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. doi:10.1177/0956797612437607

Mitchell, K. J., Johnson, M. K., & Mather, M. (2003). Source monitoring and suggestibility to misinformation: Adult age-related differences. *Applied Cognitive Psychology*, **17**, 107–119. doi:10.1002/acp.857

Popham, L. E., & Hess, T. M. (2015). Age differences in the underlying mechanisms of stereotype threat effects. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, **70**, 225–34. doi:10.1093/geronb/gbt093

Roediger, H. L., Jacoby, J., & McDermott, K. (1996). Misinformation effects in recall: Creating false memories through repeated retrieval. *Journal of Memory and Language*, **35**, 300–318. doi:10.1006/jmla.1996.0017

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

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*. Pittsburgh, PA: Psychology Software Tools.

Seibt, B., & Förster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, **87**, 38–56. doi:10.1037/0022-3514.87.1.38

Shipley, W. C. (1946). *Institute of Living Scale*. Los Angeles, CA: Western Psychological Services.

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**, S206–S215. doi:10.1093/geront/gnx057

Stein, R., Blanchard-Fields, F., & Hertzog, C. (2002). The effects of age-stereotype priming on the memory performance of older adults. *Experimental Aging Research*, **28**, 169–181. doi:10.1080/03610730252800184

Thomas, A. K., & Bulevich, J. B. (2006). Effective cue utilization reduces memory errors in older adults. *Psychology and Aging*, **21**, 379–389. doi:10.1037/0882-7974.21.2.379

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

Thomas, A. K., Gordon, L. T., & Bulevich, J. B. (2014). Uniting theory to empirical evidence: How to understand memory of the elderly witness. In M. P. Toglia, D. F. Ross, J. Pozzulo, & E. Pica (Eds.), *The elderly eyewitness in court* (pp. 308–335). New York, NY: Psychology Press.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, **37**, 498–505. doi:10.3758/BF03192720

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

Yonelinas, A. P., & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory & Cognition*, **40**, 663–680. doi:10.3758/s13421-012-0205-5

Zuber, S., Ihle, A. Blum, A., Desrichard, O., & Kliegel, M. (2017). The effect of stereotype threat on age differences in prospective memory: Differential effects on focal versus nonfocal tasks. *Journal of Gerontology, Series B: Psychological Science and Social Science*. doi:10.1093/geronb/gbx097