THE GERONTOLOGICAL SOCIETY OF AMERICA®

The Gerontologist
cite as: *Gerontologist*, 2017, Vol. 57, No. S2, S193–S205
doi:10.1093/geront/gnx056

OXFORD

Literature Review

Stereotype Threat Effects on Older Adults' Episodic and Working Memory: A Meta-Analysis

Bonnie Armstrong, MA, Sara N. Gallant, MA,* Lingqian Li, MA, Khushi Patel, BA, and Brenda I. Wong, MA

Department of Psychology, Ryerson University, Toronto, Canada.

*Address correspondence to: Sara N. Gallant, MA, Ryerson University, 350 Victoria Street, Toronto, ON, Canada, M5B 2K3.
E-mail: s7gallan@psych.ryerson.ca

Received September 10, 2016; Editorial Decision Date March 30, 2017

Decision Editor: Catherine W. Gillespie, PhD

Abstract

Background and Objectives: Prior research has shown that exposure to negative age-based stereotype threat (ST) can undermine older adults' memory performance. The objective of the current meta-analysis was to examine the reliability and magnitude of ST effects on older adults' episodic and working memory performance—two forms of memory that typically show the greatest age-related declines. In addition, we examined potential moderators of age-based ST including type of ST manipulation, type and timing of memory task, participant age and education level.
Research Design and Method: A total of 23 samples for episodic memory and 15 samples for working memory were derived from 19 published and 4 unpublished articles and analyzed in two separate meta-analyses.
Results: Analyses revealed a reliable effect of ST on both older adults' episodic (d = 0.253) and working memory performance (d = 0.373). Interestingly, the age-based ST effect was only significant when blatant ST manipulations were used with episodic memory tasks or when subtle ST manipulations were used with working memory tasks. Moreover, within episodic memory, the ST effect only reached significance for recall but not cued-recall or recognition performance, and for immediate but not delayed tests of memory. Neither age nor level of education moderated the association between ST and older adults' memory performance.
Discussion and Implications: These results highlight the vulnerability of both older adults' episodic and working memory performance to age-based ST. When measuring older adults' memory performance in a research context, we must therefore be wary of exposing participants to common stereotypes about aging and memory.

Keywords: Stereotype threat, Aging, Episodic memory, Working memory

The population is graying as baby boomers age and we continue to live longer lives (World Health Organization, 2015). Despite our aging society, however, many individuals have a rigid schema of the aging process, often resulting in the automatic and implicit categorization of older adults according to their physical and functional attributes (Perdue & Gurtman, 1990). While some may attribute positive qualities to older adults, such as being approachable and warm, the greater assumption involves a lack of competence (Cuddy, Norton, & Fiske, 2005). These perceptions of aging have shown to develop in children as young as 3-years old (Seefeldt, Jantz, Galper, & Serlock, 1977) and are perpetuated throughout the lifespan by the negative portrayal of older adults in the media and use of colloquial terms such as "senior's moment" to describe lapses in memory (Vasil & Wass, 1993). Consequently, older adults are often viewed as incapable, slow, or forgetful, resulting in social categorizations which reinforce the prevalence of age-related stereotypes in society (Cuddy, Norton, Fiske, 2005; Hummert, 2011).

Given the prevalence of age-based stereotypes, the question arises as to how the endorsement of such negative beliefs might impact behaviour. While a large body of research documents age-related declines on laboratory-based memory tasks (e.g., Craik & McDowd, 1987; Park, 2000), there is considerable evidence that at least some of older adults' memory decrements may be accounted for by exposure to stereotypes about aging and memory (for a review see Barber & Mather, 2014). This phenomenon, referred to as stereotype threat (ST), describes a predicament in which a threat is "in the air" and the targeted individual is aware that their performance could confirm a negative stereotype about their group (Steele & Aronson, 1995; Steele, 1997). This was illustrated in Steele and Aronson's (1995) seminal studies in which African American students performed below their potential on the Scholastic Aptitude Test when stereotypes about their intellectual ability were emphasized. In contrast, those who were not exposed to the threat performed on par with their Caucasian peers. In further refining the ST effect, Steele and Aronson (1995) highlighted the importance of group identity and social context such that the stigma must be self-relevant and the setting must pose a risk for performance to confirm the stereotype.

In the context of aging, studies that experimentally induce ST in older adults by exposing them to age-related stereotypes about their mental or physical functioning have similarly shown decrements to performance relative to a non-threatened group (e.g., Hess, Auman, Colcombe, & Rahhal, 2003). This includes a recent meta-analysis demonstrating that negative age-based stereotypes are detrimental to older adults' performance (Lamont, Swift, & Abrams, 2015). Specifically, across a variety of performance domains such as cognitive, memory, physical, skill acquisition, and driving performance, Lamont et al.'s results show that age-based ST reliably impairs older adults' performance (d = 0.28). Although, this evidence suggests that activation of age-related stereotypes can impair older adults' memory (d = 0.21 in the memory domain, Lamont et al., 2015), whether different forms of memory are similarly impacted by ST has yet to be determined. Thus, diverging from prior investigations (e.g., Lamont et al., 2015), the goal of the current meta-analysis is to narrow the focus by elucidating the specific impact of age-related ST within episodic and working memory.

These two types of memory were of particular interest as older adults typically show the greatest declines in these memory processes, whereas other types of memory such as semantic and procedural memory, are relatively preserved in old age (Hedden & Gabrieli, 2004; Nyberg, Lövdén, Riklund, Lindenberger, & Bäckman, 2012). Moreover, research has shown that the effect of ST on memory is most prominent when older adults place high value on their memory abilities (Hess et al., 2003) or when they show lower memory self-efficacy (Desrichard & Köpetz, 2005; ST induced through describing tasks as memory or non-memory tasks in this study). Personal experience with deterioration in episodic and working memory abilities or fear of age-related memory loss (Dark-Freudeman, West, & Viverito, 2006; Laditka et al., 2011) might therefore make these processes even more vulnerable to the effects of ST in old age.

Episodic and working memory are also the two memory systems that have been most extensively studied in the field of age-based ST; however, a stronger focus has been placed on examining the impact of ST on episodic relative to working memory performance. Whereas many studies support the impairing effects of age-based ST on older adults' episodic memory (e.g., Barber & Mather, 2013a Study 1; Barber, Mather, & Gatz, 2015 gain-based condition; Eich, Murayama, Castel, & Knowlton, 2014; Fernández-Ballesteros, Bustillos, & Huici, 2015; Hess et al., 2003; Hess & Hinson, 2006), an examination of the literature reveals that research findings on working memory are relatively mixed. Some studies have shown that when negative age-related stereotypes about memory performance are activated, older adults' working memory performance is undermined relative to controls (e.g., Abrams, Eller, & Bryant, 2006; Barber & Mather, 2013b gain-based condition; Mazerolle, Régner, Morisset, Rigalleau, & Huguet, 2012; Swift, Abrams, & Marques, 2013), whereas other studies have failed to diminish older adults' working memory under ST conditions (e.g., Hess, Hinson, & Hodges, 2009; Popham & Hess, 2015). As such, although ST effects appear to be present and reliable in older adults' episodic memory (Lamont et al., 2015), whether ST similarly affects their working memory is less clear.

Taken together, the literature discussed suggests that older adults' memory can be negatively affected by age-related ST. Although, the effect of ST on memory has been investigated in Lamont and colleagues' (2015) meta-analysis, working memory studies were not specifically examined. Therefore, in the current study we used a meta-analytic approach to examine whether there is indeed a reliable impact of ST on both older adults' episodic and working memory performance. This also provided us with an opportunity to conduct a more current analysis of age-related ST effects on episodic memory. This was necessary given the rapid growth of the age-based ST literature since Lamont et al. (2015) terminated their search in February 2013. Specifically, nine papers (six on episodic memory, two on working memory, and one paper on both memory types) have been published between February 2013 and March 2016, the date the search was terminated for the current meta-analysis.

In addition to examining the reliability of ST effects on episodic and working memory performance, we examined how certain methodologic and demographic factors might moderate the ST effect on older adults' memory including the type of ST manipulation, type of memory task, timing of memory task, participant age, and participant education level. The following section provides an overview and rationale for examining these potential moderators.

Stereotype Threat and Older Adults' Memory: Moderating Factors

Type of ST Manipulation

In the majority of age-based ST studies reviewed, age-relevant stereotypes were activated and older adults were asked to perform in a domain relevant to the stigma. Their performance was then compared to young adults or older adults not exposed to ST. It is important to note, however, that the cues used to elicit ST across these studies vary drastically, particularly in terms of the degree of exposure to the age-related stereotype. According to Steele (1997), these methodological differences are important to consider as the relevance of the stereotype to both the individual and the testing environment can account for variation in the magnitude of ST effects. As such, in the current review, we distinguish between studies that use blatant versus subtle cues to induce age-based ST (for similar distinctions see Lamont et al., 2015; Nguyen & Ryan, 2008).

Blatant ST manipulations involve a direct reference to the stereotyped groups' inferiority in the performance domain. In such studies, salient cues are used so that participants are aware of the stigma and the fact that their performance could validate the stereotype about their group. For instance, Hess and colleagues (2003) asked participants to read fake news articles that either emphasized the negative effects of aging on memory or described positive effects of aging on memory. In the negative articles, direct evidence was presented confirming that older adults have poorer memory skills than young adults. In the positive articles, age-based stereotypes were counteracted by suggesting that older adults have control over age-related changes in memory. This manipulation was effective at producing ST effects such that older adults in the negative ST group showed reduced episodic memory performance relative to young adults, older adults in the positive ST condition, and a "no-threat" control group that did not read the articles. Other blatant ST manipulations have involved telling older ST participants that the study's objective is to evaluate memory—a cognitive domain known to decline with age (Bouazzaoui et al., 2016), or that young adults perform better on the memory test they are about to complete (e.g., Bensadon, 2010; Fernández-Ballesteros et al., 2015; Hess, Emery, & Queen, 2009).

Another class of age-based ST methods involves a subtler approach, in which the threat is activated by manipulating the context of the study. For example, ST participants may be told that the memory task is diagnostic of memory ability (e.g., Kang & Chasteen, 2009), that their memory performance will be compared to that of young adults (e.g., Mazerolle et al., 2012; Wong, 2014), or they may be asked to complete the memory task in the same room as a younger adult (e.g., Kang & Chasteen, 2009). Importantly, across these subtle manipulations, the inferior memory performance of older relative to young adults is not directly communicated. In this way, the threat is still "in the air" so that participants are aware that their performance could validate the stereotype about their memory.

In addition to blatant and subtle manipulations, some studies have used subliminal cues to elicit age-based ST effects in older adults (for a meta-analysis on priming and age-related ST, see Horton, Baker, Pearce, & Deakin, 2008). Examples include priming tasks in which positive or negative age-related stereotype words (e.g., aged, old, feeble, confused, senile, and forgot) are presented to participants at speeds that allow for visual perception of the stimulus but not conscious detection (e.g., Krendl, Ambady, & Kensinger, 2015; Levy, 1996; Levy & Leifheit-Limson, 2009; Stein, Blanchard-Fields, & Hertzog, 2002). Different from blatant and subtle manipulations, these subliminal ST cues operate below the level of consciousness, meaning that participants are unaware of the stereotype activation. For this reason, studies utilizing priming techniques are not considered in the current paper.

In Lamont and colleagues (2015) meta-analysis, they showed that subtle ST manipulations negatively impacted older adults' task performance relative to blatant ST manipulations. The authors suggest a greater sensitivity to subtle over blatant ST manipulations as subtle cues of ST may increase the perceived uncertainty that a threat is present (Lamont et al., 2015; Schmader, Johns, & Forbes, 2008). This may occupy and eventually deplete limited cognitive resources, negatively impacting performance on tasks that heavily depend on cognitive resources. Lamont and colleagues (2015) thus suggest that older adults experience greater decrements to performance when exposed to subtle ST cues than when there is no ST cue. In the context of the current study, we aimed to replicate this finding and determine if it extends to working memory. Given Lamont and colleagues (2015) findings, we hypothesized that ST manipulations classified as subtle would have a greater negative impact on both episodic and working memory performance relative to ST manipulations classified as blatant.

Type of Memory Task

Within episodic memory, we further examined how the effect of ST might be moderated by the type of memory test used, including free recall, cued-recall, or recognition based tasks. Research suggests that recall tasks rely more on controlled, self-initiated processes, in which learned information is retrieved entirely without support from environmental cues (e.g., Craik, 1983, 1986). This is in contrast to recognition-based memory in which environmental cues can be used to inform the decision as to whether an item has been encoded before. In the context of aging, the ability to use recall to retrieve information from memory has shown to decline at a rate that is disproportionately greater than that of recognition memory (Danckert & Craik, 2013). As such, the goal of this analysis was to determine if the effect of ST would be greater for those studies that used recall relative to recognition based memory tests.

Timing of Memory Task

Although episodic memory in ST studies has been tested both immediately after encoding (i.e., immediate test; e.g., Barber et al., 2015) and after delayed periods (i.e., delayed test; e.g., Hess, Emery, & Queen, 2009), it is unclear whether ST disrupts both types of retrieval equally. By examining the potential moderating role of testing time point on ST effects, we can gain insight into whether early encoding-based processes are differentially impacted by ST than time-dependent consolidation-based processes. If delayed retrieval is more affected by ST than immediate retrieval, then the effect of stereotypes might not only impact older adults' encoding of materials, but also their consolidation processes that continue after the encoding phase. However, if immediate retrieval shows a greater impact of ST, this would provide evidence that encoding-based processes may be more sensitive to ST than later consolidation processes.

Demographic Characteristics

A final goal of the current meta-analysis was to explore how participants' demographic characteristics might impact older adults' memory performance under ST. First, there is evidence that age-related ST is more detrimental to "young-olds" (ages 60–70), but exerts less of an influence on "old–olds" (70s and above; e.g., Hess & Hinson, 2006; Hess, Hinson, & Hodges, 2009). According to Hess and colleagues (2006,2009), the seventh decade of life (ages 60 and above) is considered old age in the Western culture. This "old-age" label may be especially salient for individuals new to their identity within the age group, thus making young–olds most susceptible to ST effects. Participants' level of education has also shown to moderate the effect of ST on memory; however, findings to date are mixed. Hess, Hinson, and Hodges (2009) found that the effect of ST on memory is stronger among participants with higher education than those with lower education, whereas Andreoletti and Lachman (2004) found education buffer against the effect of ST. In light of this evidence, we examined the potential moderating role of these demographic factors on the effect of ST within both episodic and working memory.

Method

Literature Search and Selection

The initial search was carried out using "PsycInfo," "Pubmed," "Web of Science," and "Medline" databases. The Boolean search terms used included "older adults" OR "seniors" OR "elderly" OR "late life" OR "age*" OR "aging" OR "age" OR "later life") AND ("stereotype threat" OR "stereotype-attitudes" OR "beliefs of aging") AND ("memory" OR "episodic memory" OR "working memory" OR "recall" OR "recognition" OR "remember"). This search returned 1,194 candidate studies for screening (as of March 9, 2016). To locate grey literature such as unpublished theses, dissertations, manuscripts or conference proceedings, additional searches were carried out on the Educational Resources Information Center (ERIC) database and Google Scholar. Authors who published multiple articles on age-related ST effects on memory were also contacted for any unpublished manuscripts.

Criteria for Inclusion

A series of eligibility criteria were established to screen for selected studies. First, only studies that manipulated ST by directly activating age-related stereotypes were included, such as asking participants to read articles that describe age-related memory declines (e.g., Hess & Hinson, 2006) or stating that the purpose of the study was to confirm age-related declines in memory (e.g., Barber et al., 2015). Second, to ensure that selected studies were examining the influence of ST on the older adult population, only study samples that contained participants of at least a mean age of 60 were selected. Studies that contained participants with a mean age younger than 60 years (e.g., O'Brien, & Hummert, 2006) were excluded. Studies were excluded if they did not include age-based stereotypes (e.g., Rahhal, Hasher, & Colcombe, 2001; Desrichard, & Kopetz, 2005), if ST was induced after studying the target items (e.g., Wong & Gallo, 2015), or if the ST was induced with subconscious ST manipulations such as priming techniques (e.g., Hess, Hinson, & Statham, 2004; Krendl et al., 2015; Stein et al., 2002). Finally, if studies did not report the mean scores and standard deviations or standard errors of memory performance (e.g., Levy & Leifheit-Limson, 2009), the authors were contacted with a request for data; if no response was received, the study was excluded from the meta-analysis. Figure 1 illustrates a detailed flow diagram of the selection procedure.

Two rounds of eligibility screening were carried out before articles were selected for this meta-analysis according to the inclusion criteria. During the first round, we inspected the title and abstract of each candidate article, which resulted in 35 candidate articles. In the second round, each candidate article was then closely examined

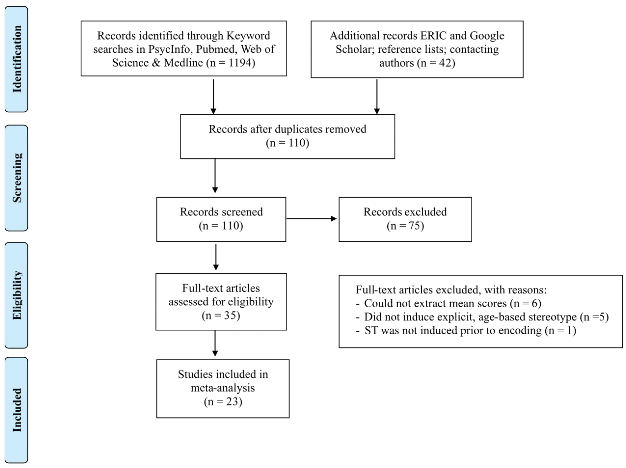

Figure 1. Flow diagram of article selection procedure.

to further exclude studies that did not fit our inclusion criteria. At each round of screening, articles were screened by at least two of the authors before decisions were made to exclude them. This resulted in a total of 19 published and 4 unpublished experimental studies included in the final analyses. Among these selected articles, 14 studies used episodic memory tasks only, 6 studies used working memory tasks only, and 3 studies used both episodic memory and working memory tasks (a total of 34 samples, 4 of which consisted of both episodic and working memory results).

Data Exaction and Management

As the focus of the current meta-analysis is memory performance in older adults, the dependent variable was the mean score of memory performance (raw number of items correctly recalled, proportion of recall accuracy, or corrected recognition scores for recognition tests) in experimental and control groups. In a few selected articles, the numerical values of means and standard deviations were not reported but instead illustrated in bar graphs and error bars. In these instances, we extracted the group means and standard errors via the online application WebPlotDigitizer (http://arohatgi.info/WebPlotDigitizer/app/), which estimated the numerical values of bars by comparing their relative length to the y-axis in the figure. We used Cohen's d as an estimate of effect size, which was calculated by dividing the difference in memory scores between control and experimental groups by the standard deviation, following the guidelines of Rosenthal (1995). All the calculations and meta-analyses were conducted using the Comprehensive Meta-Analysis (CMA) Software (Version 3.0).

In addition, the following data were extracted from selected articles: authorship, publication year, publication status, sample size, mean age of sample, mean education of sample, and whether ST was induced via blatant or subtle stereotype exposure. We adopted the same coding scheme of ST manipulation as Lamont and colleagues (2015). Studies that induced ST by presenting participants with "scientific findings" or "factual evidence" of age-related memory declines were coded as blatant (e.g., Hess & Hinson, 2006); studies were coded as subtle if they used indirect manipulations to induce threat, such as having older participants complete memory tasks in the same room with a young adult confederate (e.g., Kang & Chasteen, 2009).

For episodic memory tasks, type of memory task (e.g., free recall, cued-recall, or recognition) was also extracted, as well as whether a delay period was placed between learning and testing phases. Each study was independently reviewed by two authors, and all authors took part in coding. The interrater reliability was above ninety percent. Dispute in coding was re-coded by a third author and agreed by the entire team.

Results

Stereotype Threat Effects on Episodic Memory

A total of 23 samples were submitted to this meta-analysis (Table 1 for a list of the studies included). An overall random effects meta-analysis was conducted, which returned a significant main effect of ST induction on memory performance, with a mean weighted effect size of $d = 0.253$ ($SE = 0.091$, 95% CI = 0.074 to 0.433; $z = 2.771$, $p = 0.006$). Overall, the effect sizes were heterogeneous, $Q(22) = 50.648$, $p < .001$. See Figure 2 for detailed statistics and forest plot and Table 2 for summary of the moderator analyses.

Type of ST Manipulation

There was no statistically significant difference of ST effect between blatant and subtle ST manipulations in episodic memory, $Q_{\text{between}}(1) = 0.019$, $p = .889$. When examined separately, samples with blatant ST manipulations ($k = 20$) revealed a small but significant effect of 0.257, $z = 2.475$, $p = .013$. On the other hand, moderator analysis in the subtle ST studies did not reach significance, $d = 0.227$, $z = 1.213$, $p = .225$, possibly due to low sample size ($k = 3$).

Type of Memory Task

Overall, there was no statistically significant difference in the mean effect sizes between the different episodic memory tasks, $Q_{\text{between}}(2) = 0.099$, $p = .952$. However, an examination of the individual effects of the three types of memory tests revealed a significant effect of 0.237 on free recall memory performance ($k = 16$), $z = 2.150$, $p = .032$, whereas the ST effect was not significant for recognition tasks ($k = 4$), $d = .327$, $z = 1.206$, $p = .228$, nor cued-recall tasks ($k = 3$), $d = 0.230$, z = 0.879, $p = .380$. Due to low sample sizes in the recognition and cued-recall samples, no definitive conclusion could be drawn about the respective ST effects in these tasks.

Immediate Versus Delayed Memory

Another moderator in our analysis was the duration of delay between encoding and retrieval in memory tasks (i.e., immediate vs. delayed memory). The majority of study samples employed an immediate memory paradigm ($k = 20$). A few samples included both immediate and delayed tasks. To avoid interference from practice effects, only data from the immediate tasks were extracted. There were three samples that only included delayed tasks. For these tasks, the delay was 10 (2 study samples from Hess, Emery, & Queen, 2009) to 15 minutes (Kang & Chasteen, 2009) in duration. In both studies, ST was reinstated during the delay period, as participants were asked to indicate their age prior to the memory tests. Overall, there was no significant difference between immediate and delayed memory, $Q_{\text{between}}(1) = 0.281$, $p = .596$. However, examination of the individual effects of the two time points revealed a significant effect of ST on immediate memory ($k = 20$), $d = 0.228$, $z = 2.414$, $p = .016$, whereas the ST effect was not significant for delayed tests of memory ($k = 3$), $d = 0.426$,

Table 1. Summary of Selected Studies in the Episodic Memory Meta-Analysis

Study name	*N*	*d*	Age (years)	Education (years)	Memory Task	Delay	ST manipulation	Publication
Andreoletti & Lachman (2004)	50	0.14	69.3	—	Free recall	Immediate	Blatant	Experimental Aging Research
Barber & Mather (2013a) Study 1	31	0.88	70.42	—	Free recall	Immediate	Blatant	The Quarterly Journal of Experimental Psychology
Barber & Mather (2013a) Study 2	64	−0.64	71.81	16.78	Free recall	Immediate	Blatant	The Quarterly Journal of Experimental Psychology
Barber, Mather, & Gatz (2015): Gain-related	40	0.30	69.35	17.23	Free recall	Immediate	Blatant	J Gerontol B Psychol Sci Soc Sci
Barber, Mather, & Gatz (2015): Loss-related	40	0.17	69.73	18.32	Free recall	Immediate	Blatant	J Gerontol B Psychol Sci Soc Sci
Bensadon (2010)	61	0.29	67.10	15.80	Free recall	Immediate	Blatant	Unpublished
Bouazzaoui et al. (2016)	92	0.43	70.11	12.81	Free recall	Immediate	Blatant	Memory
Eich, Murayama, Castel, & Knowlton (2014): Early-aging	35	0.85	67.80	14.09	Cued-recall	Immediate	Blatant	Social Cognition
Eich, Murayama, Castel, & Knowlton (2014): Late-aging	35	−0.20	84.60	14.09	Cued-recall	Immediate	Blatant	Social Cognition
Fernández-Ballesteros, Bustillos, & Huici (2015)	90	0.31	61.63	—	Free recall	Immediate	Blatant	Experimental Aging Research
Haslam, Morton, Haslam, Varnes, Graham, & Gamaz (2012)	34	1.13	65.60	16.45	Free recall	Immediate	Blatant	Psychology and Aging
Hess, Auman, Colcombe, & Rahhal (2003)	32	0.71	68.90	15.60	Free recall	Immediate	Blatant	J Gerontol B Psychol Sci Soc Sci
Hess, Emery, Queen (2009): Deadline condition	45	0.89	71.04	16.60	Recognition	Delayed	Blatant	J Gerontol B Psychol Sci Soc Sci
Hess, Emery, Queen (2009): Unlimited condition	37	−0.31	70.86	16.24	Recognition	Delayed	Blatant	J Gerontol B Psychol Sci Soc Sci
Hess & Hinson (2006): Old–old	32	−0.12	80.00	—	Free recall	Immediate	Blatant	Psychology and Aging
Hess & Hinson (2006): Young–old	32	0.47	68.00	—	Free recall	Immediate	Blatant	Psychology and Aging
Hess, Hinson, & Hodges (2009): Old–old	51	−0.39	75.35	16.25	Free recall	Immediate	Blatant	Experimental Aging Research
Hess, Hinson, & Hodges (2009): Young–old	52	0.14	64.15	15.35	Free recall	Immediate	Blatant	Experimental Aging Research
Horton, Baker, Pearce, & Deakin (2010)	64	−0.17	67.33	14.30	Free recall	Immediate	Blatant	Educational Gerontology
Kang & Chasteen (2009)	42	0.68	71.50	15.10	Free recall	Delayed	Subtle	Int'l. J. Aging and Human Development
Mazerolle, Régner, Morisset, Rigalleau, & Huguet (2012)	110	0.12	69.01	—	Cued-recall	Immediate	Subtle	Psychological Science
Mazerolle, Régner, Rigalleau, & Huguet (2015)	38	−0.03	69.13	13.42	Recognition	Immediate	Subtle	Experimental Psychology
Wong (2014) Study 2	84	0.66	75.25	16.57	Recognition	Immediate	Blatant	Unpublished

Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andreoletti & Lachman (2004)	0.142	0.283	0.080	-0.413	0.697	0.500	0.617
Barber & Mather (2013a) Study 1	0.879	0.376	0.142	0.141	1.616	2.335	0.020
Barber & Mather (2013a) Study 2	-0.644	0.256	0.066	-1.146	-0.141	-2.511	0.012
Barber, Mather, & Gatz (2015): Gain-related	0.304	0.318	0.101	-0.319	0.928	0.956	0.339
Barber, Mather, & Gatz (2015): Loss-related	0.174	0.317	0.100	-0.447	0.795	0.549	0.583
Bensadon (2010)	0.286	0.258	0.066	-0.219	0.791	1.111	0.266
Bouazzaoui et al. (2015)	0.426	0.211	0.044	0.012	0.839	2.019	0.044
Eich, Murayama, Castel, & Knowlton (2014): Early-aging	0.845	0.353	0.125	0.154	1.537	2.395	0.017
Eich, Murayama, Castel, & Knowlton (2014): Late-Aging	-0.203	0.339	0.115	-0.867	0.462	-0.598	0.550
Fernandez-Ballesteros, Bustillos, & Huici (2015)	0.309	0.219	0.048	-0.120	0.737	1.412	0.158
Haslam, Morton, Haslam, Varnes, Graham, & Gamaz (2012)	1.134	0.370	0.137	0.410	1.859	3.069	0.002
Hess & Hinson (2006): Old-old	-0.118	0.354	0.125	-0.811	0.576	-0.333	0.739
Hess & Hinson (2006): Young-old	0.471	0.358	0.128	-0.231	1.174	1.315	0.188
Hess, Auman, Colcombe, & Rahhal (2003)	0.706	0.364	0.133	-0.008	1.420	1.937	0.053
Hess, Emery, & Queen (2009): Deadline	0.887	0.313	0.098	0.274	1.500	2.838	0.005
Hess, Emery, & Queen (2009): Unlimited	-0.305	0.331	0.109	-0.954	0.343	-0.923	0.356
Hess, Hinson, & Hodges (2009): Old-old	-0.385	0.283	0.080	-0.939	0.169	-1.361	0.174
Hess, Hinson, & Hodges (2009): Young-old	0.143	0.278	0.077	-0.402	0.688	0.514	0.607
Horton, Baker, Pearce, & Deakin (2010)	-0.167	0.250	0.063	-0.658	0.324	-0.666	0.506
Kang & Chasteen (2009)	0.677	0.317	0.101	0.055	1.299	2.134	0.033
Mazerolle, Regner, Morisset, Rigalleau, & Huguet (2012)	0.124	0.191	0.036	-0.250	0.498	0.648	0.517
Mazerolle, Regner, Rigalleau, & Huguet (2015)	-0.034	0.324	0.105	-0.670	0.602	-0.106	0.916
Wong (2014) Study 2	0.663	0.224	0.050	0.224	1.103	2.960	0.003
	0.253	0.091	0.008	0.074	0.433	2.771	0.006

Statistics for each study

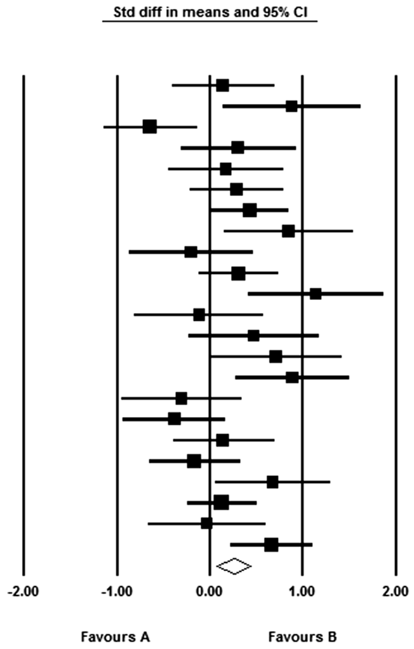

Figure 2. Effect size statistics and forest plot of selected articles in the episodic memory analysis.

Table 2. Meta-Analytic Results for Moderators in Episodic Memory Samples

	k	*d*	*SE*	95% CI for *d* Lower	95% CI for *d* Upper	*Q*
ST manipulation						
Blatant	20	0.257*	0.104	0.054	0.461	47.665***
Subtle	3	0.227	0.187	−0.140	0.595	2.939
Type of memory task						
Cued recall	3	0.230	0.262	−0.283	0.743	4.912
Free recall	16	0.237*	0.110	0.021	0.453	34.153**
Recognition	4	0.327	0.271	−0.204	0.857	10.146*
Delay in memory task						
Delayed	3	0.426	0.362	−0.284	1.136	7.681*
Immediate	20	0.228*	0.094	0.043	0.412	41.570**

Note: k = number of samples included; d = point estimates of population effect size; Q = homogeneity of variance ($d = k-1$).
$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. Two-tailed.

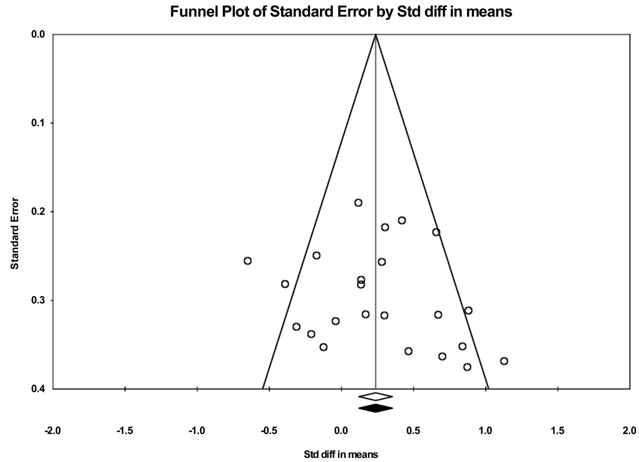

Figure 3. Funnel plot of on the distribution of standard errors around the mean effect size. The symmetrical shape of the distribution shows no signs of publication bias in the samples used in episodic memory analysis.

$z = 1.176$, $p = .239$. Despite the lack of statistical significance in this analysis, it should be noted that the reported effect size for delayed memory tests was larger than that of immediate memory tests. The fact that the effect size for delayed memory tests failed to reach our threshold for statistical significance is likely the result of a lack of power due to the small number of samples within this condition. These results should therefore be interpreted with caution as more samples are required to make definitive conclusions.

Age and Education

Two separate random model meta-regressions were conducted to examine the potential moderating role of age and education level in the observed ST effect. In the current study samples, five studies (six samples) did not report mean years of education, and hence these samples were excluded from our analysis on education. The first model used age as the predictor, with the result suggesting that age could not explain variance in d ($df = 21$, $\beta = -0.03$, $p = .165$). Level of education was the predictor in the second model, and this variable also could not account for the variance in d ($df = 15$, $\beta = -0.001$, $p = .988$).

Risk of Publication Bias

Due to the low number of unpublished manuscripts ($k = 2$), we could not directly run a moderator analysis to examine the effect of publication status on the mean weighted effect size. Instead, a publication bias analysis was run in CMA. The CMA program was used to produce the funnel plot (Figure 3), which did not generate any imputed sample. The funnel plot is designed to check for publication bias and displays the distribution of standard errors around the mean effect size, with each dot representing a single study. In the absence of any bias, the studies with high precision should cluster close to the mean whereas those with low precision (e.g., resulting from smaller samples) will be scattered widely at the bottom, creating a symmetrical inverted funnel shaped distribution. If publication bias is present, the plot will deviate from this shape with an asymmetrical distribution (Sedgwick, 2013). Since assessment of symmetry in the funnel plot is largely subjective, we used Egger's regression test of the intercept to formally test the significance of asymmetry in the funnel plot, which failed to detect any significant bias, $(B0) = 1.384$, $t(21) = 0.875$, p(1-tailed) $= .196$. In addition, after using the Duval and Tweedie's trim and fill procedure to adjust for any unrepresented samples, the imputed point estimate remained unchanged as $d = 0.253$ (0.074, 0.433).

Stereotype Threat Effects on Working Memory

Fifteen samples were submitted to this meta-analysis (Table 3 for a list of the studies included). The results returned a significant random effect, with a mean weighted effect size of $d = 0.373$ ($SE = 0.170$, 95% CI = 0.040 to 0.707; $z = 2.192$, $p = .028$). The effect sizes were significantly heterogeneous, $Q(14) = 71.750$, $p < .001$. See Figure 4 for detailed statistics and forest plot.

Type of ST Manipulation

There was a significant difference in mean effect sizes between blatant and subtle ST manipulations in working

Table 3. Summary of Selected Studies in the Working Memory Meta-analysis

Study Name	*N*	*d*	Age (years)	Education (years)	ST Manipulation	Publication
Abrams, Eller, & Bryant (2006): Less contact	45	1.77	74.81	—	Subtle	Psychology and Aging
Abrams, Eller, & Bryant (2006): More contact	47	0.50	74.81	—	Subtle	Psychology and Aging
Barber & Mather (2013b) Study 1a: Gain-based	28	1.00	69.65	16.27	Blatant	Psychological Science
Barber & Mather (2013b) Study 1a: Loss-based	28	−0.98	68.93	15.27	Blatant	Psychological Science
Barber & Mather (2013b) Study 1b: Gain-based	28	0.99	67.50	15.77	Blatant	Psychological Science
Barber & Mather (2013b) Study 1b: Loss-based	28	−1.04	63.72	15.93	Blatant	Psychological Science
Barber, Niblett & Seliger (n.d.)	29	1.02	68.97	16.48	Subtle	Unpublished
Hess, Hinson, & Hodges (2009): Old–old	51	−0.32	75.69	14.74	Blatant	Experimental Aging Research
Hess, Hinson, & Hodges (2009): Young–old	52	−0.02	64.18	16.04	Blatant	Experimental Aging Research
Kominsky (2003)	40	0.13	75.03	—	Blatant	Unpublished
Mazerolle, Regner, Morisset, Rigalleau, & Huguet (2012)	110	0.76	69.01	—	Subtle	Psychological Science
Popham & Hess (2015)	63	−0.20	70.95	16.45	Blatant	J Gerontol B Psychol Sci Soc Sci
Swift, Abrams, & Marques (2013)	80	0.98	76.5	15.00	Subtle	J Gerontol B Psychol Sci Soc Sci
Wong (2014) Study 1	84	0.44	74.64	17.12	Blatant	Unpublished
Wong (2014) Study 2	84	0.49	75.25	16.57	Blatant	Unpublished

Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Abrams, Eller, & Bryant (2006): More contact	0.501	0.298	0.089	-0.083	1.084	1.680	0.093
Abrams, Elller, & Bryant (2006): Less contact	1.773	0.352	0.124	1.083	2.462	5.037	0.000
Barber & Mather (2013b) Study 1a: Gain	0.996	0.401	0.161	0.210	1.781	2.485	0.013
Barber & Mather (2013b) Study 1a: Loss	-0.983	0.400	0.160	-1.768	-0.199	-2.458	0.014
Barber & Mather (2013b) Study 1b: Gain	0.985	0.400	0.160	0.200	1.769	2.461	0.014
Barber & Mather (2013b) Study 1b: Loss	-1.038	0.403	0.162	-1.827	-0.249	-2.578	0.010
Barber, Niblett, & Seliger (n.d.)	1.023	0.395	0.156	0.249	1.798	2.589	0.010
Hess, Hinson, & Hodges (2009): Old-old	-0.321	0.282	0.079	-0.873	0.232	-1.138	0.255
Hess, Hinson, & Hodges (2009): Young-old	-0.019	0.278	0.077	-0.563	0.525	-0.068	0.946
Kominsky (2003)	0.126	0.317	0.100	-0.494	0.747	0.400	0.689
Mazerolle, Regner, Morisset, Rigalleau, & Huguet (2012)	0.763	0.198	0.039	0.376	1.150	3.862	0.000
Popham & Hess (2013)	-0.204	0.253	0.064	-0.699	0.291	-0.807	0.420
Swift, Abrams, & Marques (2013)	0.984	0.237	0.056	0.520	1.448	4.155	0.000
Wong (2014) Study 1: Reading Span	0.436	0.221	0.049	0.004	0.869	1.977	0.048
Wong (2014) Study 2: Digit Span	0.485	0.221	0.049	0.051	0.919	2.191	0.028
	0.373	0.170	0.029	0.040	0.707	2.192	0.028

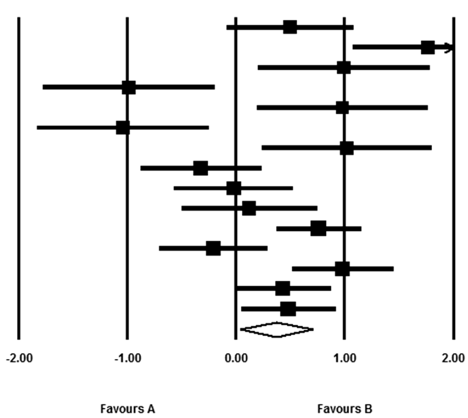

Figure 4. Effect size statistics and forest plot of selected articles in the working memory analysis.

memory, $Q_{between}(1) = 11.852$, $p = .001$. In particular, samples with subtle ST manipulations ($k = 5$) yielded a significantly larger effect size of 0.964, $z = 5.195$, $p < .001$, than the samples with blatant ST manipulation ($k = 10$), $d = 0.059$, $z = 0.315$, $p = .753$.

Age and Education

Similar to the episodic memory analyses, age and years of education were included as predictors in two random model meta-regressions to examine the role of these variables in the effect of ST on working memory. Mean age was reported in all working memory studies included in this meta-analysis. Our model revealed that age could not account for the variance in d ($df = 13$, $\beta = 0.063$, $p = .141$). Years of education were reported in seven studies (11 samples) in this review and also could not explain variance in d ($df = 9$, $\beta = 0.242$, $p = .422$).

Risk of Publication Bias

The number of samples derived from unpublished manuscripts was low ($k = 4$) and so moderator analyses were not run to examine the effect of publication status on the mean weighted effect size. The CMA program was used to produce the funnel plot (Figure 5). Egger's regression test of revealed no significant bias, $(B0) = -1.345$, $t(13) = 0.538$, p(1-tailed) = .300. After Duval and Tweedie's trim and fill, the imputed point estimate remained at $d = 0.373$ (0.040, 0.707).

Table 4. Moderator Analysis in Working Memory Samples

	k	d	SE	95% CI for d Lower	Upper	Q
ST manipulation						
Blatant	10	0.059	0.186	−0.307	0.424	34.408***
Subtle	5	0.964***	0.186	0.600	1.328	8.625

Note: k = number of samples included; d = point estimates of population effect size; Q = homogeneity of variance ($df = k-1$).
*$p < .05$; **$p < .01$; ***$p < .001$. Two-tailed.

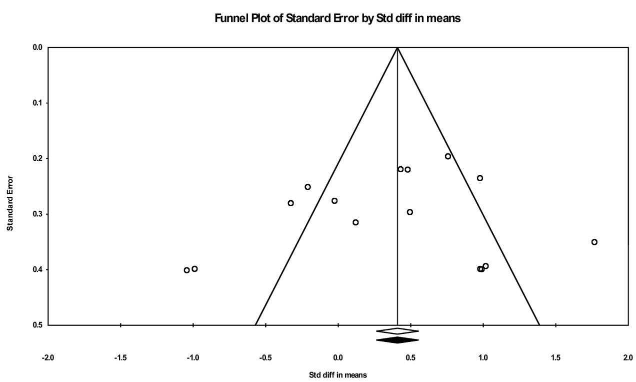

Figure 5. Funnel plot of on the distribution of standard errors around the mean effect size. The symmetrical shape of the distribution shows no signs of publication bias in the samples used in working memory analysis.

Discussion

The objective of the current meta-analysis was to provide a detailed examination of age-based ST effects across older adults' episodic and working memory performance. Overall, our meta-analysis of the literature revealed that age-based ST has a reliable negative effect on older adults' episodic and working memory performance. Specifically, studies investigating the impact of ST on older adults' episodic memory yielded a small effect size of $d = 0.25$, whereas those examining working memory produced a small-to-medium effect size of $d = 0.37$ (Cohen, 1988).

Examining the varying degrees of ST manipulations and their impact on older adults' episodic and working memory performance was of importance in the current paper given evidence that both stereotype relevance and differences in testing environments may account for variability in ST (Steele, 1997). Lamont and colleagues (2015) recently showed that subtle ST manipulations negatively impact older adults' cognitive and memory performance. However, memory in this paper was broadly defined as a performance domain and important differences across memory types (i.e., episodic and working memory) were overlooked. Our results indicate that working memory performance in older adults was only impacted by subtle manipulation of ST but not by blatant induction of ST, in which factual statements about age-based declines in tasks were explicitly conveyed. This finding aligns with Lamont et al.'s (2015) hypothesis that ambiguity in subtle ST induction may increase distracting thoughts during the experiment, which consumes cognitive resources from the task at hand and may impair performance on tasks that require working memory abilities. However, because age stereotypes are explicitly defined under blatant ST manipulations, participants are less likely to engage in distracting thoughts, which Lamont et al. hypothesized should be less detrimental to older adults' performance. Lamont and colleagues further suggested that the effect of blatant and subtle ST inductions might differ in tasks that are less reliant on working memory abilities. Our results provided some support for this speculation, as older adults' episodic memory was also impaired by ST when threat was induced blatantly. It is important to note, however, that only three samples used subtle ST manipulations in the investigation of ST effects on episodic memory. Subtle ST manipulations may therefore be underrepresented compared to the 20 blatant manipulations submitted to this analysis. We were therefore not able to draw confident conclusions on the effect of subtle ST induction on episodic memory in the current analysis. Further, these results shed light on the minimal literature examining how older adults' episodic memory is affected by subtle ST manipulations. Future research should test how the continuous varying degrees of

ST exposure (e.g., extremely subtle to extremely blatant ST manipulations) affect different memory types in old age as well as other stereotyped domains.

Interestingly, we did not find evidence that age or education moderated the association between ST and memory performance in old age, as suggested in previous studies (e.g., Andreoletti & Lachman, 2004; Hess & Hinson, 2006; Hess, Hinson, & Hodges, 2009). These results could imply that ST effects are pervasive and occur regardless of age or education level; however, we acknowledge that the age ranges and years of education in the studies included in our analysis were relatively conservative. More specifically, the mean age in most studies ranged from 65 to 75 years of age, and the average amount of years of education was between 14 and 16, equivalent to some levels of post-secondary education. To further verify the effects of age and education, future studies should include more participants in the old-old group (ages 75 and above), and examine samples with varied levels of education (less than or more than post-secondary education). This would allow for a better understanding of how specifically these demographic characteristics impact ST effects.

Within studies on episodic memory, we also analyzed the difference in ST effects across different types of memory tests. This allowed us to determine whether ST effects might be larger in tasks that require greater cognitive effort, such as in recall tasks that rely heavily on self-initiated processing, relative to recognition tasks which rely more on environmental supports to cue memory. We found that induction of negative stereotypes had a significant effect on older adults' free recall performance, which was the most common type of memory assessment in the age-based ST literature. In contrast, the effect of ST on memory was not significant under cued-recall or recognition tests. These findings suggest that older adults' memory may be more impaired by the induction of ST when there is a need to rely on controlled, self-initiated processing in memory tasks. This is consistent with Craik's (1983, 1986) idea that older adults show the most impairment when their memory is measured using recall tests relative to cued-recall or recognition tests, because performance on free recall tests rely primarily on controlled processes with very little environmental support. Cued-recall and recognition tests, on the other hand, provide cues (e.g., reminders of stimulus through use of cues or response options) to support retrieval of learned information. We would like to point out that there were only four study samples in which memory was measured using recognition tests and three study samples in which cued-recall tests were employed. The standard error of the average effect size across these studies was also relatively high. Therefore, it is crucial to have more empirical research on the effect of ST on cued-recall and recognition memory.

Finally, we examined the difference in ST effects across tests of immediate and delayed episodic memory. The goal of this analysis was to assess whether the effect of ST varied depending on the stage of memory formation, with immediate tests of memory reflecting relatively early encoding-based processes and delayed tests reflecting later consolidation- or storage-based processes. Results revealed a robust effect of ST in immediate tests of memory, whereas the effect of ST in delayed memory failed to reach significance. These findings may suggest that encoding-based processes in the formation of episodic memories are more sensitive to the impact of ST, whereas time-dependent consolidation processes are more resistant to the effects of ST. The results can also not be accounted for by a differential impact of age on immediate versus delayed memory as older adults show equivalent rates of decline across both testing time points (e.g., Haaland, Price, & Larue, 2003). However, the results should be interpreted with caution due to the small number of delayed memory tests included in our analysis and the high standard error. Further research is thus required to elucidate the impact of age-related ST across different stages of memory.

While we present a novel finding of a reliable ST effect in older adults' working memory, future research is required to further refine our understanding of ST effects during aging. Given that not all forms of memory are equally impacted by the aging process, future research should consider the effect ST has on forms of memory that are differentially impacted by age. For instance, whereas the episodic memory system shows age-related deficits, semantic memory (e.g., vocabulary or general knowledge) and implicit memory for information have shown to be relatively preserved with aging (Mather, 2010). As such, the magnitude of age-related ST effects may be reduced for the types of memory that remain intact in late life. Investigating whether the magnitude of ST effects varies across cultures with differing perceptions of older adults would also inform socio-psychological theories concerning the impact of ST. Additionally, it is also important to determine whether ST effects are generalizable to pathological populations such as those with mild cognitive impairment and dementia.

Limitations

Although, we found a reliable effect of age-based ST on episodic and working memory performance in older adults, there are some limitations to the current meta-analysis. First, the current literature does not allow us to statistically compare differences in effect sizes between episodic and working memory. This limitation is due to the lower number of studies on the effect of ST on working memory, as there is a stronger emphasis in current literature to examine the ST effect on episodic memory. Adding to the problem, almost half of the study samples in our meta-analysis on working memory were also included in our meta-analysis on episodic memory, hence there were not sufficient study samples that investigated the two types of memory independently to allow us to run statistical comparisons between the ST effects on the two types of memory. Despite this limitation, the finding that both episodic and working memory in older adults are impaired by the effect of age-based ST is by itself an important contribution to the field, particularly considering working memory has not been extensively investigated in the current literature. To our knowledge, this paper is the first meta-analysis that reviewed the effects of ST on the two types of memory separately.

Additionally, this meta-analysis may be restricted by the context of the articles reviewed, as there is visible overlap in authorship, such that some studies reviewed for this analysis were conducted by the same researchers and the majority studies were conducted in North America. Therefore, the results reported here may not be generalizable beyond these contexts. This is important to consider, given that varying cultures may have different ideations of aging. As such, the effects observed in these studies, which are primarily from Western cultures where age-based stereotypes remain prevalent, may differ from cultures that tend to view the aging process in a more positive manner.

Conclusion

Taken together, this meta-analysis shows that exposure to negative age-based stereotypes negatively impacts both episodic and working memory performance in older adults. Additionally, the two types of memory appear to be differentially impacted by blatant and subtle manipulations of ST. The ST effect on episodic memory might also be influenced by differences in testing procedures, such as the type of memory tests employed or the timing of the memory tests. However, overall ST effects on memory were not moderated by age or level of education. Negative stereotypes associated with aging are common in Western societies. Considering the detrimental effects of these negative stereotypes, future cognitive aging research must take this into account by ensuring their experimental design minimizes exposure to popular stereotypes about aging and memory, especially when examining the performance of older adults in an experimental setting.

Acknowledgments

We would like to thank Dr. Leslie Atkinson and Dr. Sarah Barber for their constructive comments and feedback on this meta-analysis and manuscript. Author Note: All authors contributed equally to this manuscript and are listed in alphabetical order.

Funding

This paper was published as part of a supplement sponsored and funded by AARP. The statements and opinions expressed herein by the authors are for information, debate, and discussion, and do not necessarily represent official policies of AARP.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Abrams, D., Eller, A., & Bryant, J. (2006). An age apart: The effects of intergenerational contact and stereotype threat on performance and intergroup bias. *Psychology and Aging*, **21**(4), 691–702. doi:10.1037/0882-7974.21.4.691

*Andreoletti, C., & Lachman, M. E. (2004). Susceptibility and resilience to memory aging stereotypes: Education matters more than age. *Experimental Aging Research*, **30**(2), 129–148. doi:10.1080/03610730490274167

*Barber, S. J., & Mather, M. (2013a). Stereotype threat can reduce older adults' memory errors. *Quarterly Journal of Experimental Psychology (2006)*, **66**(10), 1888–1895. doi:10.1080/17470218.2013.840656

*Barber, S. J., & Mather, M. (2013b). Stereotype threat can both enhance and impair older adults' memory. *Psychological Science*, **24**(12), 2522–2529. doi:10.1177/0956797613497023

Barber, S. J., & Mather, M. (2014). Stereotype threat in older adults: When and why does it occur, and who is most affected? In P. Verhaeghen & C. Hertzog (Eds.), *The Oxford Handbook of Emotion, Social Cognition, and Everyday Problem Solving During Adulthood*. New York: Oxford University Press.

*Barber, S. J., Mather, M., & Gatz, M. (2015). How Stereotype threat affects healthy older adults' performance on clinical assessments of cognitive decline: The key role of regulatory fit. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, **70**(6), 891–900. doi:10.1093/geronb/gbv009

*Barber, S. J., Niblett, S., & Seliger, J. (n.d.). *Testing the efficacy of value affirmations in eliminating age-based stereotype threat*. Unpublished raw data.

*Bensadon, B. A. (2010). *Memory self-efficacy and stereotype effects in aging (Unpublished doctoral dissertation)*. University of Florida, Florida.

*Bouazzaoui, B., Follenfant, A., Ric, F., Fay, S., Croizet, J. C., Atzeni, T., & Taconnat, L. (2016). Ageing-related stereotypes in memory: When the beliefs come true. *Memory (Hove, England)*, **24**(5), 659–668. doi:10.1080/09658211.2015.1040802

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society, London, Series B*, **302**, 341–359. doi: 10.1098/rstb.1983.0059

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities, mechanisms and performances* (pp. 409–422). North-Holland, The Netherlands: Elsevier.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 474–479. doi: 10.1037/0278-7393.13.3.474

Cuddy, A. J. C., Norton, M. I., & Fiske, S. T. (2005). This old stereotype: The pervasiveness and persistence of the elderly stereotype. *Journal of Social Issues*, **61**, 265–283. doi: 10.1111/j.1540-4560.2005.00405.x

Danckert, S. L., & Craik, F. I. (2013). Does aging affect recall more than recognition memory? *Psychology and Aging*, **28**(4), 902–909. doi:10.1037/a0033263

Dark-Freudeman, A., West, R. L., & Viverito, K. M. (2006). Future selves and aging: Older adults' memory fears. *Educational Gerontology*, **32**, 85–109. doi: 10.1080/03601270500388125

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. doi: 10.1002/ejsp.249

*Eich, T. S., Murayama, K., Castel, A. D., & Knowton, 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. doi: 10.1521/soco.2014.32.6.559

*Fernández-Ballesteros, R., Bustillos, A., & Huici, C. (2015). Positive perception of aging and performance in a memory task: compensating for stereotype threat? *Experimental Aging Research*, **41**(4), 410–425. doi:10.1080/0361073X.2015.1053757

Haaland, K. Y., Price, L., & Larue, A. (2003). What does the WMS-III tell us about memory changes with normal aging? *Journal of the International Neuropsychological Society: JINS*, **9**(1), 89–96. doi: 10.1017/S1355617703910101

*Haslam, C., Morton, T. A., Haslam, S. A., Varnes, L., Graham, R., & Gamaz, L. (2012). "When the age is in, the wit is out": age-related self-categorization and deficit expectations reduce performance on clinical tests used in dementia assessment. *Psychology and Aging*, **27**(3), 778–784. doi:10.1037/a0027754

Hedden, T., & Gabrieli, J. D. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews. Neuroscience*, **5**(2), 87–96. doi:10.1038/nrn1323

*Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype threat on age differences in memory performance. *Journal of Gerontology: Psychological Sciences*, **55B**, 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*, **64**(4), 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**(3), 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**(2), 153–177. doi: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**(3), 495–505. doi:10.1037/0882-7974.19.3.495

Horton, S., Baker, J., Pearce, W., & Deakin, J. M. (2008). On the malleability of performance: Implications for seniors. *Journal of Applied Gerontology*, **27**, 446–465. doi:10.1177/0733464808315291

*Horton, S., Baker, J., Pearce, W., & Deakin, J. M. (2010). Immunity to popular sterotypes of aging? Seniors and stereotype threat. *Educational Gerontology*, **36**, 353–371. doi:10.1080/03601270903323976

Hummert, M. L. (2011). Age stereotypes and aging. In K. W. Schaie & S. L. Willis (Eds.), *Handbook of the psychology of aging*, 7th ed (pp. 249–262). San Diego, CA: Elsevier Academic Press.

*Kang, S. K., & Chasteen, A. L. (2009). The moderating role of age-group identification and perceived threat on stereotype threat among older adults. *International Journal of Aging & Human Development*, **69**(3), 201–220. doi:10.2190/AG.69.3.c

*Kominsky, T. K. (2003). *The effects of stereotype threat on older adults' prospective memory performance (Unpublished doctoral dissertation)*. University of Sioux Falls, Sioux Falls, SD.

*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**(2), 103–109. doi:10.1016/j.jarmac.2015.02.001

Laditka, J. N., Laditka, S. B., Liu, R., Price, A. E., Wu, B., Friedman, D. B., … & Logsdon, R. G. (2011). Older adults' concerns about cognitive health: commonalities and differences among six United States ethnic groups. *Ageing and Society*, **31**, 1202–1228. doi:10.1017/S0144686X10001273

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**(1), 180–193. doi:10.1037/a0038586

Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology*, **71**(6), 1092–1107.

Levy, B. R., & Leifheit-Limson, E. (2009). The stereotype-matching effect: Greater influence on functioning when age stereotypes correspond to outcomes. *Psychology and Aging*, **24**(1), 230–233. doi:10.1037/a0014563

Mather, M. (2010). Aging and cognition. *Wiley interdisciplinary reviews. Cognitive science*, **1**(3), 346–362. doi:10.1002/wcs.64

*Mazerolle, M., Régner, I., Morisset, P., Rigalleau, F., & Huguet, P. (2012). Stereotype threat strengthens automatic recall and undermines controlled processes in older adults. *Psychological Science*, **23**(7), 723–727. doi:10.1177/0956797612437607

*Mazerolle, M., Régner, I., Rigalleau, F., & Huguet, P. (2015). Stereotype threat alters the subjective experience of memory. *Experimental Psychology*, **62**(6), 395–402. doi:10.1027/1618-3169/a000303

Nguyen, H. H., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *The Journal of Applied Psychology*, **93**(6), 1314–1334. doi:10.1037/a0012702

Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, **16**(5), 292–305. doi:10.1016/j.tics.2012.04.005

O'Brien, L. T., & Hummert, M. L. (2006). Memory performance of late middle-aged adults: Contrasting self-stereotyping and stereotype threat accounts of assimilation to age stereotypes. *Social Cognition*, **24**, 338–358. doi:10.1521/soco.2006.24.3.338

Park, D. C. (2000). The basic mechanism accounting for age-related decline in cognitive function. In D.C. Park & N. Schwarz (Eds.), *Cognitive aging: A primer* (pp. 3–21). Philadelphia: Psychology Press.

Perdue, C. W., & Gurtman, M. B. (1990). Evidence for the automaticity of ageism. *Journal of Experimental Social Psychology*, **26**, 199–216. doi: 10.1016/0022-1031(90)90035-K

*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**(2), 225–234. doi:10.1093/geronb/gbt093

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**(4), 697–706. doi:10.1037/0882-7974.16.4.697

Rosenthal, R. (1995). Writing meta-analytic Reviews. *Psychological Bulletin*, **118**(2), 183–192. doi:10.1037/0033-2909.118.2.183

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

Sedgwick, P. (2013). Meta-analyses: How to read a funnel plot. *British Medical Journal*, **345**, 1–2. doi: 10.1136/bmj.f1342

Seefeldt, C., Jantz, R. K., Galper, A., & Serock, K. (1977). Using pictures to explore children's attitudes toward the elderly. *Gerontologist*, **17**(6), 506–512. doi:10.1093/geront/17.6.506

Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *The American Psychologist*, **52**(6), 613–629. doi:10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, **69**(5), 797–811. doi: 10.1037/0022-3514.69.5.797

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**(2), 169–181. doi:10.1080/03610730252800184

*Swift, H. J., Abrams, D., & Marques, S. (2013). Threat or boost? Social comparison affects older people's performance differently depending on task domain. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, **68**(1), 23–30. doi:10.1093/geronb/gbs044

Vasil, L., & Wass, H. (1993). Portrayal of the elderly in the media: A literature review and implications for educational gerontologists. *Educational Gerontology: An International Quarterly*, **19**, 71–85. doi: 10.1080/0360127930190107

*Wong, J. (2014). *Ageist stereotype threat effects on memory and metacognition in older adults (Unpublished doctoral dissertation or master's thesis)*. The University of Chicago, Chicago, Illinois.

Wong, J. T., & Gallo, D. A. (2015). Stereotype threat reduces false recognition when older adults are forewarned. *Memory (Hove, England)*, **24**(5), 1–9. doi:10.1080/09658211.2015.1036885

World Health Organization (2015, April). Ageing and Health. Retrieved from: http://www.who.int/mediacentre/factsheets/fs404/en/