Published in final edited form as:
Exp Aging Res. 2009 ; 35(2): 153–177. doi:10.1080/03610730802716413.

Moderators of and Mechanisms underlying Stereotype Threat Effects on Older Adults' Memory Performance

Thomas M. Hess, **Joey T. Hinson**, and **Elizabeth A. Hodges**
Department of Psychology North Carolina State University Raleigh, NC 27695-7650

Abstract

Recent research has suggested that negative stereotypes about aging may have a detrimental influence on older adults' memory performance. This study sought to determine whether stereotype-based influences were moderated by age, education, and concerns about being stigmatized. Possible mechanisms underlying these influences on memory performance were also explored. The memory performance of adults aged 60 to 70 years and 71 to 82 years was examined under conditions designed to induce or eliminate stereotype threat. Threat was found to have a greater impact on performance in the young-old than in the old-old group, whereas the opposite was observed for the effects of stigma consciousness. In both cases, the effects were strongest for those with higher levels of education. Further analyses found little evidence in support of the mediating roles of affective responses or working memory. The only evidence of mediation was found with respect to recall predictions, suggesting a motivational basis of threat effects on performance. These findings highlight the specificity of stereotype threat effects in later adults as well as possible mechanisms underlying such effects.

Keywords

Aging; stereotype threat; memory; affect; stigma consciousness

---

Aging-related stereotypes have always been of interest in the field of gerontology. Relatively recently, however, attention has shifted from what might be characterized as externally driven forces associated with stereotypes (e.g., the effects of being treated in a stereotype-consistent fashion) to more internally based factors (e.g., the impact of self knowledge of aging-related stereotypes) (Hertzog & Hultsch, 2000; Hess, 2006). In the study of memory, a traditional emphasis has been on how aging stereotypes might affect performance through their impact on one's beliefs (e.g., control, self-efficacy) and the subsequent influence of these beliefs on memory-related behaviors, and vice versa (e.g., Lachman, 2000). More recently, the emphasis has shifted slightly to understanding what might be thought of as more direct influences of stereotypes without specific reference to intervening alterations in beliefs.

The direct influence of stereotypes on memory has been studied in two ways. The first involves implicit (i.e., without the individual's awareness) priming of both positive and negative aging stereotypes. Although results from this body of research have been somewhat mixed (e.g., Stein, Blanchard-Fields, & Hertzog, 2002), there is evidence that older adults remember less and age differences in performance are exacerbated when negative rather than neutral or positive aging stereotypes are activated (Hess, Hinson, & Statham, 2004; Levy, 1996). This

Correspondence should be addressed to Thomas M. Hess, Department of Psychology, North Carolina State University, Raleigh, NC 27695-7801. Office phone: (919) 515-1729 Fax: (919) 515-1716 Electronic mail: thomas_hess@ncsu.edu.

suggests that older adults' memory performance may be unknowingly affected by subtle cues in the environment that are associated with and activate aging-related schemas.

A second line of research involves more explicit activation of stereotypes, in which individuals are aware of relevant cues. Most of this research has been conducted using the stereotype threat framework developed by Steele (1997; Steele, Spencer, & Aronson, 2002) as a starting point. Stereotype threat occurs in situations where the individual is conscious of the fact that one's behavior may confirm negative views of a group to which one belongs. This can lead to the subjective experience of threat which, in turn, may negatively affect performance through mechanisms such as increased anxiety or reductions in working memory capacity. Research in this area has suggested that threat may be operative in old age (Chasteen, Bhattacharyya, Horhota, Tam, & Hasher, 2005; Desrichard & Köpetz, 2005; Hess, Auman, Colcombe, & Rahhal, 2003; Hess & Hinson, 2006; Hess et al., 2004; Rahhal, Hasher, & Colcombe, 2001), with memory performance being negatively affected when cues in the experimental context highlight the stereotype diagnosticity of the test situation. This research has also highlighted the fact that the impact of threat—as reflected in responses to experimental manipulations designed to induce it—is not universal. The effects appear most evident in those who (a) value their memory (Hess et al., 2003), (b) are at the beginning of what is considered old age (e.g., 60 to 70; Hess & Hinson, 2006), and (c) subjectively report feelings of threat (Chasteen et al., 2005). These findings are all consistent with stereotype threat predictions in that the negative effects of threat are most evident in those whose group membership is most salient, who identify with the stereotyped domain, and who consciously experience threat. The first two effects are of particular interest in that they may reflect the ease with which self stereotypes relating to aging are activated. For example, those who are just entering old age may have relatively heightened sensitivity to their new group membership status and associated cues in the environment (e.g., the Association for the Advancement of Retired Persons begins sending membership forms to people when they turn 50). This may, in turn, make individuals aware of their age-related status and make relevant cues (e.g., listings for senior citizen movie ticket prices) more salient. This, in turn, may make them more susceptible to threat due to the lower thresholds of activation for aging schemas.

An important question in this research on stereotype threat and aging concerns the mechanisms underlying the effects. Steele et al. (2002) note a number of different possibilities and also suggest that the operative factor may vary across situations and stereotypes. In aging-related work, Hess et al. (2003) obtained evidence suggesting that threat effects on word list recall were partially mediated by strategy use. This finding is consistent with some other research on stereotype threat (e.g., Quinn & Spencer, 2001) and appears to fit with a hypothesis that the negative effects of stereotype threat are in part due to reductions in working memory that affect the efficiency of processing (Croizet et al., 2004; Schmader & Johns, 2003). Note, however, that later work by Hess and Hinson (2006) failed to replicate this result.

The finding by Chasteen et al. (2005) that subjective experiences of threat negatively influence performance is consistent with another possible mechanism having to do with the disruptive effects associated with anxiety. (It should be noted, however, that this study failed to find systematic effects due to an experimental manipulation of threat.) In particular, the worry component of anxiety has been implicated as mediating threat effects on performance, perhaps through the negative impact of intrusive thoughts on working memory capacity (e.g., Cadinu, Maass, Rosabianca, & Kiesner, 2005). There is also research on aging that has identified a link between self-reported affect and free recall performance (Hill, van Boxtel, Ponds, Houx, & Jolles, 2005), suggesting that emotional responses to the test situation may be important influences on older adults' memory performance.

The goal of the present study was to extend our understanding of stereotype-based influences on memory performance in older adults. In particular, we were interested in examining moderators of stereotype threat effects associated with aging as well as the mechanisms underlying such effects. To do so, we tested older adults aged 60 and older using a free recall task used in prior research (e.g., Hess et al., 2003). Threat was manipulated through a relatively simple procedure in which participants were told either that the test was being used to examine aging effects on memory (threat condition) or that it was constructed so that it controlled for biases that might be associated with age (non-threat). This manipulation was designed to highlight the diagnosticity of the memory task with respect to stereotyped age-related abilities. In addition, to further enhance the salience of the manipulation, we asked participants in the threat condition to write down their age immediately after reading the task instructions.

As noted before, a primary goal of the present study was to identify moderators of the anticipated negative impact of threat on memory performance. Three primary factors were investigated. First, we predicted that the threat manipulation would have a stronger impact on the performance of those in the younger half of the sample than in the older half. Such a finding would replicate the findings of an initial examination of such effects by Hess and Hinson (2005) using a continuous age range, but the present study provided a more focused test of this hypothesis. Second, and relatedly, we were also interested in whether people who believed that they or members of their group are being stigmatized by others were more susceptible to threat. Brown and Pinel (2003) examined math performance in women and found that individuals high in stigma consciousness performed worse under conditions of threat than did those who were low in stigma consciousness. Similar to our hypotheses regarding age, high stigma consciousness may increase awareness of threat-related cues in the environment due to decreased threshold of activation for aging self stereotypes. Finally, we also investigated the moderating impact of education, which we assumed might be a proxy for the importance attached to one's cognitive abilities. If true, threat effects on memory should be greater for those with higher levels of education.

We were also interested in examining three types of mechanisms presumed to underlie threat-based effects. The first related to cognitive capacity. Steele and colleagues had hypothesized that threat may impact performance in part by reducing cognitive resources as individuals focus attention on threat and ways of managing it, thereby reducing available capacity for task performance. Schmader and Johns (2003) found specific support for this by demonstrating reductions in working memory performance following threat. A similar mechanism may underlie anxiety-based effects on performance effects when task-irrelevant thoughts (e.g., worry) occupy working memory and interfere with performance (e.g., Cadinu et al., 2005). In the present case, we tested the hypothesis that threat influences performance through cognitive resources by examining working memory capacity using a computation span task. We chose this specific task not only for its ability to tap into working memory, but also because we thought that it might not resemble participants' notion of a typical memory task. This was important because if the task was perceived as one testing memory, then any threat-related effects that may be observed could be attributed to its being perceived as diagnostic of memory ability rather than to capacity reductions. To further influence participants' perceptions of this task, it was never explicitly labeled as one involving memory and instead was referred to as a test of quantitative skills. If threat does affect capacity, then performance on this task should be negatively affected. In addition, variation in performance on this task should account for threat-related variability on our primary memory task. We also examined strategy use—as reflected in semantic clustering during recall—as a related, but indirect expression of working memory.

The second mechanism of interest involved affective responses. Threat has been hypothesized to be associated with increased negative affect, which in turn may negatively influence

NIH-PA Author Manuscript

NIH-PA Author Manuscript

performance through associated negative thoughts or increased levels of arousal (Smith, 2004). We examined this hypothesis by measuring state anxiety during the test session as well as the extent to which individuals were currently experiencing positive or negative affect. The use of these measures was designed to provide a more broad-based approach to assessing affect than has been used in most studies of stereotype, where anxiety as been the primary variable of interest. It was expected that threat would be associated with increases in negative affect.

We also examined arousal using skin conductance responses (SCR) as an index of autonomic nervous system functioning. SCR is associated with threat and exposure to negative stimuli (e.g., Auman, Bosworth, & Hess, 2005; Bradley, Moulder, & Lang, 2005), and thus was expected to be higher in the threat than in the non-threat condition. Previous research involving implicit activation of aging stereotypes (Levy, Hausdorff, Hencke, & Wei, 2000) has shown larger increases in SCR for negative than for positive stereotypes. The present study extends this to a threat-related situation involving activation of aging self-stereotypes through more explicit cues. The hypothesized relationship between physiological markers of arousal and memory also builds on recent reports demonstrating a relationship between stress-related responses and performance in older adults (Neupert, Miller, & Lachman, 2006; Wright, Kunz-Ebrecht, Iliffe, Foese, & Steptoe, 2005). These studies have found that older adults exhibit stronger responses to cognitive testing situations than do younger adults, and that variations in stress responses are related to performance. In addition, Wright et al. observed that these relationships were stronger for memory tasks than for other cognitive tasks. Such relationships might be thought of in terms of aging-specific threat-related responses to testing situations, particularly when the task involves a strongly stereotyped skill (i.e., memory). Of interest in the present case is the possibility that situational factors (e.g., task instructions) and individual characteristics (e.g., stigma consciousness) might be associated with the strength of such responses and account for individual variability in performance. If threat does negatively influence affective responses, and these factors in turn interfere with performance, it would be expected that variation in subjective reports of affect as well as psychophysiological responses would be predictive of threat effects on memory performance.

Finally, we examined a motivational mechanism in the form of performance expectations. Recent research has suggested that threat effects may be more based in motivational factors than in working memory or affective responses (Desrichard & Köpetz, 2005; Jamieson & Harkins, 2007; Seibt & Förster, 2004). To provide a test of this hypothesis, we had participants predict how well they would do on the memory task, and then examined the role of such predictions in mediated threat effects.

Method

Participants

The study sample included 103 adults who responded to recruitment advertisements in the local newspaper, each of who received $20 for their participation. Participants fell into two broad age groups: (a) young-old—60 to 70 years of age (27 men, 26 women; *M* age = 64.2) and (b) old-old—71 to 82 years of age (25 men, 25 women; *M* age = 75.4). None of these individuals had participated in prior studies in our lab.

Materials

Computation span task—This task was modeled after that used by Salthouse and Babcock (1991). Participants were presented with sets of simple arithmetic equations consisting of the addition or subtraction of two numbers. Three sets of three, four, five, six, and seven problems were created. For each set, an answer sheet was created containing three alternative answers for each problem in the set.

Free recall task—The memory task was identical to that used in Hess et al. (2003). The list of 30 to-be-remembered words consisted of five moderate to strong exemplars from each of six different semantic categories. The words were randomly ordered in two vertical columns of 15 words each on a sheet of standard white paper.

Measures of affective responses—State anxiety was assessed using the10-item short form of the state anxiety subscale of the State-Trait Anxiety Inventory (STAI-S; Speilberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). The Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) was used to evaluate participants' current affective states.

Stigma consciousness—We adapted the Stigma Consciousness Questionnaire (SCQ) developed by Pinel (1999) to examine concerns regarding aging stereotypes. The SCQ consists of 10 items (e.g., “Stereotypes about older adults have not affected me personally.”) responded to using a 7-point Likert scale (0 = strongly disagree; 6 = strongly agree). Internal consistency for this scale in the present study was reasonable (Cronbach's alpha = .71). Lower scores on the SCQ represented greater stigma consciousness.

Other measures—A standard background questionnaire was used to collect basic demographic data, and health status was assessed using the SF-36 Health Survey (Ware, 1993). Two additional ability tests were also administered. Processing speed was assessed using Salthouse and Coon's (1994) letter and pattern comparison tasks, and verbal ability was measured with the Vocabulary Test II (parts 1 and 2) from the Kit of Factor-Referenced Tests (Ekstrom, French, Harman, & Derman, 1976).

Equipment

The ProComp Infiniti Model SA 7500 multi-modality encoder (Thought Technology, Ltd., 2004) was used to collect SCR data. This unit was connected to a microcomputer using the BioGraph Infiniti Software (Thought Technology, Ltd., 2004) interface, which was used to record and store the psychophysiological data. Responses were collected using two Ag/Ag/Cl electrodes attached to the index and middle fingers of the participant's left hand.

Procedure

Participants were contacted by telephone in order to schedule a time for their participation. A few days prior to testing, they received a packet containing the background questionnaire, SF-36 survey, the SCQ, and several attitudes questionnaires unrelated to the present study. Participants returned the completed questionnaires when they arrived at the lab for testing.

Upon their arrival at the lab, participants read and signed an informed consent form. The SCR recording devices were then attached to the ring and index fingers of their nondominant hand. The recording unit was then turned on, and data was recorded throughout the test session. Participants were instructed to place their nondominant hand palm down on an outline of a hand on the table in front of them, and were told to try to keep movement of this hand to a minimum. As each new phase of the study was begun, the experimenter pressed the space bar on the computer, which in turn inserted a marker in the electronic record of the test session.

Participants were randomly assigned to the threat and non-threat conditions. In the young-old group, there were 14 women and 13 men in the threat condition and 13 women and 12 men in the non-threat condition. In the old-old group, there were 13 women and 12 men in the threat condition and 13 women and 13 men in the non-threat condition. At the beginning of the session, all participants were informed that we were interested in examining those factors that predict performance on tests of memory ability. In the threat condition, participants were further informed that a primary purpose of the study was to try to explain why younger and older adults perform so differently on memory tests. They then were asked to write down their age on a sheet of paper. In the non-threat condition, participants were informed that, in order to eliminate any contaminating factors, the test they were going to take was free of age-related biases such that adults of various ages performed similarly on the test. They were not asked to indicate their age at this point.

The PANAS was administered next, with participants being instructed to rate the extent to which each of the 20 items reflected how they felt at the present moment. This was followed by the computation span task, which was described as a measure of quantitative skills in order to minimize the potential bias introduced by the using the word "memory". In this task, participants were presented with a series of simple arithmetic problems (e.g., "7 plus 2"), which were read to them by the experimenter. They then circled the correct answer on the sheet in front of them. Instructions were also given to remember the second number in each problem (e.g., "2" in the aforementioned problem). After all the problems at a specific set size were presented, participants turned the answer sheet over and wrote down the second numbers from each of the problems in that set in they order in which they were presented. Participants were given practice with a set size of two, with testing in the primary task proceeding through the three sets at each level from smallest to largest.

State anxiety was assessed next using the STAI-S. The memory test was then described in detail, and participants were asked to predict how many words they would remember. To provide a common anchor for these judgments, they were informed that the average person recalls about 15 words. This reflects the approximate lower bound of performance observed in previous studies (e.g., Hess & Hinson, 2006), and thus was not an unreasonable standard. [1] It was assumed that the use of the anchor would reduce variability across participants, thereby creating more statistical power to identify threat effects. The memory test was presented next. Participants were allowed to study the list of words for 2 min, after which the list was taken away. They were then given a blank piece of paper and allowed a minimum of 3 min to write down as many words as they could remember. If they were still writing at the end of this period, an additional 30 s was given.

Following the memory task, participants were debriefed. The vocabulary test and letter and pattern comparison tasks were then given, after which participants were paid and thanked for their participation.

Results

Participant characteristics are presented in Table 1 by age group and experimental condition. Prior to testing our hypotheses, we conducted a series of 2 × 2 (Age Group × Instructional Condition) analyses of variance (ANOVAs) to see if there were any between-group differences on relevant background characteristics (e.g., education, ability, health) that could affect the results. Notable differences were only found for two variables. Education varied significantly as a function of age, $F(1,99) = 4.54$, $p = .04$, $\eta_p^2 = .04$, and its interaction with condition, $F(1,99) = 4.56$, $p = .04$, $\eta_p^2 = .04$. These effects were due to young-old participants in the threat condition having significantly lower levels of education ($M = 14.7$ years) than those in the other three Age × Condition groups (Ms = 16.0 to 16.5 years). Processing speed was also related to age, with those in the young-old group exhibiting higher levels of performance than those in the old-old group, $F(1,99) = 9.62$, $p = .01$, $\eta_p^2 = .09$.

[1]In spite of these data, it is possible that informing participants about these norms may have negatively affected performance by increasing performance anxiety. This should be reflected in a recall predictions being negatively associated with our measure of negative affect. Although this correlation was negative, it was not significant ($r = -.11$, $p > .26$), suggesting that this was not a strong concern.

Memory Performance

In our first set of analyses, we were interested in examining the influence of age, threat condition, stigma consciousness (as indexed by SCQ scores), and education on our primary outcome measure: free recall. The first two variables were treated as categorical variables whereas stigma consciousness and education were treated as continuous variables centered through standardization. Given the previously reported group differences, speed was considered as a possible covariate, but its inclusion did not alter any results. Thus, it was not considered further as a covariate.

Free recall—The proportion of words correctly reported in the free recall task was of primary interest, with one significant effect relating to our threat manipulation emerging from the analysis of these data: a significant Age × Condition × Education interaction, $F(1,87) = 7.76$, $p = .01$, $\eta_p^2 = .08$. The Age × Condition interaction approached significance, $F(1,87) = 3.86$, $p = .053$, $\eta_p^2 = .04$. Analyses conducted within age groups revealed that these effects were based in the young-old group, where the Condition × Education interaction was significant, $F(1,44) = 9.43$, $p = .01$, $\eta_p^2 = .18$. We were somewhat concerned that this effect was driven by an individual in the threat condition with 12 years of education and whose recall performance was .13 higher than anyone else's. Exclusion of this outlier weakened the effect somewhat, but the interaction was still significant $F(1,43) = 7.69$, $p = .01$, $\eta_p^2 = .15$ (see Figure 1). Further examination of performance in the young-old group discovered that the interaction was due to recall decreasing with education in the threat condition ($\beta = -.55$, $p = .01$); this relationship was in the opposite direction, but was not significant in the non-threat group ($\beta = .31$, $p = .16$). Thus, the anticipated threat effect on memory was primarily evident in young-old participants with higher levels of education.[2]

The only other significant effect obtained in the main analysis was the Age × Stigma × Education interaction, $F(1,87) = 5.07$, $p = .03$, $\eta_p^2 = .06$. In contrast to the previous effects, this interaction was based in the old-old group, where the Stigma × Education interaction was significant, $F(1,43) = 7.76$, $p = .01$, $\eta_p^2 = .15$. To examine this interaction more closely, participants in the old-old group were divided into high and low stigma consciousness groups (based on a median split of SCQ scores). Education was negatively related to recall performance in the high stigma consciousness group, $\beta = -.37$, $p = .05$, but unrelated to performance in the low stigma consciousness group, $\beta = .11$, $p = .59$ (see Figure 2).

Interestingly, stigma consciousness was not related to performance in the young-old group, although the Condition × Stigma interaction in this group did approach significance, $F(1,44) = 3.02$, $p = .09$, $\eta_p^2 = .06$. The power to detect this effect [.40], however, was rather low in the present analysis. Given this fact plus our original interest in this issue, we decided to decompose this interaction by examining the impact of stigma conscious within conditions. Stigma consciousness was unrelated ($p = .55$) to recall in the non-threat condition. In contrast, a marginally significant negative relationship between recall and SCQ scores was observed in the threat condition, $\beta = -.35$, $p = .06$. This finding is consistent with the expectation that stigma consciousness makes one more susceptible to stereotype threat.

In sum, the effects of stereotypes were evident through the impact of both experimental manipulations of threat and self-reported concerns about aging-related stigma. Interestingly, the threat manipulation only affected performance in young-old participants, whereas the

[2]It is important to note that the ranges of education across the four Age X Threat Condition groups were comparable, with a lower bound of 12 years in each of the four groups and a higher bound ranging form 19 – 21. When those individuals with education levels greater than 19 years ($n = 4$) were excluded from the analyses to equate the ranges of education within groups, the three-way interaction was unaffected.

impact of stigma consciousness was stronger in the old-old participants. The impact of both variables, however, was observed primarily at higher levels of education.

Clustering—The other dependent variable obtained from this task was output organization by semantic category during recall, as measured by the adjusted ratio of clustering (ARC; Roenker, Thompson, & Brown, 1971). The only effect approaching significance was the Age × Condition interaction, $F(1,87)= 2.99$, $p = .09$, $\eta_p^2 = .03$. Tests within age groups revealed significantly lower ARC scores in the threat condition than in the non-threat condition in the young-old group (*M*s = .54 vs. 74), $F(1,44) = 4.56$, $p = .04$, but no difference between these two conditions in the old-old group (adjusted *M*s = .68 vs. 66), $F < 1$.

Recall predictions—Expectations about one's own performance were examined in the same manner as recall. The only significant effects to emerge from this analysis were the main effect of education, $F(1,87) = 16.07$, $p < .001$, $\eta_p^2 = .16$, and the Condition × Education interaction, $F(1,87) = 8.41$, $p = .01$, $\eta_p^2 = .09$. The interaction was decomposed by examining the impact of instructional condition at education levels 1 *SD* above and below the sample mean. A significant condition effect, $F(1,87) = 4.79$, $p = .03$, was observed at the higher level of education, with predicted recall lower in the threat than in the non-threat condition (*Ms* = 10.8 vs. 11.4). In contrast, this effect only approached significance at the lower level of education, $F(1,87) = 3.88$, $p = .06$.

Given that education only moderated condition effects in the young-old group when memory performance was examined, we were curious as to whether there was evidence of this same effect in predicted recall. Although the three-way interaction was not significant, we examined predictions within age groups. The results mirrored those with recall, with a significant Education × Condition interaction being observed in the young-old group, $F(1,44) = 6.05$, $p = .02$, but not in the old-old group, $F(1,43) = 2.43$, $p = .13$. Examination once more of the impact of instructional condition at high and low levels of education within the young-old group revealed that predicted scores were significantly greater in the control than in the threat condition at high levels of education, $F(1,44) = 5.25$, $p = .03$, but not at low levels of education ($p = .23$).

Hypothesized Mediators

We next performed the same analyses on those measures reflecting constructs representing potential mediators of threat effects.

Working memory—We first examined performance on the computation span task to test the prediction that threat results in reductions in working memory. Performance was assessed by the total number of items from the equations that were correctly recalled across all set sizes. (The observed effects were similar using alternative ways of scoring, such as highest level at which participants correctly recalled all items from two of the three sequences of equations.) Education was the only factor found to influence performance, $F(1,86) = 15.95$, $p < .001$, $\eta_p^2 = .16$, with span scores increasing with education ($\beta = .29$). No significant effects involving condition or stigma consciousness emerged (*p*s > .23), providing no support for the hypothesis that threat effects operate through working memory.

Self-Reported Affect—Subjective reports reflecting affective responses were available in the form of three measures. Positive and negative affect scores were obtained from the PANAS, whereas state anxiety was indexed by STAI-S scores. Although given at different times during the study (i.e., before and after the working memory task), these measures were significantly intercorrelated (*p*s < .03), and thus arguably tapping the same affective reaction to the test situation. In addition, the fact that state anxiety and negative affect scores were strongly

NIH-PA Author Manuscript

NIH-PA Author Manuscript

correlated with each other (r = .50) and similarly correlated with performance on the computation span task (rs = −.20 and −.22, respectively) suggests that STAI-S scores did not differentially reflect responses to the working memory task. Therefore, to simplify our analyses, we subjected these three measures to factor analysis, where a single factor relating to negative affect was obtained that accounted for 58% of the variance. Factor loadings were .90 for anxiety, .56 for negative affect, and −.40 for positive affect. Factor scores were then calculated and analyzed in the same fashion as the other dependent measures. A significant main effect of stigma consciousness was obtained, $F(1,87) = 10.49$, $p = .002$, $\eta_p^2 = .11$, with negative affect increasing with stigma consciousness. Stigma consciousness also interacted with education, $F(1,87) = 5.71$, $p = .02$, $\eta_p^2 = .06$, with the just-described relationship being significant ($p = .001$) at low levels of education (i.e.,1 *SD* below the sample mean), but not at higher levels ($p = .76$).

The Age × Condition × Education interaction just missed significance, $F(1,87) = 3.85$, $p = .053$, $\eta_p^2 = .04$. As with recall performance, the impact of condition was strongest at the highest levels of education in the young group, where the Condition × Education interaction approached significance ($p = .07$); this same effect was not significant in the old group ($p = .41$). The pattern of performance in the young-old group was similar to that observed for recall, with threat-based effects being stronger at higher levels of education. This was reflected in the positive relation between negative affect and education in the threat condition ($B = .16$) and a negative relation in the control condition ($B = -.41$), although neither effect was significant (ps > .10).

Thus, stigma consciousness was related to self-reports of negative affect, with those expressing higher levels of consciousness also reporting greater negative affect. The induction of threat also had a marginal influence on affective responses, as reflected in higher negative affect. Interestingly, as with memory performance, the effects of threat were primarily evident in those with more education in the young-old group.

Psychophysiological Responses—Two summary measures were obtained for SCR during each of four different phases of the test session: baseline, threat instructions, free recall study, and free recall test. The first measure was the mean level of response, whereas the second was variability in this response, as measured by the standard deviation. Given that each phase of testing lasted a different period of time, we decided to use only the first 60 s of data following the start of each phase (e.g., the first 60 s of data following presentation of the word list to the participant in the study phase of the free recall task). Eight data points were available per second, making a total of 480 observations per test phase. Prior to calculating the *M* and *SD* for each phase, extreme responses that were 3 *SD*s above or below a participant's mean were eliminated from consideration. Due to equipment malfunctions, data for 8 participants were lost. The resulting scores were examined using Age Group × Condition × Stigma Consciousness × Education × Test Phase ANOVAs, with test phase being a repeated measures variable.

We first examined mean SCRs. Initial analyses revealed that the data of two participants—one in each condition—unduly influenced the results of our statistical tests, and thus they were excluded from further consideration. Significant effects were observed for test phase, $F(3,237) = 111.38$, $p < .001$, $\eta_p^2 = .59$, and its interaction with age group, $F(3,237) = 3.65$, $p = .01$, $\eta_p^2 = .04$. These effects reflected the facts that (a) SCR increased from baseline to free recall study, and then remained stable through test, and (b) this effect was stronger for the young-old group than for the old-old group. A significant main effect due to condition was also observed, $F(1,79) = 5.73$, $p = .02$, $\eta_p^2 = .07$, due to mean SCRs being greater in the threat condition (3.12) than in the non-threat condition (2.31). There was some evidence of those in the threat condition exhibiting stronger responses to the instructional manipulation (Table 2), but the interaction between condition and test phase was not significant ($p > .20$).

Examination of SCR variability (Table 2) revealed that variability was higher in the threat condition ($M = .19$) than in the non-threat condition ($M = .13$), $F(1,79) = 4.63$, $p = .03$, $\eta_p^2 = .06$, and increased from baseline to free recall test, $F(3,237) = 18.39$, $p < .001$, $\eta_p^2 = .19$. In addition, a significant Age × Condition × Stigma Consciousness × Test Phase, $F(3,237) = 3.09$, $p = .03$, $\eta_p^2 = .04$, interaction was obtained. Analyses within conditions revealed that this interaction was due to variations within the threat condition. Variability increased over test phases in both age groups (*p*s <.01), but stigma consciousness moderated this variability in the old-old group, with variability increasing with stigma consciousness during the first two test phases.

Mediators of Threat Effects on Performance

In our final set of analyses, we were interested in examining potential mediators of threat effects on performance. Given that the relationship between our instructional condition and the primary outcome measure—free recall—varied as a function of age and education, a simple mediation analysis was not possible. Thus, we focused on the young-old group, where there was evidence of some impact of threat.[3] Given the interaction between instructional condition and education, we decided to test for mediated moderation (Muller, Judd, & Yzerbyt, 2005), which involves determining the mediating process responsible for the moderating impact of education on the threat manipulation. Within the context of our study, there were five potential mediators that we examined: working memory, predicted recall, clustering scores, negative affect, and arousal (i.e., SCR).

According to Muller et al., three conditions must occur to demonstrate mediated moderation. First, there must be evidence of moderation in the primary analyses. The interaction between our treatment (instructional condition) and moderator (education) was significant, thereby satisfying this condition, $F(1,47) = 7.33$, $p = .01$, $\eta_p^2 = .14$. (Education was centered by subtracting the mean for this age group.) Second, a significant effect of the treatment or its interaction with the moderator must be obtained for the hypothesized mediators. This condition was only satisfied for (a) predicted recall, where a significant Education × Condition effect was obtained, $F(1,47) = 7.40$, $p = .01$, $\eta_p^2 = .14$, and (b) clustering (i.e., ARC scores), where a significant main effect of condition was obtained, $F(1,47) = 6.29$, $p = .02$, $\eta_p^2 = .12$. The final condition for mediated moderation is tested by adding the mediator and its interaction with the moderator to the initial model examining recall. (The mediator was centered by subtracting out the group mean for each variable.) In the case of clustering, where a main effect of the treatment was observed, mediation would be supported by a significant Clustering × Education interaction, along with a reduction in the size of the Education × Condition effect. The former interaction was not significant ($p = .56$), providing no support for mediation. For predicted recall, mediation would be reflected in (a) a significant main effect of the mediator, which was obtained, $F(1,45) = 4.47$, $p = .04$, $\eta_p^2 = .09$, and (b) a reduction in the size of the Education × Treatment interaction, which was also observed, $F(1,45) = 3.81$, $p = .06$, $\eta_p^2 = .08$. Using the Goodman modification of the Sobel test (see MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002), we found a significant reduction in the regression coefficient associated with the Education × Condition interaction when predicted recall was controlled, $z = 6.38$, $p < .001$. This finding is consistent with performance expectations mediating between the observed moderated effect of education on condition and memory performance.

[3]The previously described outlier in the threat group was excluded from these analyses to control for any undue influence.

Discussion

The present study was designed to extend recent investigations into the impact of stereotype-based influences on memory performance in older adults. Several notable effects emerged from this study. First, relative to our goal of examining moderators of stereotype-based effects on memory, we found that the impact of our threat manipulation on recall was most evident in our young-old group, which included individuals at the early stage (7th decade of life) of what is thought of as old age in western society. This effect is consistent with age trends observed by Hess and Hinson (2006).[4] The stronger impact of threat in the 60s could reflect the salience of membership in the stereotyped group of “older adult”, which may be particularly high as individuals enter this group. Self-related salience of social categories has been shown to be associated with stronger stereotype-based effects on performance (e.g., Shih, Ambady, Richeson, Fujita, & Gray, 2002). This salience argument could also explain the fact that the threat manipulation on memory performance increased in strength with increasing education. It is reasonable to assume that those who are higher in education value their cognitive skills the most, and thus might be most susceptible to threats to their ability. Identification with the stereotyped group or domain has been shown to bolster threat effects in past research in a variety of domains (e.g., Hess et al., 2003; Keller, 2007; Schmader, 2002). Although we did not assess domain identification in the present study, it could be argued that highly educated young-old adults might be most sensitive to and reactive to threat manipulations due to the perceived self-relevance of both the stereotyped group and domain.

This explanation is bolstered somewhat by our examination of affective responses. Specifically, those who were higher in education tended to show higher levels of self-reported negative affect when subjected to threat. It is further bolstered when recall predictions are examined, which are also lower for young-old, highly educated participants under conditions of threat. Threat-based effects on predictions might be viewed as a self-protective mechanism that is consistent with a self-regulatory view of stereotype threat (Seibt & Förster, 2004—see below). The findings with respect to education also fit in with recent research by Neupert et al. (2006), who found that stress responses to cognitive testing situations were strongest in those older adults with the most education.

We also investigated the effects of stigma consciousness, which Brown and Pinel (2003) found exacerbated the effects of threat on performance. Consistent with this research, we found that threat-based effects in the young-old group were stronger in those with higher levels of stigma consciousness. No such interaction was observed in the old-old group. We did find, however, that stigma consciousness was associated with memory in this group, and that this effect was moderated by education. Thus, in a somewhat similar effect to that observed in the young-old group, concerns about being stereotyped negatively influenced performance, with this impact being greater at higher levels of education. We also found that stigma consciousness was associated with higher levels of negative affect, regardless of age or condition.

The variations in the impact of the two stereotype-related factors across age is interesting, and suggests potential differences in the role of situational versus more trait-like factors. As originally hypothesized, old-old adults may become more immune to situationally based threat factors, perhaps due to decreases in salience of the old-age label or devaluation of stereotyped cognitive skills. General trait-like factors associated with social perceptions, however, might persist and continue to influence performance.

[4]The time between the threat manipulation and memory test was relatively long in the present study, leading to the possibility that the null threat effects in the old-old group may have reflected differential dissipation (e.g., forgetting) of the experimental manipulation across ages. Although this can not be ruled out, it should be noted that Hess and Hinson (2006) observed a similar age trend using a much shorter interval between threat and the memory test, thereby supporting the reliability of this age trend.

The second major goal of the research was to investigate mechanisms underlying stereotype-based effects on memory performance. We specifically examined the hypothesis that threat has its impact on performance through its impact on working memory, but failed to replicate earlier research by Schmader and Johns (2003) that demonstrated such a linkage. The present failure is especially noteworthy given that the working memory assessment occurred sooner after the threat manipulation than did the main memory task, but that only the latter task exhibited effects associated with threat. It may be possible that the effects on working memory are primarily present in situations in which the stereotyped ability is being assessed. Thus, our attempts to minimize the labeling of the computation span task as one involving memory may have actually reduced threat-based effects by focusing on a skill with minimal associations with the negative aging stereotype. Examination of the Schmader and Johns research revealed that, in two experiments, the working memory task was explicitly linked to the negatively stereotyped skill through labeling and/or task content (e.g., the use of an operation span in examining gender-based stereotypes about mathematical ability). In a third experiment, this linkage was eliminated, and evidence for working memory mediating threat was still obtained. It is noteworthy, however, that the effect size associated with threat on working memory in the examination of gender and math was about 40% larger when the operation span task was used than when a working memory task with less obvious relevance to math was employed. This does suggest that the effects may be stronger when the task is explicitly linked to the stereotyped skill, which in turn calls into question the extent to which working memory mediates threat effects on performance. Concurrent assessments of working memory (e.g., divided attention tasks) during primary task performance might prove more beneficial in testing this mediational hypothesis.

We also investigated the impact of stereotype threat on affective and psychophysiological responses that have been associated with stress and threat in past research. Little evidence in support of a systematic relationship between threat and these variables was found, however. For example, SCR and variability therein were associated with the instructional manipulation, but no systematic differences in change over time were observed across conditions, as would be expected. As noted before, there was some relationship observed with self-reported negative affect. Subsequent failure to find evidence for a mediating role of these reports, however, suggests that such responses are independent of the mechanisms underlying threat effects on performance. The failure to find strong linkages between these variables and performance is consistent with much other work on stereotype threat (see Smith, 2004).

Of final interest in the present study were the factors mediating performance across threat-based conditions. Given that instructional condition did not influence performance in all age groups, and that education mediated the impact of condition, analyses of mediated moderation (Muller et al., 2005) were conducted only within the young-old group. Consistent with the primary analyses, there was no evidence for working memory or affective responses serving as mediators of memory performance. In contrast, performance expectations were observed to satisfy the conditions for mediation. Specifically, as with recall, education moderated the impact of instructional conditions, with recall predictions being lower in those with higher levels of education in the threat group. In addition, recall predictions were predictive of performance, and when these relationships were controlled, the interaction between education and condition was reduced to marginal significance. In other words, expectations were related to performance, but these expectations were primarily affected by the threat manipulation in those with high levels of education.

This relationship is consistent with recent findings in the aging literature by Desrichard and Kopetz (2005), and fits in with a growing body of results consistent with a motivational explanation of stereotype threat effects (e.g., Jamieson & Harkins, 2007; Seibt & Förster, 2004). For example, Seibt and Förster presented evidence that threat influences regulatory

NIH-PA Author Manuscript
NIH-PA Author Manuscript

focus, which in turn affects the manner in which individuals approach a task. Within this framework, negative stereotype activation is thought to result in a prevention focus whereas positive stereotype activation induces a promotion focus. A prevention focus associated with stereotype threat can have either positive or negative effects on performance depending upon the match between the task demands and approach to task associated with prevention. Interestingly, prevention focus is associated with a more cautious approach to task in an effort to reduce errors, which results in less inclusive memory searches (Friedman & Förster, 2001; Seibt & Förster, 2004). Lower recall predictions may, in part, be a reflection of such cautiousness, which in turn gets translated into a specific approach to the memory task that results in few mistakes, but low levels of performance. Although the present results appear consistent with a regulatory focus explanation, further research is necessary using experiments specifically designed to test this hypothesis.

In conclusion, the present study extends our understanding of moderators of and mechanisms underlying stereotype threat effects on memory in adulthood. Age, education, and stigma conscious were all found to moderate the impact of threat, and findings that recall expectations mediated the relationship between instructional condition and performance provides some clues regarding the mechanisms behind threat effects. The fact that age, education, and stigma consciousness all moderated performance further highlights the fact that the effect of stereotype threat on older adults' memory performance is not universal across ages or individuals. The general finding that affective responses and working memory were not strongly predictive of performance following threat, but that performance expectations were, is consistent with emerging evidence regarding the importance of motivational mechanisms associated with the activation of positive and negative stereotypes in performance situations.

Acknowledgements

Support for this research was provided by grant AG20153 from the National Institute on Aging.

References

Auman C, Bosworth HB, Hess TM. The Effects of Health-Related Stereotypes on Physiological Responses of Hypertensive Older Adults. Journal of Gerontology: Psychological Sciences 2005;60B:P3–P10.

Bradley MM, Moulder B, Lang PJ. When good things go bad: the reflex physiology of defense. Psychological Science 2005;16:468–473. [PubMed: 15943673]

Brown RP, Pinel EC. Stigma on my mind: Individual differences in the experience of stereotype threat. Journal Experimental Social Psychology 2003;39:626–633.

Cadinu M, Maass A, Rosabianca A, Kiesner J. Why do women underperform under stereotype threat? Evidence for the role of negative thinking. Psychological Science 2005;16:572–578. [PubMed: 16008792]

Chasteen AL, Bhattacharyya S, Horhota M, Tam R, Hasher L. How feelings of stereotype threat influence older adults' memory performance. Experimental Aging Research 2005;31:235–260. [PubMed: 16036721]

Croizet JC, Despres G, Gauzins ME, Huguet P, Leyens JP, Meot A. Stereotype threat undermines intellectual performance by triggering a disruptive mental load. Personality and Social Psychology Bulletin 2004;30:721–731. [PubMed: 15155036]

Desrichard O, Köpetz C. 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 2005;35:537–552.

Ekstrom, RB.; French, JW.; Harman, H.; Derman, D. Kit of factor- referenced cognitive tests. Educational Testing Service; Princeton, NJ: 1976. rev. ed.

Friedman RS, Förster J. The effects of promotion and prevention cues on creativity. Journal of Personality and Social Psychology 2001;81:1001–1013. [PubMed: 11761303]

Hertzog, C.; Hultsch, DF. Metacognition in adulthood and old age. In: Craik, FIM.; Salthouse, TA., editors. The handbook of aging and cognition. Lawrence Erlbaum; Mahwah, NJ: 2000. p. 417-466.

Hess, TM. Attitudes toward aging and their effects on behavior. In: Birren, JE.; Schaie, KW., editors. Handbook of the psychology of aging. Vol. 6th edition. Academic Press; San Diego: 2006. p. 379-406.

Hess TM, Auman C, Colcombe SJ, Rahhal TA. The impact of stereotype threat on age differences in memory performance. Journal of Gerontology: Psychological Sciences 2003;58:P3–P11.

Hess TM, Hinson JT. Age-related variation in the influences of aging stereotypes on memory in adulthood. Psychology and Aging 2006;21:621–625. [PubMed: 16953723]

Hess TM, Hinson JT, Statham JA. Implicit and explicit stereotype activation effects on memory: Do age and awareness moderate the impact of priming? Psychology and Aging 2004;19:495–505. [PubMed: 15382999]

Hill RD, van Boxtel MPJ, Ponds R, Houx PJ, Jolles J. Positive affect and its relationship to free recall memory performance in a sample of older Dutch adults from the Maastricht Aging Study. International Journal of Geriatric Psychiatry 2005;20:429–435. [PubMed: 15852435]

Jamieson JP, Harkins SG. Mere effort and stereotype threat performance effects. Journal of Personality and Social Psychology 2007;93:544–564. [PubMed: 17892331]

Keller J. Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. British Journal of Educational Psychology 2007;77:323–338. [PubMed: 17504550]

Lachman, ME. Promoting a sense of control over memory aging. In: Hill, RD.; Bäckman, L.; Stigsdotter-Neely, A., editors. Cognitive rehabilitation in old age. Oxford University Press; New York: 2000.

Levy B. Improving memory in old age by implicit self-stereotyping. Journal of Personality and Social Psychology 1996;71:1092–1107. [PubMed: 8979380]

Levy BR, Hausdorff JM, Hencke R, Wei JY. Reducing cardiovascular stress with positive self-stereotypes of aging. Journal of Gerontology: Psychological Sciences 2000;55B:P205–P213.

MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychological Methods 2002;7:83–104. [PubMed: 11928892]

Muller D, Judd CM, Yzerbyt VY. When moderation is mediated and mediation is moderated. Journal of Personality and Social Psychology 2005;89:852–863. [PubMed: 16393020]

Neupert SD, Miller LMS, Lachman ME. Physiological reactivity to cognitive stressors: Variations by age and socioeconomic status. The International Journal of Aging and Human Development 62:221–235.in press

Pearman, A.; Andreolletti, C. The role of anxiety about aging in the physiological and memory functioning of younger and older adults; Poster session presented at the annual meeting of the Gerontological Society of America; Orlando, FL. Nov. 2005

Pinel EC. Stigma consciousness: The psychological legacy of social stereotypes. Journal of Personality & Social Psychology 1999;76:114–128. [PubMed: 9972557]

Quinn DM, Spencer SJ. The interference of stereotype threat with women's generation of mathematical problem-solving strategies. Journal of Social Issues 2001;57:55–71.

Rahhal TA, Hasher L, Colcombe SJ. Instructional manipulations and age differences in memory: Now you see them, now you don=t. Psychology and Aging 2001;16:697–706. [PubMed: 11766922]

Roenker DL, Thompson CP, Brown SC. Comparison of measures for the estimation of clustering in free recall. Psychological Bulletin 1971;76:45–48.

Salthouse TA, Babcock RL. Decomposing adult age differences in working memory. Developmental Psychology 1991;27:763–776.

Salthouse TA, Coon VE. Interpretation of differential deficits: The case of aging and mental arithmetic. Journal of Experimental Psychology: Learning, Memory, and Cognition 1994;20:1172–1182.

Schmader T. Gender identification moderates stereotype threat effects on women's math performance. Journal of Experimental Social Psychology 2002;38:194–201.

Schmader T, Johns M. Converging evidence that stereotype threat reduces working memory capacity. Journal of Personality and Social Psychology 2003;85:440–452. [PubMed: 14498781]

Seibt B, Förster J. Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. Journal of Personality and Social Psychology 2004;87:38–56. [PubMed: 15250791]

Shih M, Ambady N, Richeson JA, Fujita K, Gray HM. Stereotype performance boosts: The impact of self-relevance and the manner of stereotype activation. Journal of Personality and Social Psychology 2002;83:638–647. [PubMed: 12219859]

Smith JL. Understanding the process of stereotype threat: A review of mediational variables and new performance goal directions. Educational Psychology Review 2004;16:177–206.

Spielberger, CD.; Gorsuch, RL.; Lushene, R.; Vagg, PR.; Jacobs, GA. Manual for the State-Trait Anxiety Inventory (Form Y). Mind Garden; Palo Alto, CA: 1983.

Steele CM. A threat in the air: How stereotypes shape intellectual identity and performance. American Psychologist 1997;52:613–629. [PubMed: 9174398]

Steele, CM.; Spencer, SJ.; Aronson, J. Contending with group image: The psychology of stereotype and social identity threat. In: Zanna, MP., editor. Advances in experimental social psychology. Vol. 34. Academic Press; San Diego: 2002. p. 379-440.

Stein R, Blanchard-Fields F, Hertzog C. The effects of age-stereotype priming on memory performance in older adults. Experimental Aging Research 2002;28:169–181. [PubMed: 11928527]

Thought Technology, Ltd.. 2004.

Ware, JE, Jr.. SF-36 Health Survey. The Health Institute, New England Medical Center; Boston: 1993.

Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality & Social Psychology 1988;54:1063–1070. [PubMed: 3397865]

Wright CE, Kunz-Ebrecht SR, Iliffe S, Foese O, Steptoe A. Physiological correlates of cognitive functioning in an elderly population. Psychoneuroendocrinology 2005;30:826–838. [PubMed: 15975730]

NIH-PA Author Manuscript
NIH-PA Author Manuscript

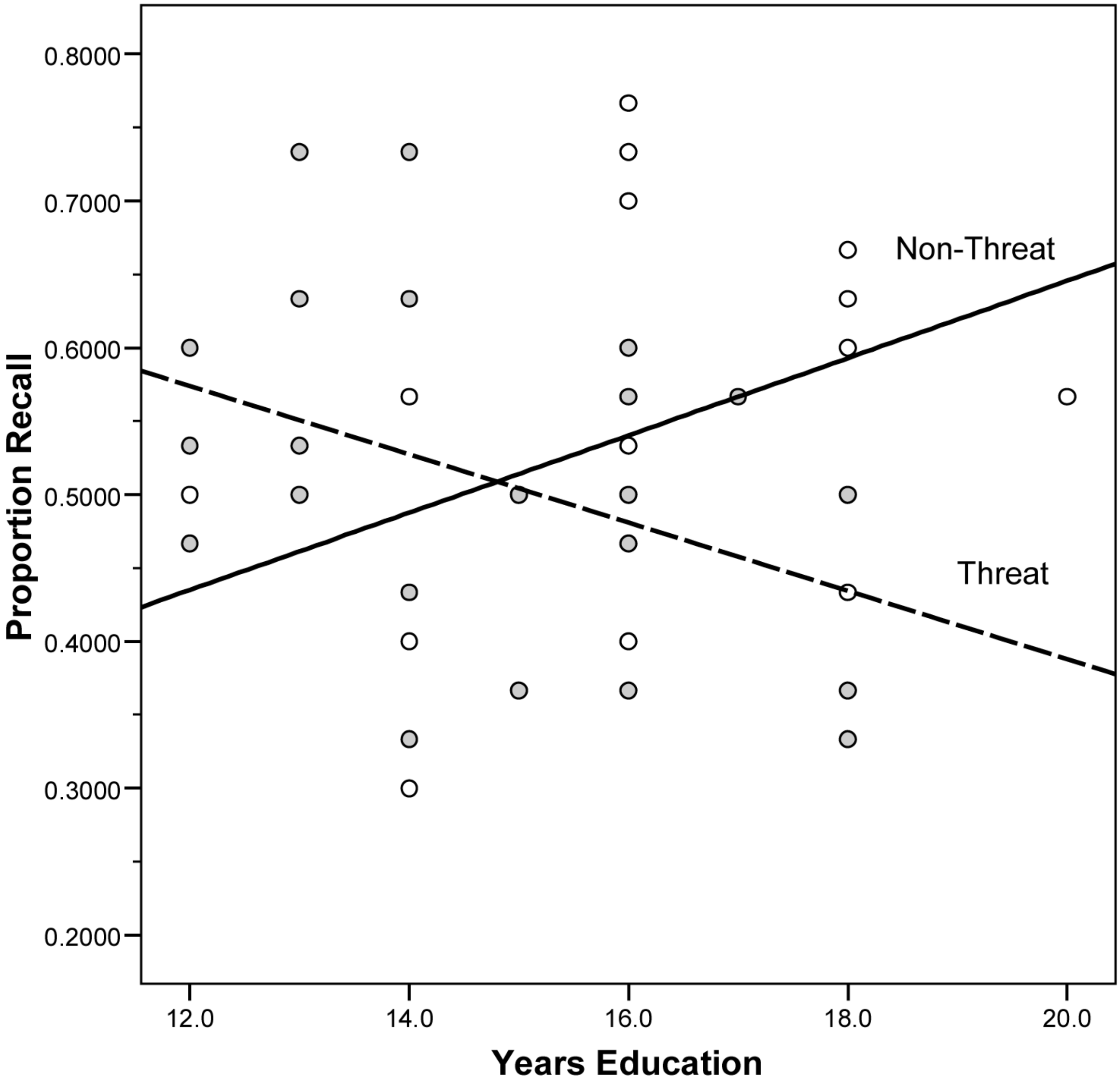

Figure 1.
Recall in the young-old group as a function of education and condition (threat = filled circles; non-threat = open circles).

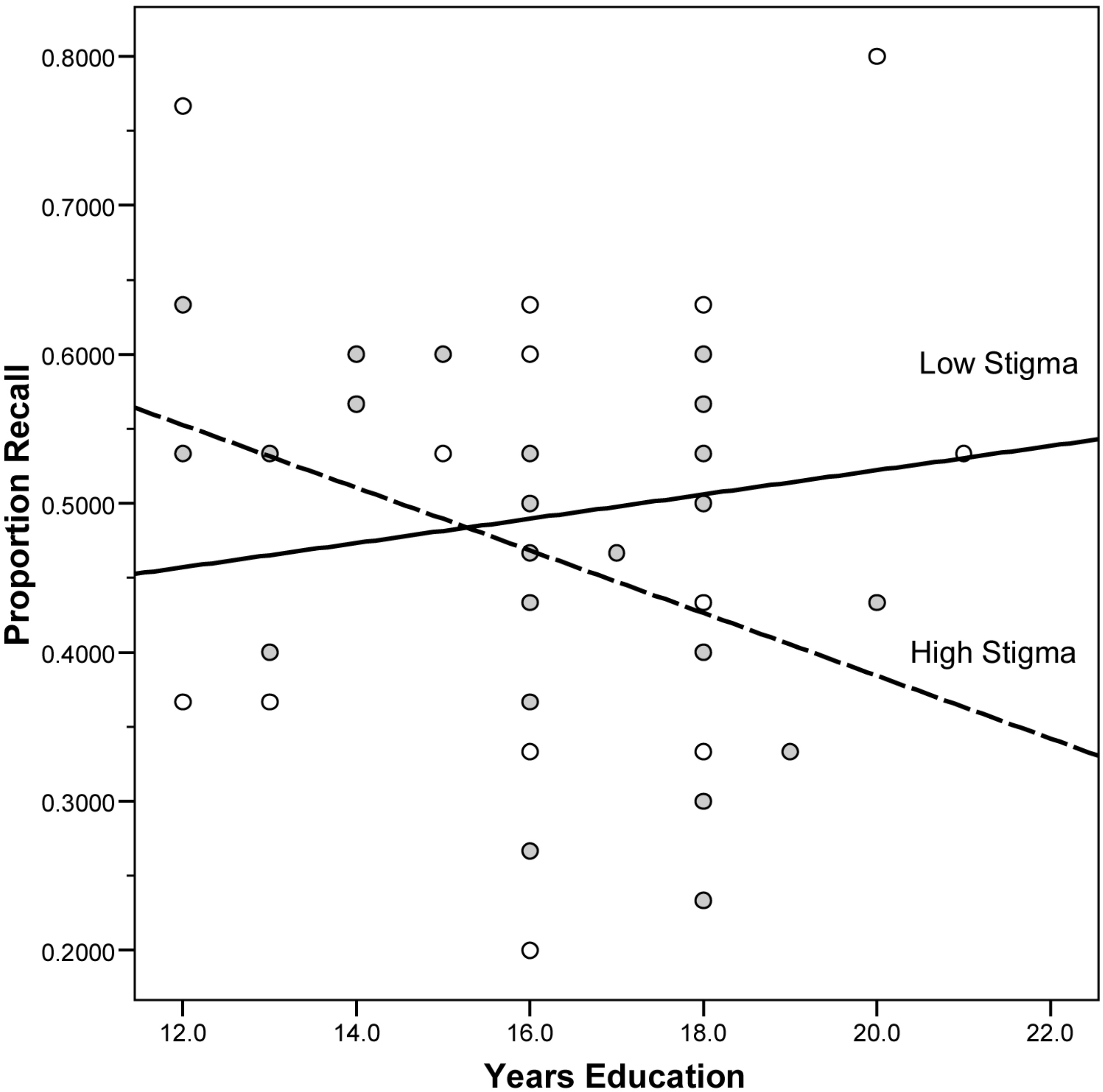

Figure 2.
Recall in the old-old group as a function of education and stigma consciousness (high = filled circles; low = open circles).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Participant Characteristics

	Age Group and Condition			
	Young-old		Old-old	
Variable	**Threat**	**Non-Threat**	**Threat**	**Non-Threat**
Age (years)	64.0	64.3	75.0	75.7
Education (years)	14.7	16.0	16.5	16.0
SF-36 Physical Health	44.0	44.9	43.1	43.7
SF-36 Mental Health	49.8	56.2	56.3	55.9
Vocabulary	29.4	30.1	31.0	30.0
Speed	.16	.35	−.04	−.26
Stigma Consciousness	33.7	34.2	33.5	32.0

Note: SF-36 scores represent norm-based *T*-scores. Vocabulary scores could range from 0 to 36. Speed scores are means of sample-based standardized scores for the letter and pattern comparison tasks. Stigma consciousness scores could range from 0 to 60, with lower scores reflecting greater stigma consciousness.

Table 2

Skin Conductance as a Function of Threat Condition and Time

Condition		Baseline	Instructions	Study	Recall
		Skin Conductance			
Threat	*M*	2.02	2.36	4.08	4.01
	SD	1.19	1.43	2.24	2.13
Non-Threat	*M*	1.45	1.71	2.95	2.97
	SD	.80	1.02	1.65	1.65
		Skin Conductance Variability			
Threat	*M*	.14	.13	.25	.24
	SD	.12	.14	.20	.19
Non-Threat	*M*	.08	.10	.16	.19
	SD	.09	.09	.12	.19