# Task Demands Moderate Stereotype Threat Effects on Memory Performance

Thomas M. Hess, Lisa Emery, and Tara L. Queen

Department of Psychology, North Carolina State University, Raleigh.

**Previous research has demonstrated that older adults' memory performance is adversely affected by the explicit activation of negative stereotypes about aging. In this study, we examined the impact of stereotype threat on recognition memory, with specific interest in (a) the generalizability of previously observed effects, (b) the subjective experience of memory, and (c) the moderating effects of task demands. Older participants subjected to threat performed worse than did those in a nonthreat condition but only when performance constraints were high (i.e., memory decisions had to be made within a limited time frame). This effect was reflected in the subjective experience of memory, with participants in this condition having a lower ratio of "remember" to "know" responses. The absence of threat effects when constraints were minimal provides important boundary information regarding stereotype influences on memory performance.**

**Key Words:** Aging—Memory—Stereotype threat.

RESEARCH has demonstrated that older adults' memory performance is adversely affected when negative stereotypes about aging are activated in the performance context (for review, see Hess, 2006). The construct of stereotype threat (Steele, Spencer, & Aronson, 2002) has been used in several studies to characterize the mechanisms underlying such effects. In these studies, negative stereotypes are explicitly activated through the labeling of the task (e.g., "your memory is going to be tested") or by highlighting the aging-related bases of performance (e.g., "this study will examine aging effects on performance"), and older adults subjected to such cues typically demonstrate lower levels of performance than similarly aged individuals receiving less stereotype-relevant information (e.g., Hess, Auman, Colcombe, & Rahhal, 2003; Rahhal, Hasher, & Colcombe, 2001). Whereas existing research has identified individual characteristics and responses that moderate such effects (e.g., Chasteen, Bhattacharyya, Horhota, Tam, & Hasher, 2005; Hess et al.), less is known about the impact of threat on the subjective experience of memory and the specific task characteristics associated with threat-based effects.

The present study addressed three specific questions related to these issues. First, we were interested in the generalizability of threat effects across memory tasks. Previous studies have demonstrated that threat as well as perceptions of one's control over memory—which is related to threat—affect memory performance through their impact on strategic behavior (Hess & Hinson, 2006; Hess et al., 2003; Lachman & Andreoletti, 2006). These studies have relied on free recall tests, which have a strong effortful and strategic component to them. If threat effects are mediated by strategic behavior, through either disruptions or diversions of resources, then such effects may be minimized on tasks

that have a less dominant strategic component. To this end, we examined threat effects on recognition memory performance, which is less demanding than free recall and less negatively affected by age.

Second, we were also interested in whether threat effects would increase with task demands. Increasing constraints on performance could elevate demands on cognitive resources, even in tasks such as recognition, which are often associated with minimal age effects. This, in turn, could further highlight the diagnosticity of the test with respect to the stereotyped skill, increasing the probability of threat effects. To test this hypothesis, we manipulated the test context so that individuals were either required to make recognition responses within a limited time frame or were given unlimited time to respond. It was hypothesized that threat effects would be greater in the former condition, with stereotype-based effects on older adults' memory attenuated when task demands were reduced.

Finally, we also examined whether threat might influence one's subjective experience of memory. Schmader, Johns, and Forbes (2008) argue that affective and self-regulatory responses triggered by threat negatively influence performance by diverting working memory resources normally devoted to controlled processing. This, in turn, may affect conscious recollection processes that support memory performance, perhaps altering the nature of encoding and retrieval processes. For example, threat-induced thoughts about one's own abilities might negatively affect the ability to retrieve specific attributes of encoded information, resulting in retrieval of disproportionately more general memories. We test this hypothesis by examining remember (R) versus know (K) judgments for positive recognition responses (Gardiner & Richardson-Klavehn,

2000). These judgments distinguish between an individual's retrieval of a specific episodic event versus a more general feeling of familiarity. If threat affects subjective experience, this may be manifested in relatively fewer *R* judgments.

## METHODS

### Participants

Ninety-five older adults were originally recruited for the study through newspaper advertisements and were paid $20 for their time. None of these individuals had participated in previous studies examining stereotype threat. Eleven participants were dropped due to data issues (e.g., answering "yes" to all items, missing data). Two additional participants were dropped: one due to a high score (>6) on the short Blessed test (Katzman et al., 1983) and the other due to an SF-36 Mental Health score 1 *SD* below the rest of the sample. The remaining 39 men and 43 women ranged in age from 60 to 86 years.

### Materials and Procedure

Individuals completed a set of background questionnaires, including the SF-36 Health Survey (Ware, 1993), prior to coming to the laboratory. Participants were tested individually and were assigned to one of four experimental conditions representing the crossing of stereotype condition (threat vs. nonthreat) and response condition (deadline vs. unlimited response time). Those in the threat condition received the following instructions prior to the memory task:

> One goal of this study is to examine age differences in memory ability. I am now going to examine your memory ability using a test that has been used extensively by researchers to study aging effects on memory. Younger adults typically do much better than older adults on this task.

They also typed their age into the computer prior to both study and test to reinforce threat. Those in the nonthreat condition received the following instructions:

> One goal of this study is to examine individual differences in ability and the factors that account for those differences. I am now going to examine your ability to process verbal information. In an effort to reduce potential biases, we will be using a task that has been shown to be appropriate for individuals of all ages. Interestingly, older adults have been shown to do quite well on this task.

Participants were then instructed to study 50 words (five moderately high exemplars from each of 10 different semantic categories) which were presented in random order at a 2-s rate by computer. They then performed several unrelated tasks for approximately 10 min, after which their memory was tested.

The recognition test consisted of the 50 targets interspersed with 50 distracters (five new exemplars from each of the same 10 categories), with one target and one distracter from each category presented in each quintile of the test list. Participants made a yes/no recognition response to each item. If they responded "yes," they were then prompted to make a remember, know, or guess (RKG) response using the instructions contained in Gardiner and Richardson-Klavehn (2000, p. 240). Participants in the deadline condition were given 2,500 ms to make a recognition response, whereas those in the other condition were given unlimited time. If participants in the deadline condition did not respond in the allotted time, the computer displayed the message "too slow" and moved to the next test item. No time constraints were placed on RKG responses. Note that the same study and test lists were used for all subjects. Although not an uncommon practice, this could limit generality of the findings.

Following the recognition test, participants rated how well they performed using a 7-point scale. They then completed the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) and a threat assessment scale (Chasteen et al., 2005), after which they were debriefed. Finally, they completed the Wechsler Adult Intelligence Scale III (WAIS III) Vocabulary, Letter–Number Sequencing, and Digit–Symbol subtests (Wechsler, 1997).

## RESULTS

### Preliminary Analyses

We conducted a series of $2 \times 2$ (Stereotype Condition $\times$ Response Condition) analyses of variance (ANOVAs) on age, self-rated health, years of education, and scores on the WAIS III subtests to determine if there were any significant variations—and, thus, potential confounds—across conditions (Table 1). The only significant effect that emerged was an interaction effect for vocabulary, $F(1, 78) = 5.01$, $p = .03$, $\eta_p^2 = .06$, due to a significantly higher mean score in the threat-unlimited condition relative to the other conditions. Because controlling for it did not affect the results of the following analyses, vocabulary was not included as a covariate.

### Memory

Participants in the deadline condition responded within the allotted time to 96.4% of the targets and 94.9% of the distracters, with response rate not varying across Stereotype Conditions, $F < 1$. Memory performance was determined by calculating corrected recognition scores (*p.* hits − *p.* false alarms), which were then examined using a $2 \times 2$ (Stereotype Condition $\times$ Response Condition) ANOVA. Performance in the deadline condition was significantly worse than in the unlimited condition ($Ms = 0.47$ vs. 0.57), $F(1, 78) = 6.09$, $p = .02$, $\eta_p^2 = .07$, and the predicted interaction approached significance, $F(1, 78) = 3.87$, $p = .053$, $\eta_p^2 = .05$. As seen in Figure 1, the negative effects of threat on performance were evident under only deadline conditions. Follow-up

Table 1. Participant Characteristics: $M$ ($SD$)

| | Test conditions | | | |
| --- | --- | --- | --- | --- |
| | Threat | | Nonthreat | |
| Variable | Deadline ($n = 23$) | Unlimited ($n = 19$) | Deadline ($n = 22$) | Unlimited ($n = 18$) |
| Age (years) | 70.4 (7.2) | 70.1 (6.8) | 71.7 (6.7) | 71.3 (5.6) |
| Education (years) | 16.4 (2.8) | 16.0 (1.8) | 16.8 (2.3) | 16.5 (2.5) |
| SF-36 Physical Health | 42.6 (4.9) | 42.0 (3.3) | 42.9 (3.8) | 41.9 (3.4) |
| SF-36 Mental Health | 52.1 (3.7) | 53.0 (5.0) | 53.3 (5.3) | 54.0 (3.2) |
| Vocabulary | 50.4 (8.3) | 55.9 (5.1) | 51.9 (7.4) | 49.1 (11.6) |
| Letter–Number Sequencing | 10.1 (3.1) | 11.6 (2.6) | 10.2 (3.0) | 10.3 (2.0) |
| Digit–Symbol | 47.5 (9.8) | 48.8 (8.6) | 46.2 (7.9) | 42.6 (12.0) |

*Note*: Scores on the SF-36 were norm-based $T$ scores. Vocabulary scores could range from 1 to 66, Letter–Number Sequencing scores could range from 1 to 21, and Digit–Symbol substitution scores could range from 1 to 133.

contrasts testing our specific predictions revealed that stereotype threat reduced memory in the deadline condition, $F(1, 43) = 4.59$, $p = .04$, but not in the unlimited response condition, $F < 1$.

We next examined subjective experiences of memory. To simplify these analyses, we calculated corrected recognition responses by subtracting the proportions of $R$ and $K$ responses to distracters from the same proportions to targets. (One statistical outlier was excluded.) A Stereotype Condition × Response Condition × Response Type ANOVA revealed that means for response types varied as a function of condition, as reflected in significant Response Condition × Response Type, $F(1, 77) = 4.56$, $p = .04$, $\eta_p^2 = .05$, and Stereotype Condition × Response Condition × Response Type, $F(1, 77) = 4.76$, $p = .03$, $\eta_p^2 = .06$, interactions. As seen in Figure 2, there was little variation in $K$ response rates across conditions. In contrast, follow-up analyses revealed that $R$ response rates were significantly higher in the unlimited

versus deadline condition (.50 vs. .39), $F(1, 77) = 6.86$, $p = .01$. In addition, the impact of stereotype condition was significant in the deadline response condition, $F(1, 42) = 7.03$, $p = .01$, but not in the unlimited response condition, $F < 1$.

We also examined self-assessments of performance as another subjective index of memory. Those in the deadline condition thought they performed worse than did those in the unlimited response condition ($M$s = 4.1 vs. 4.7), $F(1, 78) = 4.22$, $p = .04$, $\eta_p^2 = .05$. Threat did not reliably moderate this effect ($p = .26$), but, as before, the impact of test condition was greater in the threat condition ($M_{\text{deadline}} = 4.0$; $M_{\text{unlimited}} = 4.8$; $p = .04$) than in the nonthreat condition ($M_{\text{deadline}} = 4.3$; $M_{\text{unlimited}} = 4.6$; $p = .50$).

### Affective Responses

Examination of PANAS scores did not reveal significant effects due to either experimental condition, nor was there significant variation in perceived threat across conditions.
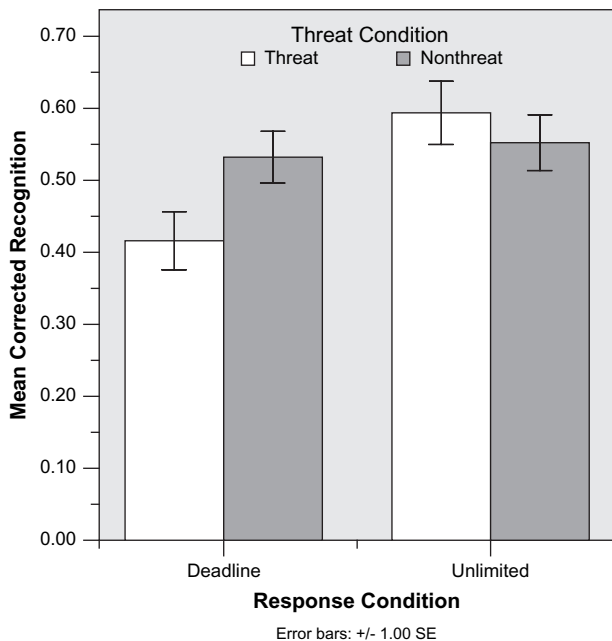


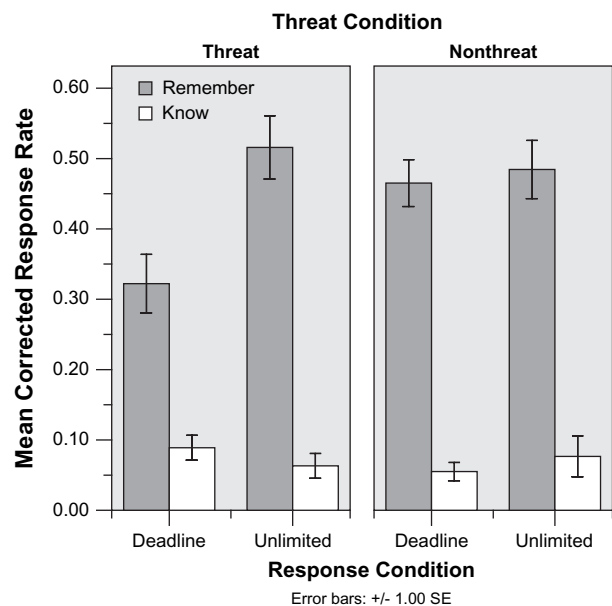Figure 1. Mean corrected recognition by stereotype and test condition.



Figure 2. Mean proportion of corrected "remember" and "know" responses.

## DISCUSSION

The results of this experiment are relatively straightforward. First, we found evidence of stereotype threat having a negative effect on performance with a different memory task than used previously. Of interest is the fact that these effects were observed in a task typically associated with both minimal aging effects and minimal strategic involvement. This result seems inconsistent with current conceptualizations of threat effects that rely on working memory as a mediator (Schmader et al., 2008). The seeming inconsistency may be resolved, however, when considering our other two major findings: Threat effects were most evident when constraints on performance were high, with the effect appearing to be tied to changes in the subjective experience of memory. Given that the stereotype effects were evident only when the response window was constrained at test, threat appears to have had its primary impact on retrieval in the present study. This leads to two possible interpretations of these effects.

First, reductions in *R* responses associated with threat in the deadline condition may reflect alterations in recollection processes related to increased demands on processing resources. For example, Johnson (1992) has suggested that details associated with recollection take time to accumulate, a process which may be affected by the confluence of threat—which may divert resources—and the limited response window. Consistent with this perspective, Skinner and Fernandes (2008) found that divided attention at retrieval altered the accuracy of *R* responses. (Note, however, that other research has found that divided attention effects are stronger for *K* than for *R* responses [e.g., Knott & Dewhurst, 2007]). An alternative—and not unrelated—explanation is based in the possibility that *R* and *K* responses reflect different levels of confidence (Dunn, 2004). Threat combined with strong constraints on performance may alter individuals' confidence in their recognition judgments, resulting in a higher proportion of *K* responses relative to *R* responses. This interpretation is in line with other research demonstrating that threat is associated with increased self-doubt (e.g., Steele & Aronson, 1995) and worry (e.g., Cadinu, Maass, Rosabianca, & Kiesner, 2005). This interpretation is also consistent with participants' self-reported assessments of performance. Distinction between these perspectives awaits future research. Of importance, however, is the finding that threat effects can be counteracted by changing the situational constraints on performance.

Although we attribute the observed effects to stereotype threat, the absence of group differences in subjective reports of threat may suggest alternative mechanisms, such as self-stereotyping (O'Brien & Hummert, 2006). This latter process operates through automatic (and unconscious) activation of stereotype-consistent behaviors, which may result in similar behavioral effects as observed with threat but without the affective response. Our preference for a threat-based interpretation relates to the fact that the observed effects were associated with our manipulation of the salience of participant age, the diagnosticity of the task, and group-based performance expectancies, all of which are associated with threat effects (Steele et al., 2002). In addition, the measure of threat employed in our study (Chasteen et al., 2005) assesses responses associated with concerns about others' reactions. As noted earlier, it is possible that threat may affect performance through other self-evaluative concerns (see also Hess & Hinson, 2006), and thus, the null effects associated with responses to the scale by Chasteen et al. are not necessarily inconsistent with the operation of stereotype threat.

In conclusion, the present study has extended previous research by demonstrating threat effects on older adults' memory performance in a new task domain. Previous research employing free recall suggested that threat effects were operative during encoding, but the present results indicate that threat effects are operative at retrieval as well. This study also adds to our understanding of the mechanisms, suggesting that threat may operate by altering one's subjective experience of memory, perhaps due to the diversion of processing resources necessary for supporting performance. Finally, we have also demonstrated that threat effects on memory are not inevitable and are most likely to occur when constraints on performance are high. Such conditions may accentuate the feelings of threat that one experiences in the performance context.

REFERENCES

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science*, *16*, 572–578.

Chasteen, A. L., Bhattacharyya, S., Horhota, M., Tam, R., & Hasher, L. (2005). How feelings of stereotype threat influence older adults' memory performance. *Experimental Aging Research*, *31*, 235–260.

Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review*, *111*, 524–542.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 229–244). New York: Oxford University Press.

Hess, T. M. (2006). Attitudes toward aging and their effects on behavior. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (6th ed., pp. 379–406). San Diego, CA: Academic Press.

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*, *58*, P3–P11.

Hess, T. M., & Hinson, J. T. (2006). Age-related variation in the influences of aging stereotypes on memory in adulthood. *Psychology and Aging*, *21*, 621–625.

Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience*, *4*, 268–280.

Katzman, R., Brown, T., Fuld, P., Schecter, R., & Schimmel, H. (1983). Validation of a short orientation-memory-concentration test of cognitive impairment. *American Journal of Psychiatry*, *140*, 734–739.

Knott, L. M., & Dewhurst, S. A. (2007). Divided attention at retrieval disrupts knowing but not remembering. *Memory*, *15*, 664–674.

Lachman, M. E., & Andreoletti, C. (2006). Strategy use mediates the relationship between control beliefs and memory performance in middle-aged and older adults. *Journal of Gerontology: Psychological Sciences*, *61*, P88–P94.

O'Brien, L., & Hummert, M. L. (2006). Age self-stereotyping, stereotype threat, and memory performance in late middle-aged adults. *Social Cognition*, *24*, 338–358.

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

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*, 336–356.

Skinner, E. I., & Fernandes, M. A. (2008). Interfering with remembering and knowing: Effects of divided attention at retrieval. *Acta Psychologica*, *127*, 211–221.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379–440). San Diego, CA: Academic Press.

Ware, J. E., Jr. (1993). *SF-36 Health Survey*. Boston: The Health Institute, New England Medical Center.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality & Social Psychology*, *54*, 1063–1070.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition*. New York: Psychological Corporation.