FALSE memories occur when a remembered event either never occurred or the memory differs substantially from reality. Events rich in meaning are typically remembered through a reconstructive process prone to memory errors. In efforts to reconstruct life events (Johnson, 2006; Schacter, 1999), one attempts to fill gaps with appropriate information, but this information may be inaccurate (Roediger & McDermott, 1995). False memories are particularly problematic when mistaken information informs real-world decision making, such as inaccurate eyewitness testimony (Johnson, 2006). Older adults are especially at risk because they not only experience lower rates of accurate memory but an increase in false memory (Schacter, Koutstaal, & Norman, 1997).

One strategy that improves memory is relating information to oneself or self-referencing (Symons & Johnson, 1997). Although self-referencing has been shown to increase memory, it also could contribute to memory errors under some circumstances. The self enables us to organize information (Symons & Johnson, 1997) enabling deeper processing (Craik & Lockhart, 1972; Hartlep & Forsyth, 2000) and thus better memory for self-related information. For example, we are more likely to retrieve words that we judged for self-descriptiveness rather than for the descriptiveness of another person (Glisky & Marquine, 2009; Gutchess, Kensinger, & Schacter, 2007, 2010; Gutchess, Kensinger, Yoon, & Schacter, 2007). Self-reference also benefits memory because the self-schema is a highly organized and elaborate construct that makes self-referencing an efficient means of processing information related to the self (Greenwald & Banaji, 1989; Klein & Loftus, 1988; Kuiper & Rogers, 1979; Rogers, Kuiper, & Kirker, 1977; Symons & Johnson, 1997). The factors that allow self-reference to benefit memory, however, may also impair memory performance. The deeper processing associated with self-reference may trigger the retrieval of additional associations, which have been shown to increase false memories (Roediger & McDermott, 1995). When participants encode a word, related words are triggered, known as the implicit associative response. As a result, one can experience intrusions of associated words and may report having previously studied them. The more associations are activated, the stronger the associations become and the greater the likelihood that intrusions will occur. As a result, items connected to the self may trigger associations with other self-related items and lead to greater false remembering. As a result of activation of these associations, self-referencing could contribute to false memory by altering response bias (Niezmanski, 2009). Participants may be more likely to say that self-related information is “old” than information that is not self-related, resulting in higher hit and false alarm rates. Information with a strong connection to the self is likely to be frequently accessed and may have increased fluency. Fluency of processing should increase the familiarity of information such that an individual may falsely believe she has previously studied a new word, especially if it is
highly self-relevant (Jacoby & Whitehouse, 1989; Thapar & Westerman, 2009). Based on this, we predict that the increased familiarity as a result of fluency associated with the self may lead to greater false memory. Although young and older adults may be equally susceptible to fluency-based intrusions (Thapar & Westerman, 2009), young adults seem to experience fewer false memories because they are better at distinguishing between information that has been internally generated versus externally perceived (Dehon & Brédart, 2004). It is also possible that greater familiarity works in conjunction with implicit associative responses, such that frequent access and deep reflection on highly self-descriptive information increases both its familiarity and its connection to other information in memory.

Although feelings of familiarity may increase false memory with age, source memory may contribute as well. Older adults have difficulty distinguishing between similar sources, resulting in memory errors (Ferguson, Hashtroudi, & Johnson, 1992; Henkel, Johnson, & DeLeonardis, 1998). Self-referencing, however, improves source memory (Dulas, Newsome, & Duarte, 2011; Hamami, Serbun, & Gutches, 2011; Leshikar & Duarte, 2011; Rosa & Gutches, 2011; Serbun, Shih, & Gutches, 2011). However, when the self is the source of both items (e.g., “I thought of this word” or “I read this word”), the relationship to the self may be the most important feature of the memory, thereby increasing the difficulty of determining the source. If older adults experience difficulty distinguishing between sources of self-related information, self-referencing could actually lead to greater false memory.

The present study examined whether the degree of self-reference contributes to false memory with age, such that highly self-descriptive information induces memory errors. Few studies examine the link between false memory and self-reference or invoke contexts in which self-referencing increases vulnerability to memory errors. Rogers, Rogers, and Kuiper (1979) found that young adults’ false memory for adjectives increased as the degree of self-reference increased. We extended this work into the study of aging based on findings that older adults are more prone to false recognition than younger adults. Although the mechanisms that contribute to greater false memories for self-referenced information may be similar across the age groups, such as the fact that the self is a well-known entity (Greenwald & Banaji, 1989; Kuiper & Rogers, 1979; Rogers et al., 1977; Symons & Johnson, 1997) that enables deeper processing (Craik & Lockhart, 1972; Hartlep & Forsyth, 2000) and increased associations, other processes may differ with age, making older adults more prone to false memories due to self-referencing. Older adults may be more likely to accept easily accessed information, such as that related to the self, as true (Jacoby & Rhodes, 2006; McDaniel, Einstein, & Jacoby, 2008). It is also possible that older adults’ increased vulnerability to intrusions caused by familiar or well-known information (Jacoby & Rhodes, 2006) in combination with difficulty correctly identifying the source of these intrusions as internal or external (Dehon & Brédart, 2004) puts older adults at a greater risk of false memory. Thus, we hypothesized that both young and older adults would not only have greater hit rates for highly self-descriptive words compared with those rated as less descriptive but that younger and older adults would exhibit greater rates of false alarms for highly self-descriptive items than less descriptive items. However, due to vulnerability to increased intrusions and difficulties in determining the source of information as external (i.e., seen) or internal (i.e., self-generated), false alarms for self-descriptive words would be greater for older than younger adults. The increase in false memory could mitigate any potential benefits in memory for highly self-relevant words, at least for older adults.

**EXPERIMENT 1**

In Experiment 1, we explore whether false alarms increase as information becomes more self-relevant (Rogers et al., 1979). We expected to see a greater increase in false alarms for highly self-relevant information for older compared with younger adults.

**Method**

Participants.—Data from 62 participants (31 young and 31 older) were included in analyses. Seven additional participants (four young and three older) were excluded due to missing data or failure to follow instructions (e.g., incomplete trials). Older adult participants were recruited from the community. Young adults were recruited from Brandeis departmental study pools. All participants provided written informed consent for a protocol approved by the Brandeis University Institutional Review Board.

Older adults demonstrated significantly higher vocabulary scores (Shipley, 1940) than younger adults, \( t(57) = 3.61, p = .001 \) (data missing from two young and one older adults), and a higher average level of education, \( t(37.65) = 6.56, p < .001 \) (equal variances not assumed; see Table 1). Younger adults demonstrated significantly faster processing speed on a digit comparison task (Hedden

**Table 1. Study 1 Participant Demographics, Including Means and Standard Deviations**

<table>
<thead>
<tr>
<th>Study 1</th>
<th>Young adults</th>
<th>Older adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>19.28 (.10)</td>
<td>75.10 (6.00)</td>
</tr>
<tr>
<td>Range</td>
<td>18–22</td>
<td>64–84</td>
</tr>
<tr>
<td>N</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>Gender</td>
<td>13M, 18 F</td>
<td>12M, 19 F</td>
</tr>
<tr>
<td>Years of education**</td>
<td>12.82 (.99)</td>
<td>16.26 (2.74)</td>
</tr>
<tr>
<td>Digit comparison**</td>
<td>76.15 (15.50)</td>
<td>58.26 (9.74)</td>
</tr>
<tr>
<td>Shipley vocabulary*</td>
<td>32.66 (3.30)</td>
<td>35.97 (3.72)</td>
</tr>
</tbody>
</table>

*Notes. *Significant difference between young and older adults, \( p < .01 \).
**Significant difference between young and older adults, \( p < .001 \).
Materials.—To control for effects that were word specific, 3 lists of 120 adjectives (e.g., “cultured,” “sensible”) were created with an equal number of positive and negative words and equated for likeability (Anderson, 1968). Words contained an average of 8.9 letters and the mean frequency scores ranged from 14.63 to 34.22 across the three lists (ratings available for 80% of words; Kucera & Francis, 1967). Lists were counterbalanced so that each appeared equally often at encoding and as lure words not presented at encoding.

Procedure.—During encoding, participants rated 120 adjectives for self-descriptiveness on a 9-point scale (1 = never describes me; 9 = always describes me). To ensure an equal amount of time was spent encoding each item, adjectives were presented for 5 s and responses were made with a keypress. Following a 15-min retention interval, participants completed a surprise, self-paced recognition test consisting of the previously seen 120 adjectives plus 120 new adjectives. Recognition was self-paced to ensure that a recognition response was recorded for each item. Participants indicated with a keypress whether each adjective was “old” or “new.” Participants then rated the 120 “new” adjectives for self-descriptiveness on the same 9-point scale in a self-paced task.

Scoring.—Ratings were collapsed into three rating levels. Words rated as 1–3 were considered “low,” 4–6 were “middle,” and 7–9 were “high.” This prevented loss of participants due to missing data, as not all participants used every number on the scale. See Table 2 for proportion of items in each category.

Results

Hits.—Hit rates were calculated as a proportion of correctly recognized old items from each level of rating. Using a factorial analysis of variance (ANOVA) with age (young/older) as the between-subject variable and rating (low/middle/high) as the within-subject variable, the main effect of age was marginal, \( F(1, 60) = 3.79, p = .06, \eta^2 = .06 \), with younger adults’ hit rates (\( M = .80 \)) higher than older adults’ (\( M = .71 \)). The main effect of rating was significant, \( F(2, 120) = 18.06, p < .001, \eta^2 = .23 \). Follow-up contrasts showed that hit rates for the entire sample were higher for highly self-descriptive words (\( M = .80 \)) than low (\( M = .72 \); \( F(1, 60) = 27.49, p < .001, \eta^2 = .31 \)) or middle self-descriptive words (\( M = .74 \); \( F(1, 60) = 23.34, p < .001, \eta^2 = .28 \)). The interaction between rating and age was also significant, \( F(2, 120) = 9.34, p < .001, \eta^2 = .14 \). Contrasts revealed that young adults had significantly higher hit rates for high (\( M = .81 \); \( F(1, 30) = 6.67, p = .02, \eta^2 = .18 \)) and middle (\( M = .81 \); \( F(1, 30) = 9.73, p = .004, \eta^2 = .25 \)) self-descriptive words than low (\( M = .76 \)) self-descriptive words. In contrast, older adults had significantly higher hit rates for words rated as highly self-descriptive (\( M = .79 \)) than words at low (\( M = .68 \); \( F(1, 30) = 21.58, p < .001, \eta^2 = .42 \)) or middle levels (\( M = .67 \); \( F(1, 29) = 33.17, p < .001, \eta^2 = .53 \); see Figure 1a). Note that all significant effects are reported throughout.

False alarms.—False alarms were calculated as a proportion of falsely recognized new items. The effect of self-reference on false alarms was tested with a factorial ANOVA with age (young/older) as the between-subject factor and rating (low/middle/high) as the within-subject factor. The main effect of age on false alarm rate was significant, \( F(1, 60) = 13.23, p = .001, \eta^2 = .18 \), such that older adults (\( M = .30 \)) committed more false alarms than younger adults (\( M = .18 \)). The main effect of rating, \( F(2, 120) = 41.96, p < .001, \eta^2 = .41 \), was significant, with contrasts revealing that highly descriptive words (\( M = .31 \)) led to more errors than words in the low (\( M = .19 \); \( F(1, 60) = 69.68, p < .001, \eta^2 = .54 \)) or middle (\( M = .21 \); \( F(1, 60) = 37.84, p < .001, \eta^2 = .39 \)) rating levels. The interaction between age and rating, \( F(2, 120) = 6.33, p = .002, \eta^2 = .10 \), was also significant. Follow-up contrasts revealed significant interactions with age between the high and low rating levels, \( F(1, 60) = 9.26, p = .003, \eta^2 = .13 \), and the high and middle rating levels, \( F(1, 60) = 7.55, p = .008, \eta^2 = .11 \). There was no significant difference between the middle and low rating levels, \( p > .10 \). As shown in Figure 1b, older adults exhibited higher false alarm rates than younger adults, particularly for the high condition.

Table 2. Means (SD) in Each Condition for Study 1

<table>
<thead>
<tr>
<th>Study 1</th>
<th>Low</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young adults</td>
<td>.30 (.09)</td>
<td>.35 (.13)</td>
<td>.33 (.07)</td>
</tr>
<tr>
<td>Older adults</td>
<td>.34 (.07)</td>
<td>.26 (.14)</td>
<td>.34 (.10)</td>
</tr>
<tr>
<td>Hit</td>
<td>.76 (.18)</td>
<td>.81 (.17)</td>
<td>.81 (.17)</td>
</tr>
<tr>
<td>Miss</td>
<td>.24 (.18)</td>
<td>.19 (.17)</td>
<td>.19 (.17)</td>
</tr>
<tr>
<td>FA</td>
<td>.15 (.08)</td>
<td>.17 (.08)</td>
<td>.22 (.12)</td>
</tr>
<tr>
<td>CR</td>
<td>.85 (.08)</td>
<td>.83 (.09)</td>
<td>.78 (.12)</td>
</tr>
<tr>
<td>Hit-FA</td>
<td>.61 (.18)</td>
<td>.64 (.16)</td>
<td>.59 (.19)</td>
</tr>
</tbody>
</table>

Notes. Bolded proportions correspond to analyses presented in text; CR = corrected recognition; FA = false alarms.

et al., 2002), \( \eta(47.02) = 5.09, p < .001 \) (equal variances not assumed). These findings suggest that our samples are representative of previous ones in the literature.
Corrected recognition.—We examined the corrected recognition scores by subtracting false alarm rates from hit rates for each individual. A factorial ANOVA with age (young/older) as the between-subject variable and rating level (high/middle-low) as the within-subject variable showed that the main effect of rating was marginally significant, $F(2, 120) = 2.63, p = .08, \eta^2_p = .04$, with contrasts showing trends for higher correct recognition scores in the low ($M = .53; F(1, 60) = 4.14, p = .05, \eta^2_p = .06$) and middle ($M = .53; F(1, 60) = 3.52, p = .07, \eta^2_p = .06$) rating levels.
than in the high rating level ($M = .49$). The main effect of age was significant, $F(1, 60) = 25.38, p < .001$, $\eta^2_p = .30$, with younger adults ($M = .62$) performing better than older adults ($M = .42$).

Response bias.—C values were used in a $2 \times 3$ ANOVA with age (young/older) as the between-subject variable and rating level (high/middle/low) as the within-subject variable to assess response bias (Stanislaw & Todorov, 1999). There was a main effect of rating, $F(2, 120) = 34.17, p < .001$, $\eta^2_p = .36$, and an interaction between age and rating, $F(2, 120) = 6.97, p = .001, \eta^2_p = .10$. For older adults, contrasts indicated that words rated as highly self-descriptive ($M = −.36$) had a more negative bias (i.e., greater use of "old") than those in both the low ($M = .15; F(1, 30) = 62.62, p < .001, \eta^2_p = .68$) and middle ($M = .12; F(1, 30) = 72.95, p < .001, \eta^2_p = .71$) rating levels. Although young adults’ response bias also was more negative for words rated as highly self-descriptive ($M = −.08$) than those in the low rating level ($M = .17; F(1, 30) = 8.13, p = .008, \eta^2_p = .21$), there was no significant difference for the middle level ($M = .05; p > .10$). Participants were more likely to say that words rated as highly self-descriptive were previously seen and this was particularly true for older adults, who showed a larger difference between the high and medium levels, compared with younger adults (see Table 2).

Discussion
Findings from Experiment 1 indicate that as information becomes more self-descriptive, hit rates increase for both young and older adults. This finding converges with other studies in demonstrating that self-referencing can improve recognition memory for young and older adults (Glisky & Marquine, 2009; Gutchess et al., 2007, 2010; Gutchess, Kensinger, Yoon, et al., 2007). However, our results suggest that older adults may be less sensitive than younger adults to information that is moderately self-descriptive. Older adults exhibited a reduced self-reference effect for the low and middle level of ratings, but performance was similar to younger adults for words in the high self-condition. This appears to indicate that for both older and younger adults, the self-reference effect similarly affected hit rates when information was highly self-relevant. The self-reference effect likely facilitated memory for information that was highly associated with the self-schema, which influences what we attend to and how we process information (Markus, 1977).

In previous studies (e.g., Glisky & Marquine, 2009; Gutchess, Kensinger, Yoon, et al., 2007), self-reference effects were driven by hit rates because there was only a single pool of new lure items that contributed to a single false alarm rate. Collecting self-relevancess ratings for lure words allowed us to assess the effect of self-referencing on false alarms across levels of self-relevance. The present study demonstrates that although self-referencing increased hit rates, it also contributed to false memory. Older adults seemed to be particularly affected by the highest level of self-descriptiveness, such that their false alarm rates were disproportionately higher at this level compared with young adults, whose rates rose more gradually across the three levels of self-descriptiveness. This study extends previous findings in young adults (Rogers et al., 1979) into aging, indicating a particularly robust false memory effect for older adults (>40% of the trials, on average) for highly self-relevant information.

Although both hit and false alarm rates tended to increase with self-descriptiveness, there was no net benefit on memory accuracy, according to corrected recognition scores. If anything, both young and older adults tended to exhibit higher corrected recognition scores for information that was less self-descriptive. High levels of self-descriptiveness did not improve the accuracy of memory and may have actually led to poorer discrimination of old from new words compared with items that were less self-descriptive. This is similar to findings from a previous study that utilized a measure of corrected recognition and found that an apparent benefit shown through hit rates can be decreased or eliminated when false alarms are considered (Koutstaal & Schacter, 1997). Rather than improving the accuracy of memory, self-referencing seems to influence the response bias such that people are more willing to endorse items that are self-descriptive as having been previously studied (Nieznanski, 2009).

Experiment 2
Without a control condition, it is difficult to tell if the findings from Experiment 1 are a result of self-referencing per se. Experiment 2 introduced a commonness judgment as a control condition to assess whether the increase in hit and false alarm rates in Experiment 1 was due to the connection of the information to the self or simply reflected endorsement of information. Commonness allows for a subjective rating of the words without explicitly referencing the self and should lead to a pattern of memory performance distinct from self-referencing (see Symons & Johnson, 1997). In a recognition task, the frequency mirror effect indicates that low frequency words are more distinctive and produce higher hit rates and lower false alarm rates compared with high frequency words, which produce lower hit rates and higher false alarm rates (Glanzer & Adams, 1985; Reder, Angstadt, Cary, Erickson & Ayers, 2002). In the present experiment, words judged to be less common should have higher hit rates and lower false alarm rates. In contrast, higher hit and false alarm rates should occur for words rated as highly self-descriptive than for words rated as less self-descriptive, as shown in Study 1 and previous work (Rogers et al., 1979). Furthermore, by comparing performance across the self and common conditions, we
predicted that memory for self-related information would be associated with better memory, as reflected in higher corrected recognition scores, as suggested by previous self-referencing studies (Symons & Johnson, 1997).

**Method**

Participants.—Data from a new sample of 69 participants (30 young and 39 older) were included in all analyses. Two additional older adult participants were excluded due to below normal performance on neuropsychological measures. Participants were drawn from the same participant pools and completed the same consent process as described in Study 1.

As shown in Table 3, our samples performed similarly to those in Study 1. Older adults demonstrated higher vocabulary scores ($M = 35.23$) than younger adults ($M = 32.20$) but this difference was only marginally significant, $t(67) = 1.94, p = .06$, and a higher average level of education, $t(58.33) = 2.28, p = .03$ (equal variances not assumed). Younger adults again demonstrated significantly faster processing speed, $t(66) = 7.69, p < .001$.

Materials.—Two lists of 120 adjectives rated for likeability (Anderson, 1968) were counterbalanced so that each list appeared equally often in the self and common conditions.

Procedure.—The methods in Study 2 replicated those from Study 1 with the following exceptions. During study, participants rated 60 adjectives for self-descriptiveness (“I am”) and 60 adjectives for commonness (“How common is the word”) on a 6-point scale (1 = never; 6 = always). Each adjective was presented for 6 s, followed by a blank screen for 1 s. Conditions alternated, with blocks of 7–8 words per condition. Words were randomly presented within each condition. A fixation cross appeared between each block to indicate a change in condition. Following a surprise, self-paced recognition test, participants rated the 120 “new” adjectives (60 for self-descriptiveness and 60 for commonness) on the same 6-point scale used during encoding, in a self-paced task.

Table 3. Study 2 Participant Demographics, Including Means and Standard Deviations

<table>
<thead>
<tr>
<th></th>
<th>Study 2 Young Adults</th>
<th>Older Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>20.73 (1.62)</td>
<td>75.62 (7.70)</td>
</tr>
<tr>
<td>Range</td>
<td>18–25</td>
<td>62–94</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>39</td>
</tr>
<tr>
<td>Gender</td>
<td>9M, 21 F</td>
<td>14M, 25 F</td>
</tr>
<tr>
<td>Years of education*</td>
<td>14.15 (1.45)</td>
<td>15.47 (2.92)</td>
</tr>
<tr>
<td>Digit comparison**</td>
<td>77.53 (13.24)</td>
<td>53.50 (12.43)</td>
</tr>
<tr>
<td>Shipley vocabulary</td>
<td>32.20 (6.72)</td>
<td>35.23 (6.21)</td>
</tr>
</tbody>
</table>

Notes. *Significant difference between young and older adults, $p = .03$. **Significant difference between young and older adults, $p < .001$.

Results

Hits.—Using a factorial ANOVA with age (young/older) as the between-subject variable and rating (low/middle/high) and reference condition (self/common) as the within-subject variables, younger adults ($M = .80$) exhibited higher hit rates than older adults ($M = .69$; $F(1, 61) = 13.67, p < .001, \eta^2 = .18$), and the main effect of reference condition was significant, $F(1, 61) = 18.62, p < .001, \eta^2 = .23$, with self-referenced words leading to higher hit rates ($M = .78$) than words in the common condition ($M = .71$). There was a significant interaction between age and reference condition, $F(1, 61) = 5.61, p = .02, \eta^2 = .08$. As shown in Figure 2a, there was a smaller difference in hit rates between older and younger adults in the self condition than in the common condition. The main effect of rating across both age groups was significant as well, $F(2, 122) = 9.79, p < .001, \eta^2 = .14$, with contrasts indicating that words rated as high ($M = .79$) led to higher hit rates than those rated as low ($M = .72$; $F(1, 61) = 15.63, p < .001, \eta^2 = .97$) or middle levels ($M = .73$; $F(1, 61) = 12.82, p < .001, \eta^2 = .94$). The interaction between rating and age was also significant, $F(2, 122) = 3.32, p = .04, \eta^2 = .05$. Older adults had significantly higher hit rates for words rated as high ($M = .76$) than those rated at low ($M = .65$; $F(1, 33) = 15.75, p < .001, \eta^2 = .32$) and middle ($M = .67$; $F(1, 33) = 11.73, p = .002, \eta^2 = .26$) levels. For younger adults, the rating levels did not significantly differ ($ps > .10$). To test whether self-reference uniquely benefitted memory beyond the effects of item endorsement, we found the critical interaction between reference and rating reached significance, $F(2, 122) = 5.10, p = .01, \eta^2 = .08$. When this was further broken down using contrasts to examine performance across ratings for self and common separately, each of the ratings levels significantly differed in the self condition, such that high ($M = .84$) led to higher hit rates than middle ($M = .78$), which was higher than low ($M = .73$, $ps < .02$). However, in the common condition, only the high ($M = .74$) and middle ($M = .68$) rating levels significantly differed, $F(1, 62) = 8.04, p = .01, \eta^2 = .12$. See Figure 2a.

False alarms.—The role of self-referencing on false alarms was tested with a factorial ANOVA with age (young/older) as the between-subject factor and rating (low/middle/high) and reference (self/common) as the within-subject factors. The main effect of age on false alarm rate was significant, $F(1, 50) = 4.53, p = .04, \eta^2 = .08$, with older adults...
(M = .33) committing more false alarms than younger adults (M = .24). The main effect of rating was significant, F(2, 100) = 22.05, p < .001, η² = .31, with words in the high condition leading to more errors (M = .35) than words in the low (M = .24; F(1, 46) = 49.43, p < .001, η² = .52) or middle (M = .26; F(1, 46) = 32.25, p < .001, η² = .44) rating levels. The interaction between age and rating was also significant, F(2, 100) = 5.85, p = .004, η² = .11. For older adults, the high rating level (M = .42) significantly differed from low (M = .29; F(1, 24) = 17.73, p < .001, η² = .43) and middle (M = .27; F(1, 24) = 30.28, p < .001, η² = .56) levels. For younger adults, low (M = .19) significantly differs from both the high (M = .29; F(1, 26) = 14.26, p = .001, η² = .35) and middle (M = .25; F(1, 26) = 10.67, p = .003, η² = .29) rating groups. The interaction between reference and rating was marginally significant, F(2, 100) = 2.84, p = .06, η² = .05. However, there was a significant three-way interaction between age, reference, and rating, F(2, 100) = 3.08, p = .05, η² = .06. As shown in Figure 2b, older adults had more false alarms in the high self condition (M = .47) compared with the high common condition (M = .38; F(1, 24) = 9.95, p = .004, η² = .29), whereas this was not true for younger adults, p > .10.

Corrected recognition.—A factorial ANOVA with age (young/older) as the between-subject variable and rating level (high/middle/low) and reference (self/common) as the within-subject variables revealed a main effect of reference, F(1, 49) = 14.22, p < .001, η² = .23, and a main effect of age, F(1, 58) = 19.51, p < .001, η² = .29. Higher memory scores occurred for self (M = .51) relative to common (M = .42) and younger (M = .57) relative to older adults (M = .37). The interaction between age and reference was only marginally significant, F(1, 49) = 3.44, p = .07, η² = .07, and showed a pattern similar to that found in hits. Rating interacted with age, F(2, 98) = 5.24, p = .007, η² = .10. Younger adults displayed better corrected memory scores for the low condition (M = .62) than for the middle (M = .54; F(1, 25) = 13.61, p = .001, η² = .35) and high (M = .55; F(1, 25) = 5.67, p = .03, η² = .19) conditions, whereas older adults’ corrected recognition scores vary little across the rating levels, ps > .05. Finally, there was a significant three-way interaction between age, rating, and reference, F(2, 98) = 3.80, p = .03, η² = .07. A contrast of self versus common shows that although older adults receive a greater benefit from self than common across all three rating levels, F(1, 24) = 10.84, p = .003, η² = .31, the high (M = .37) self condition leads to worse overall memory compared with the low (M = .45) and middle levels (M = .49; F(1, 24) = 8.94, p = .006, η² = .27). Contrasts show that there is no significant difference for older adults across the three ratings in the common condition, ps > .10.

Response bias.—To assess response bias, C values were used in a factorial ANOVA with age (young/older) as the between-subject variable and rating level (high/middle/low) and reference (self/common) as the within-subject variables. There was a main effect of rating, F(2, 96) = 19.49, p < .001, η² = .29, such that words rated in the high condition had a more negative bias (i.e., tendency to respond “old”; M = -.31) than the middle (M = .01) and low (M = .07) conditions. There was also a main effect

### Table 4. Means (SD) in Each Condition for Study 2

<table>
<thead>
<tr>
<th>Study 2</th>
<th>Young adults</th>
<th></th>
<th></th>
<th>Older adults</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>Low</td>
<td>Middle</td>
<td>High</td>
<td>Low</td>
<td>Middle</td>
<td>High</td>
</tr>
<tr>
<td>Proportion of items in category</td>
<td>.34 (.06)</td>
<td>.34 (.12)</td>
<td>.31 (.10)</td>
<td>.36 (.10)</td>
<td>.30 (.15)</td>
<td>.32 (.12)</td>
</tr>
<tr>
<td>Hit</td>
<td>.77 (.16)</td>
<td>.82 (.14)</td>
<td>.86 (.14)</td>
<td>.68 (.13)</td>
<td>.75 (.15)</td>
<td>.82 (.13)</td>
</tr>
<tr>
<td>Miss</td>
<td>.23 (.16)</td>
<td>.18 (.14)</td>
<td>.14 (.14)</td>
<td>.31 (.13)</td>
<td>.23 (.15)</td>
<td>.18 (.13)</td>
</tr>
<tr>
<td>CR</td>
<td>.20 (.16)</td>
<td>.23 (.18)</td>
<td>.29 (.15)</td>
<td>.25 (.15)</td>
<td>.26 (.19)</td>
<td>.47 (.19)</td>
</tr>
<tr>
<td>FA</td>
<td>.80 (.16)</td>
<td>.77 (.18)</td>
<td>.71 (.15)</td>
<td>.75 (.15)</td>
<td>.74 (.19)</td>
<td>.53 (.19)</td>
</tr>
<tr>
<td>Hit-FA</td>
<td>.58 (.27)</td>
<td>.60 (.21)</td>
<td>.60 (.21)</td>
<td>.44 (.15)</td>
<td>.49 (.25)</td>
<td>.36 (.22)</td>
</tr>
<tr>
<td>Response bias C values</td>
<td>.02 (.45)</td>
<td>-.09 (.43)</td>
<td>-.41 (.45)</td>
<td>.08 (.40)</td>
<td>-.09 (.57)</td>
<td>-.57 (.47)</td>
</tr>
<tr>
<td>Common</td>
<td>Low</td>
<td>Middle</td>
<td>High</td>
<td>Low</td>
<td>Middle</td>
<td>High</td>
</tr>
<tr>
<td>Proportion of items in category</td>
<td>.19 (.11)</td>
<td>.38 (.15)</td>
<td>.42 (.16)</td>
<td>.15 (.13)</td>
<td>.35 (.19)</td>
<td>.49 (.26)</td>
</tr>
<tr>
<td>Hit</td>
<td>.82 (.16)</td>
<td>.75 (.13)</td>
<td>.78 (.15)</td>
<td>.61 (.28)</td>
<td>.60 (.21)</td>
<td>.69 (.20)</td>
</tr>
<tr>
<td>Miss</td>
<td>.18 (.16)</td>
<td>.25 (.13)</td>
<td>.22 (.15)</td>
<td>.39 (.28)</td>
<td>.40 (.21)</td>
<td>.31 (.20)</td>
</tr>
<tr>
<td>FA</td>
<td>.18 (.19)</td>
<td>.27 (.18)</td>
<td>.28 (.15)</td>
<td>.33 (.29)</td>
<td>.27 (.16)</td>
<td>.38 (.21)</td>
</tr>
<tr>
<td>CR</td>
<td>.82 (.19)</td>
<td>.73 (.18)</td>
<td>.72 (.15)</td>
<td>.67 (.29)</td>
<td>.73 (.16)</td>
<td>.62 (.21)</td>
</tr>
<tr>
<td>Hit-FA</td>
<td>.65 (.20)</td>
<td>.48 (.18)</td>
<td>.51 (.23)</td>
<td>.27 (.33)</td>
<td>.32 (.24)</td>
<td>.31 (.19)</td>
</tr>
<tr>
<td>Response bias C values</td>
<td>-.02 (.76)</td>
<td>.04 (.46)</td>
<td>-.11 (.41)</td>
<td>.17 (.89)</td>
<td>.20 (.41)</td>
<td>-.13 (.69)</td>
</tr>
</tbody>
</table>

**Notes.** Bolded proportions correspond to analyses presented in text; CR = corrected recognition; FA = false alarms.
FALSE MEMORY AND SELF-REFERENCE

of reference, $F(1, 48) = 10.87$, $p = .002$, $\eta^2_p = .19$, with words in the self condition leading to a more negative bias ($M = −.17$) than those in the common condition ($M = .03$). The interaction between reference and rating was significant as well, $F(2, 96) = 4.33$, $p = .02$, $\eta^2_p = .08$. Within the self condition, contrasts indicated a stronger negative bias in the high condition ($M = −.49$) than in the low ($M = .05$; $F(1, 48) = 118.80$, $p < .001$, $\eta^2_p = .71$) and middle conditions.
(M = −.09; F(1, 48) = 31.23, p < .001, η² = .39). However, in the common condition, the only significant difference was between the more positive bias for middle (M = .12) than high conditions (M = −.12; F(1, 48) = 13.26, p = .001, η² = .22). Furthermore, there was a significant difference between self and common in the high, F(1, 48) = 24.08, p < .001, η² = .35, and middle conditions, F(1, 48) = 7.04, p = .01, η² = .13.

Discussion

Experiment 2 succeeded in distinguishing self-reference effects from overall effects of ratings on memory. Self-referencing led to higher hit rates compared with commonness judgments, and elevated hit rates for older adults such that they performed more like younger adults than in the common condition. This finding is consistent with previous research demonstrating that older adults benefit from self-referencing (Glisky & Marquine, 2009; Gutchess et al., 2007; Gutchess, Kensinger, Yoon, et al., 2007). Hit rates were elevated for highly self-descriptive words, relative to less self-descriptive words; this pattern was unique to self-referencing and did not characterize the common condition. Therefore, it appears as though the elevated hit rates seen in the self condition are not simply the result of endorsement but rather are due to the connection with the self. Although we expected to see higher hit rates for words in the low common condition (Glanzer & Adams, 1985; Reder et al., 2002), these did not differ from the high common conditions.

As in Experiment 1, older adults continued to experience higher rates of false alarms compared with younger adults; this was particularly true for highly self-descriptive words. In the critical comparison against the high common condition, older adults committed more false alarms for highly self-descriptive words. Self-reference selectively increased false alarms, especially for older adults when information was highly self-relevant. It is possible that the increased fluency-based familiarity (Jacoby & Whitehouse, 1989; Thapar & Westerman, 2009) and association (Roediger & McDermott, 1995) with the self led to the higher false alarm rates in the high self condition compared with the high common condition.

In terms of corrected recognition scores, overall memory for self-referenced words was better than common words for both younger and older adults, consistent with previous studies of self-reference compared with control conditions (Symons & Johnson, 1997). However, memory is poorer for highly self-descriptive words, particularly for older adults. Although there was more of a response bias to endorse highly self-descriptive words as “old,” unlike Experiment 1, this was true for both younger and older adults. Therefore, although the response bias may have contributed to higher false alarm rates, the bias was similar for younger and older adults so that this alone does not account for age differences in false memories. It appears as though older adults simply commit more false alarms for highly self-relevant information.

General Discussion

These studies demonstrate that although self-referencing can benefit memory, by elevating hit rates as words are rated as more self-descriptive (Experiments 1 & 2) and corrected recognition scores when words are rated as highly self-descriptive compared with less self-descriptive. The strong connection to the self may lead to fluent processing and feelings of familiarity that both help and hurt memory performance. By contrasting self-referencing against commonness judgments (Experiment 2), the results indicate that it is the connection to the self, and not simply endorsement of a word, that leads to heightened memory errors.

As explained by the implicit associative model of memory (Roediger & McDermott, 1995), the deeper processing caused by self-referencing may have triggered greater associations for self-referenced information that can lead to greater false memory of these associated items. Alternatively, because the self-reference effect increases the level of familiarity, this could also elevate intrusions (Jacoby & Whitehouse, 1989; Thapar & Westerman, 2009). Older adults could experience more difficulty determining the basis of the fluency (e.g., internally generated or externally presented; Dehon & Brédart, 2004; Ferguson et al., 1992; Henkel et al., 1998). Older adults experienced more false memories than younger adults and this difference was magnified as self-relevance increased. Based on the present study, it appears as though self-reference does provide a benefit in memory relative to other conditions; but under some conditions, this benefit may begin to diminish as information becomes increasingly self-relevant.

Although the present study links the self-referencing and false memory literatures in the study of memory in aging, it is limited in that the design may not be readily applicable to real-world situations. It is important to assess the relevance of these findings in richer everyday contexts (e.g., Rosa & Gutchess, 2011) because older adults’ susceptibility to false memory makes them vulnerable to scams. Despite establishing a link between the self and false memory in the present study, future efforts should be aimed at the identification of techniques to reduce the incidence of false
memory from self-referencing, such as strategies that help in differentiating external sources from internal intrusions. Overall, the results of this study suggest that although self-reference improves memory for young and older adults (Glusky & Marquine, 2009; Gutchess et al., 2007; Hamami et al., 2011), it affects response bias, which also increases false memory. This is especially true for older adults. The findings indicate that false memory is an important factor to consider for self-referencing. Although older adults are likely to receive a benefit for information related to the self, they are also likely to experience increased false memory, and our results indicate that this can occur at strikingly high levels.

Funding
This work was supported by the National Institute on Aging (R21 AG032382 to A. H. Gutchess) and the Lifespan Initiative on Healthy Aging (LIHA) at Brandeis University (to N. M. Rosa). Portions of this research were conducted while A. H. Gutchess was a fellow of the American Federation for Aging Research.

Acknowledgments
The authors thank Liat Zabludovsky, Pete Millar, and Shanny Shmuel for their research assistance and acknowledge the helpful comments and suggestions made by Dr. Leslie Zebrowitz, Dr. Elizabeth Kensinger, Dr. Eric Leschikar, and Margeaux Aulner.

Correspondence
Correspondence should be addressed to Nicole M. Rosa, MA, MSW, Department of Psychology, Brandeis University, 415 South Street, MS 062, Waltham, MA 02454-9110. E-mail: nrosa@brandeis.edu.

References

FALSE MEMORY AND SELF-REFERENCE


