Running head: PROCESSING DEPTH AND FALSE RECOGNITION IN STM AND LTM

Get the Gist? The Effects of Processing Depth on False Recognition

in Short-Term and Long-Term Memory

Kristin E. Flegal[1] and Patricia A. Reuter-Lorenz[2]

[1] Center for Neuroscience, University of California, Davis, CA

[2] Department of Psychology, University of Michigan, Ann Arbor, MI

Abstract

Gist-based processing has been proposed to account for robust false memories in the converging associates task. The deep encoding processes known to enhance verbatim memory also strengthen gist memory and increase distortions of long-term memory (LTM). Recent research demonstrates that compelling false memory illusions are relatively delay-invariant, also occurring under canonical short-term memory (STM) conditions. To investigate the contribution of gist to false memory at short and long delays, processing depth was manipulated as participants encoded lists of four semantically-related words and were probed immediately following a filled 3-4 second retention interval or approximately 20 minutes later in a surprise recognition test. In two experiments, the encoding manipulation dissociated STM and LTM on the frequency, but not the phenomenology, of false memory. Deep encoding at STM increases false recognition rates at LTM, but confidence ratings and Remember/Know judgments are similar across delay and do not differ as a function of processing depth. These results suggest that some shared and some unique processes underlie false memory illusions at short and long delays.

Introduction

The susceptibility of memory to distortion has been empirically studied using the converging associates or DRM paradigm (Deese, 1959; Roediger & McDermott, 1995), in which unstudied semantic associates ("related lures") to lists of studied words are later endorsed as studied items. The ease with which false memories can be reliably induced using the DRM procedure is impressive, but it also invites the practical question of how the incidence of such memory errors can be controlled. Previous DRM studies have tested the effects of varying word list construction (McDermott, 1996; Robinson & Roediger, 1997), encoding instructions (Rhodes & Anastasi, 2000; Thapar & McDermott, 2001; Toglia, Neuschatz, & Goodwin, 1999), testing conditions (Coane & McBride, 2006; McDermott, 1996), and even providing explicit warnings about the nature of the word lists and related lures (Anastasi, Rhodes, & Burns, 2000; Gallo, Roberts, & Seamon, 1997; McDermott & Roediger, 1998; Neuschatz, Payne, Lampinen, & Toglia, 2001). The persistent finding is that while false memory rates can be attenuated through such manipulations, they are rarely if ever eliminated.

Recent research from our lab has further demonstrated the recalcitrance of the false memory illusion by documenting its occurrence within several seconds of studying brief four-item lists (Atkins & Reuter-Lorenz, 2008; Flegal, Atkins, & Reuter-Lorenz, 2010; see also Coane, McBride, Raulerson, & Jordan, 2007). Flegal et al. (2010) used a new experimental paradigm to directly compare false memories in the same participants under short-term memory (STM) and long-term memory (LTM) conditions. Participants viewed lists of four semantic associates and were probed immediately following a filled 3-4 second retention interval (i.e., STM) or approximately 20 minutes later in a surprise

recognition test (i.e., LTM). A unique advantage to this task design is its ability to dissociate retrieval processes operating under STM and LTM conditions, for items studied in the same encoding period under equivalent encoding conditions.

Based on traditional models postulating separable memory systems, Flegal et al. (2010) expected that false recognition rates would be higher under canonical LTM conditions than at STM testing. In particular, the subspan memory load, brief retention interval, and surface-level coding associated with STM were expected to minimize false recognition. In contrast, the semantic, meaning-based coding thought to characterize LTM, coupled with the greater memory load and longer delay, were expected to maximize false recognition. The results defied these expectations, however. Measurements of both quantity (the relative incidence of false recognition) and quality (the accompanying phenomenology, or subjective experience) of memory distortions did not reliably differ between short-term and long-term testing. The finding of stable false memory effects across time is inconsistent with predictions based on models of multiple memory systems, suggesting instead that the generation of false memories is due to the operation of the same processes regardless of delay.

These results pose a challenge to at least one prominent theoretical account of false memory. "Fuzzy trace" theory (Brainerd & Reyna, 2002; Reyna & Brainerd, 1995) proposes that two types of information are encoded into memory in parallel: verbatim, item-specific traces which faithfully record surface features and details of an experience, and gist representations which store general themes or semantic content. For accurate memory of semantically-related material, verbatim and gist traces reinforce each other, but gist traces also promote false memory for related information, especially when

unopposed by shorter-lasting verbatim traces. According to this theory, access to both types of traces declines over time, but gist traces are slower to fade. Thus, verbatim traces should be readily available to oppose gist traces during the STM test, thereby minimizing the incidence of false memories. At longer delays, the weakened verbatim traces should leave the more durable gist-based traces unopposed, thereby increasing the incidence of false memories. Yet Flegal et al. (2010) observed false memory effects that were relatively delay-invariant.

In our previous study (Flegal et al., 2010), no explicit encoding instructions were given, so the strategies participants adopted were likely aimed merely at supporting short-term recognition because the LTM test was unbeknownst to them. Consequently, gist-based coding may have been weak and merely incidental, which would also explain the relatively low memory performance observed at long-term testing. A more rigorous test of the fuzzy trace theory in the context of our hybrid DRM task can be rendered with more systematic control of encoding strategies. This is the aim of the present study.

According to the levels-of-processing theory (Craik & Lockhart, 1972), information encoded at a *deep*, semantic level will be better remembered than information encoded at a *shallow*, perceptual level. However, prior research using the DRM paradigm has shown that deep processing of semantic associates (e.g., making pleasantness ratings, or concrete/abstract decisions) increases the rate of false recall (Rhodes & Anastasi, 2000; Toglia et al., 1999) and false recognition (Thapar & McDermott, 2001) of related lure words, and also increases accurate memory for studied items. Fuzzy trace theory claims that deep, meaning-based encoding strengthens durable gist traces, accounting for the resultant increases in both accurate

memory and false memory observed in DRM studies. In the present study, we further test predictions from fuzzy trace theory by assessing the effects of encoding manipulations on false memories at short as well as long delays. At short-term testing, even if gist traces are better established with deep encoding, coexisting verbatim traces should still be strong and oppose false memory. At long-term testing, verbatim traces will be relatively weak, leading to minimal opposition and instead increased reliance on gist, and a corresponding increase in false memory. An additional prediction is that the effects of encoding manipulations on accurate memory will be minimal at STM, but the codes established at this same encoding episode will show a dramatic effect on accuracy when later probed in LTM, with deep encoding potentially bringing LTM accuracy levels close to what is seen when probing at STM.

The prediction that an encoding manipulation will dissociate memory accuracy between short-term and long-term testing has support from recent studies using a novel "levels-of-processing span" paradigm. Rose et al. (2010) found that visual, phonological, or semantic processing of to-be-remembered words did not influence immediate recall accuracy (i.e., STM), but a classic levels-of-processing effect emerged on a delayed recognition test (i.e., LTM), where accuracy was highest for words that received deep encoding, and lowest for words that received shallow encoding. A follow-up study (Rose & Craik, 2012) replicated the LTM results and showed that a levels-of-processing effect at short-term testing could be induced only when an immediate recall test was unexpected, in the context of incidental encoding that discouraged the surface-level coding and active maintenance processes characteristic of STM. Taken together, these data challenge strict models of unitary memory systems, by demonstrating the

involvement of different processes under STM and LTM conditions. However, while the work of Rose and colleagues is informative about effects of processing depth on *accurate* memory at short and long delays, it is unclear whether *false* memory would be similarly influenced. In the present study, we re-examine the effects of encoding manipulations on false memory focusing specifically on whether they will operate similarly in STM and LTM, or if their effects will dissociate these systems. In addition to assessing memory accuracy, we include phenomenological measurements to acquire converging evidence about reliance on verbatim versus gist memory.

The phenomenology of memory distortions is evident in compelling, but illusory, feelings of recollection often found to accompany false memories for semantically-related lure words, at long delays in typical DRM tasks (Lampinen, Neuschatz, & Payne, 1998; Lampinen, Ryals, & Smith, 2008; Payne, Elie, Blackwell, & Neuschatz, 1996). In our previous study (Flegal et al., 2010), surprisingly, such "illusory recollection" effects were observed at short delays as well: measures of both confidence and recollective experience for falsely recognized lure words were statistically equivalent under STM and LTM conditions. Gist-based processing offers the best account for the stable rates of false memory previously demonstrated in our paradigm under unconstrained encoding instructions, although high-confidence and recollection-based responses would be expected to decline over time along with the availability of verbatim traces. Here our goal is to further interrogate the subjective experience of false recognition by experimentally controlling processing depth. If deep encoding promotes gist-based responding at long-term testing, then decreased access to verbatim detail should be

apparent in confidence ratings and Remember/Know judgments associated with false memory illusions.

Experiment 1

Method

Participants

32 individuals (18-24 yrs old; $M = 19.8$) participated for course credit or payment. Eight additional participants were tested but excluded from analysis for recognition accuracy scores > 2.5 standard deviations from the mean at STM and/or LTM, math task accuracy < 0.70 (during the STM trial retention interval), or computer malfunction. Research protocols were approved by the University of Michigan Institutional Review Board, and all participants provided written informed consent.

Materials

The memory sets were 96 lists of four semantically related words, all associates of a common theme word (e.g., SLEEP for the list containing the associates nap, doze, bed, awake) which served as the probe on every trial. As depicted in Figure 1, the three probe types were: *related lure*, the unstudied theme word associated with a studied list; *unrelated lure*, an unstudied theme word associated with a nonpresented list; and *target*, the theme word associated with, and present in, a studied list (replacing one of the four associates). Theme words (e.g., sleep) were inserted into studied lists on target probe trials, rather than using a studied associate (e.g., bed) as the probe, because theme words in the DRM paradigm are unique in that the other memory set items converge upon them. The ordinal position of the item replaced by the theme word was

balanced across target probe trials. Theme words as probes appeared only once during the experiment; no list was probed in both STM and LTM trials.

The 96 lists were divided into four groups of 24 four-word lists (Groups A-D) equated in mean backward associative strength ($M = 0.40$) derived from the University of South Florida Free Association Norms (Nelson, McEvoy, & Schreiber, 1998). Probe type was counterbalanced with word lists across participants, so that for one quarter of all participants, lists in Group A were paired with related lure probes, B with unrelated lure probes, and C with target probes. Theme words associated with Group D lists served as the unrelated lure probes. Because the 24 lists in Group D were not presented, this resulted in a total of 72 STM trials (see Design and Procedure, below).

Each of the four groups of 24 lists was further divided into two subgroups of 12 lists, following the same parameters, to balance the status of each list as short-term versus long-term memoranda across participants. Thus, for half of the participants in each of the four counterbalanced orders, the first subgroup of lists from each group (e.g., A1) was probed during the STM trials, and the second subgroup of lists (e.g., A2) during the LTM trials; the assignment was reversed for the other half of the participants. This procedure ensured that all participants encountered the same probes – all theme words, but in different contexts – as related lures, unrelated lure probes, or target probes, and as STM or LTM probes (see also Flegal, et al., 2010).

Finally, the blocked order of encoding instructions was counterbalanced across participants, so that half of the participants in each of the orders experienced an ABBA design beginning with "shallow" encoding instructions, and the other half experienced a BAAB design beginning with "deep" encoding instructions.

Design and procedure

The four-word memory sets were probed either within the same trial (i.e., STM) or in a surprise recognition test following completion of all STM trials (i.e., LTM), in order to examine short-term and long-term memory distortions concurrently and within-subjects (see Figure 1). As in our previous study (Flegal et al., 2010), each STM trial started with a brief study period followed by a 3,000-4,000 ms retention interval filled with a math equation verification task, then (a) the probe word to respond to (*yes* or *no*) followed by a four-point confidence rating (*very low*, *somewhat low*, *somewhat high*, or *very high*), or (b) cues for two arbitrary button presses with corresponding response mappings (for memory sets that would subsequently be probed at LTM). Instead of simultaneous presentation of the entire memory set at the beginning of each STM trial, however, here the four words were presented serially (for 250 ms each, with a 2,000 ms inter-stimulus interval) to allow for individual encoding judgments.

Shallow encoding instructions directed participants to "count the total number of ascenders and descenders in each word," thereby requiring attention to the morphological characteristics of the memory set items (always presented in lower-case letters). Participants made a four-point response to each of the four words on these trials, indicating the sum of its ascenders and descenders: 0, 1, 2, or 3+. Ascenders (b, d, f, h, k, l, and t) and descenders (g, j, p, q, y) were defined for participants in a practice session prior to the start of the experiment.

Deep encoding instructions directed participants to "rate how much you like the meaning of each word," thereby requiring attention to the semantic characteristics of the

memory set items. Participants made a four-point liking response to each of the four words on these trials: *dislike a lot*, *dislike a little*, *like a little*, or *like a lot*.

The STM trials consisted of 72 trials divided into four blocks, two for each set of encoding instructions. Instructions for both encoding conditions were introduced before starting the first STM block. Half of the participants began with a shallow encoding block and the other half began with deep encoding block. All participants received a 1-min break between the second and third blocks. Each block of trials started with a reminder screen describing the encoding instructions and response options, and an identifying colored border (shallow = blue; deep = pink) remained on the screen throughout the block. Participants completed 18 trials in each of the 4 blocks, presented in random order. Of the 9 trials in each block probed at STM, 3 were of each probe type (related lure, unrelated lure, and target). Thus, a total of 36 memory sets were probed at STM.

A 2-min break followed completion of the STM trials, then participants were informed about, and given instructions for, the LTM recognition test. Each participant completed 72 LTM trials, 36 of which tested memory sets that were not probed at STM (12 each – 3 from each STM block – of the related lure, unrelated lure, and target probe type). Additionally, to match the proportion of correct "yes" responses between STM and LTM testing, there were 12 trials of studied associates from memory sets that were probed at STM (never including theme words from target probe trials), and 24 trials of unstudied, unrelated foils, matched for frequency and word length with the corpus of theme words used in the experiment. Each LTM trial started with the probe word to respond to (by indicating whether or not it had appeared during the STM trials: *yes* or *no*) followed by a four-point confidence rating.

Results

Encoding time

The mean response latency to memory set items in the STM trials (i.e., the amount of time to make encoding judgments) was significantly longer in the shallow encoding condition (M = 1,001 ms, SEM = 26 ms) than in the deep encoding condition (M = 809 ms, SEM = 26 ms), $t(31)$ = 5.69, p < 0.001. Although this may indicate systematic differences in the average amount of encoding time, it should be noted that the memory set items were present on the computer screen for the same duration (250 ms each) in each condition. The latency difference is also in the opposite direction from that which might predict better memory associated with deeper encoding and longer encoding time: response times were actually faster for liking ratings (deep processing) than letter counting (shallow processing). Nevertheless, as the data below indicate, deeper processing led to better memory, suggesting that the perceptual demands of the shallow encoding task simply made it more time-consuming.

Accuracy

Mean math task accuracy during the STM trial retention interval was 0.89; it did not significantly differ between shallow encoding (M = 0.89) and deep encoding (M = 0.88) blocks. As shown in Table 1, a classic levels-of-processing effect was evident for studied items, as correct recognition of target probes in the deep encoding condition remained nearly as high on LTM trials as on STM trials, compared to a steep drop in target probe accuracy from STM to LTM in the shallow encoding condition. Paired comparisons confirmed that the encoding manipulation reliably influenced accurate

memory at long-term testing ($t(31) = 6.12$, $p < .001$) but not short-term testing ($t(31) =$ 1.55, $p = .13$) where hit rates were high in both conditions. The encoding condition (shallow, deep) $\times$ delay (STM, LTM) interaction for target probe accuracy was highly significant, $F(1,31) = 21.40$, $p < .001$, $\eta_p^2 = .41$.

At short-term testing, there was a significant false memory effect (the difference between "yes" responses to unstudied, related lure probes and unstudied, unrelated lure probes) for both shallow encoding ($M = 0.16$) and deep encoding ($M = 0.11$). At long-term testing, however, the false memory effect no longer reliably differed from zero for lists that received shallow encoding ($M = 0.04$), while the false memory effect for lists that received deep encoding was more than double in size at LTM ($M = 0.27$) compared to STM. Thus, although the incidence of false recognition did not differ between encoding conditions at STM ($t(31) = 0.98$, $p = .33$), it was significantly greater in the deep encoding than the shallow encoding condition at LTM ($t(31) = 2.96$, $p < .01$). The baseline (unrelated lure probe) false alarm rate at LTM was equivalent for shallow and deep encoding conditions (see Table 1), indicating that the significant difference in the false memory effect at LTM was driven by a disproportionate increase in false alarms to related lure probes belonging to lists that received deep encoding at STM. This time-dependent effect of levels of processing on false memory formation was evident in a significant lure type (related, unrelated) $\times$ encoding condition (shallow, deep) $\times$ delay (STM, LTM) interaction, $F(1,31) = 10.16$, $p < .01$, $\eta_p^2 = .25$.

Phenomenological experience

Stable levels of confidence in false recognition across time were observed in both encoding conditions. As shown in Figure 2, confidence in "yes" responses to

related lures did not change from STM to LTM for lists that received shallow encoding ($t(13) = 1.38$, $p = .19$) or lists that received deep encoding ($t(14) = 0.49$, $p = .63$). This finding replicates and extends results from our previous study under unconstrained encoding instructions (Flegal et al., 2010). In contrast, but also consistent with our earlier work, confidence in "no" (correct) responses to related lures significantly decreased from short-term to long-term testing in both the shallow ($t(30) = 3.43$, $p < .01$) and deep ($t(30) = 3.37$, $p < .01$) encoding conditions. The response (*yes, no*) × delay (STM, LTM) interaction for related lure confidence ratings was marginally significant in both the shallow ($F(1,12) = 3.92$, $p = .07$) and deep ($F(1,14) = 4.26$, $p = .06$) encoding conditions. Thus, the effects of levels of processing did not dissociate STM and LTM on participants' confidence in their responses to related lure probes. For all correct responses, however, there was a significant main effect of encoding condition ($F(1,29) = 33.87$, $p < .001$), such that confidence ratings were higher overall for lists that were studied under deep than under shallow encoding instructions.

Discussion

In summary, Experiment 1 revealed that the effect of delay on the incidence of false recognition depends on encoding strategy. A direct comparison of short-term and long-term testing indicates that deep encoding preserved accurate memory across delay and increased false memory at LTM (consistent with results from DRM studies; e.g., Thapar & McDermott, 2001), while the encoding strategy manipulation had minimal effects at STM (consistent with results from the levels-of-processing span task; Rose, et al., 2010). Confidence ratings for false recognition of related lures were statistically

equivalent at short and long delays (see also Flegal, et al., 2010), and did not differ as a function of encoding strategy. In other words, false memories were endorsed with similar levels of confidence whether they occurred after short or long delays, and regardless of how deeply their associates were encoded.

Because evidence suggests that deep processing strengthens both verbatim and gist memory traces, independent contributions of the two types of memorial information are difficult to discern in the data from Experiment 1. While verbatim traces fortified by deep encoding might be expected to oppose the illusion of false memories at short-term testing, it is possible that confidence ratings are not sufficiently sensitive to such differences (Lampinen et al., 1998). In an effort to better capture differences in false memory phenomenology than the confidence ratings used in Experiment 1, Remember/Know judgments (Gardiner, 1988; Tulving, 1985; Yonelinas, 2002) were employed in Experiment 2. Recollection of vivid, specific details of a past experience is considered to be the hallmark of "remember" judgments, made on the basis of verbatim memory traces, while "know" judgments are associated with feelings of familiarity that lack episodic detail, and may arise from gist memory instead (Geraci & McCabe, 2006). In the DRM paradigm, manipulations designed to increase semantic processing (e.g., presenting word lists blocked by theme rather than randomly intermixed) lead to increases in measures of false memory phenomenology (Dewhurst, Bould, Knott, & Thorley, 2009; Mather, Henkel, & Johnson, 1997). Deep, semantic processing is also known to promote recollection-based responses in accurate memory (Gardiner, 1988; Rajaram, 1993). If processing depth can reveal delay-dependent effects on the incidence of false memory, as shown in Experiment 1, then we would expect a similar

dissociation for rates of "illusory recollection", in that gist-based responding in the LTM recognition phase should result in fewer *remember* responses to related lures due to a greater likelihood of *know* responses.

In Experiment 1, rates of accurate memory in the deep encoding condition were nearly equivalent at short-term and long-term testing, which fuzzy trace theory would attribute to verbatim traces at STM and gist traces in the absence of verbatim traces at LTM. If gist memory contributes more to retrieval decisions at longer delays for lists that receive deep encoding, then we would expect estimates of false recollection to be greater at LTM than STM. Our previous study (Flegal et al., 2010) found that when encoding strategies were unconstrained, the proportion of *remember* responses assigned to related lure false alarms was relatively delay-invariant, and the present study leverages the encoding strategy manipulation to more closely examine predicted contributions of enduring gist traces to false recognition over time.

Experiment 2

Method

Participants

36 individuals (18-28 yrs old; $M = 19.4$) participated for course credit or payment. Seven additional participants were tested but excluded from analysis for recognition accuracy scores > 2.5 standard deviations from the mean at STM and/or LTM, math task accuracy < 0.70 (during the STM trial retention interval), or post-experiment questionnaire responses indicating failure to understand the Remember/Know distinction.

<u>Design and procedure</u>

The method was the same as Experiment 1 except that remember/know/guess judgments replaced confidence ratings. Following each "yes" response to a probe word, participants were prompted to indicate with a button press whether they *remember* the probe word was in the memory set (recollecting something distinctive about studying the word), they *know* the probe word was present (recognizing the word without retrieving specific details of its study), or their response had been a *guess*. As in our previous study (Flegal et al., 2010), detailed instructions explaining the Remember/Know distinction were adapted from Rajaram (1993). To equate the number of responses required on each trial, a display of three boxes appeared following each "no" response to a probe word, prompting an arbitrary button press response. Additionally, participants made three-point (letters count: 0, 1, or 2+; liking rating: *dislike*, *like a little*, or *like a lot*), rather than four-point, responses for individual encoding judgments.

Furthermore, to minimize the influence of knowledge of multiple encoding strategies from the start of the experiment, the blocked order of encoding instructions was counterbalanced across participants so that instructions for the first half of STM trials were given before the first block, and instructions for the second half of STM trials were given between the second and third blocks. Half of the participants in each of the orders experienced an AABB design beginning with shallow encoding instructions, and the other half experienced a BBAA design beginning with deep encoding instructions. All participants completed a total of four STM blocks, two with each set of encoding instructions.

Results

Encoding time

 The mean response latency to memory set items in the STM trials was slightly but significantly longer in the shallow encoding condition ($M = 809$ ms, $SEM = 20$ ms) than in the deep encoding condition ($M = 771$ ms, $SEM = 21$ ms), $t(35) = 2.17$, $p < 0.05$, replicating the pattern from Experiment 1.

Accuracy

 Mean math task accuracy during the STM trial retention interval was 0.90; it did not significantly differ between shallow encoding ($M = 0.91$) and deep encoding ($M = 0.90$) blocks. As in Experiment 1, a classic levels-of-processing effect was evident for studied items, as correct recognition of target probes in the deep encoding condition remained nearly as high on LTM trials as on STM trials, compared to a steep drop in target probe accuracy from STM to LTM in the shallow encoding condition (see Table 1). Again, as expected, the encoding manipulation reliably influenced accurate memory at long-term testing ($t(35) = 7.99$, $p < .001$) although here it also had an effect at short-term testing ($t(35) = 2.37$, $p < .05$). The encoding condition (shallow, deep) $\times$ delay (STM, LTM) interaction for target probe accuracy was highly significant, $F(1,35) = 30.55$, $p < .001$, $\eta_p^2 = .47$.

 At short-term testing, there was a significant false memory effect for both shallow encoding ($M = 0.19$) and deep encoding ($M = 0.11$). At long-term testing, the false memory effect was slightly (but not significantly) reduced for lists that received shallow encoding ($M = 0.15$), while the false memory effect for lists that received deep encoding was almost three times larger at LTM ($M = 0.31$) compared to STM.

Although the false memory effect was significantly greater than zero in both

encoding conditions at both delays, the incidence of false recognition at LTM was higher

for lists that were studied under deep than under shallow encoding instructions (t(35) =

2.83, $p < .01$), consistent with the results from Experiment 1 and prior studies from the

DRM literature. In contrast, the incidence of false recognition at STM was greater in the

shallow encoding than the deep encoding condition; an effect in this direction was

shown numerically in Experiment 1, but here the difference was significant (t(35) = 2.66,

$p = .01$). At LTM, the baseline (unrelated lure probe) false alarm rate was equivalent for

shallow and deep encoding conditions (see Table 1), indicating that the significant

difference in the false memory effect at LTM was driven by a disproportionate increase

in false alarms to related lure probes belonging to lists that received deep encoding at

STM. As in Experiment 1, this time-dependent effect of levels of processing on false

memory formation was evident in a significant lure type (related, unrelated) $\times$ encoding

condition (shallow, deep) $\times$ delay (STM, LTM) interaction, F(1,35) = 13.16, $p < .001$, η_p^2

= .27.

Phenomenological experience

Consistent with the stable confidence in related lure false alarms in Experiment

1, normalized estimates of false recollection in both encoding conditions were relatively

time-invariant. Although the absolute incidence of *remember* responses to related lures

actually increased with longer delays, this shift was confounded with a rising baseline

false alarm rate, and thus a normalized incidence measure was calculated as the

proportion of *remember* responses to related lures out of the total proportion of *yes*

responses to related lures (see also Flegal et al., 2010). As shown in Table 2, this

measure did not change from STM to LTM for lists that received shallow encoding ($t(15)$

$= 0.87$, $p = .40$) or lists that received deep encoding ($t(16) = 0.91$, $p = .38$). Likewise,

there was no significant main effect of delay on this measure collapsed across encoding

condition ($F < 1$). Thus, the subjective experience of "remembering" an item that had

never been studied appeared equally robust at short and long delays. Furthermore, and

replicating our earlier work, the normalized incidence of *remember* responses to related

lures was lower than that for *remember* responses to target probes at both STM and

LTM, for lists studied under both shallow and deep encoding instructions (all ps $< .05$).

Thus, in both encoding conditions, participants were able to differentiate between true

and false memories on the basis of recollective experience. For all *remember*

responses to target probes, a putative index of verbatim memory, there was a

significant main effect of encoding condition ($F(1,35) = 47.02$, $p < .001$, $\eta_p^2 = .57$),

indicating that rates of "true" recollection were higher at both delays for studied items

that received deep rather than shallow encoding at STM.

These results replicate findings of stability in the proportion of false memories

associated with "remember" phenomenology from STM to LTM, as reported in Flegal et

al. (2010), and are consistent with similar effects found in confidence ratings from

Experiment 1 in the present study. Nevertheless, the normalized estimates of false

recollection from Experiment 2 should be interpreted cautiously because relatively few

participants assigned *remember* responses to related lure false alarms, especially at

short-term testing, and thus the sample size was restricted to n = 10 for the encoding

condition (shallow, deep) × delay (STM, LTM) overall ANOVA.

Discussion

As in Experiment 1, accuracy data revealed time-dependent effects of processing depth on false memory formation. Deep encoding preserved accurate memory across delay and increased false memory at LTM (consistent with results from DRM studies), although in Experiment 2 the encoding strategy manipulation also affected the incidence of false recognition at STM, in the opposite direction to LTM. A possible reason why deep encoding would reduce false memory at short delays, relative to shallow encoding, is that deeper processing is assumed to strengthen both verbatim and gist traces. Thus, under STM conditions, the verbatim representations enhanced by deep encoding would oppose the simultaneously enhanced gist representations. Under LTM conditions, with decreased access to verbatim traces, the levels-of-processing effect would be reversed as longer-lasting gist traces strengthened by deep encoding lead to an increase in false memory.

Normalized rates of *remember* responses to related lures replicated our earlier findings of relatively delay-invariant measurements of false memory phenomenology at short and long delays (Flegal et al., 2010), which did not differ as a function of encoding strategy in the present study (see Table 2). In other words, false memories were accompanied by statistically equivalent rates of "illusory recollection" whether they occurred after short or long delays, and regardless of how deeply their associates were encoded.

General Discussion

In two experiments, a levels-of-processing manipulation was found to dissociate the frequency, but not the phenomenology, of false recognition between short-term and long-term testing. Processing depth had little effect on recognition performance at short delays, but deep encoding (relative to shallow encoding) elevated the rates of both accurate and false memory at longer delays. The finding from Experiment 1 that effects of an encoding strategy manipulation at STM emerged only at LTM testing is similar to the dissociation that Rose and colleagues (2010) observed in their levels-of-processing span task. The direction of the deep processing effects that appeared at LTM in both Experiment 1 and Experiment 2 (i.e., increases in false recognition as well as accurate recognition) also reproduces and extends published findings (Thapar & McDermott, 2001), demonstrating that encoding influences on false memory in the long term can originate even from four-word lists of semantic associates studied under STM conditions.

What this study uniquely demonstrates, however, is that regardless of whether processing is shallow or deep, false recognition errors occur within seconds of encoding, and the phenomenological experience of those illusory memories is relatively delay-invariant, despite their greater incidence at longer delays after deep encoding. Self-reports of confidence in false alarms to related lures in Experiment 1, and normalized rates of *remember* responses to related lures in Experiment 2, did not significantly differ between short-term and long-term testing, or as a function of processing depth. These data provide further support for the claim that compelling false memory illusions can arise within seconds of encoding, as shown by Flegal et al.

(2010), while also demonstrating conditions under which the incidence of short-term and long-term memory distortions are dissociable.

Whereas our earlier work found relatively delay-invariant false memory rates when encoding strategies were unconstrained, the present results imply that there may still be different processes underlying memory formation, monitoring, or both, at STM and LTM, when encoding strategies are constrained by instructions (see also Rose & Craik, 2012). By experimentally controlling processing depth and tracing its effects on the quantity and quality of semantic memory errors over time, the present study represents an important advancement in constraining theoretical accounts of false memory. Rates of false recognition at short and long delays were dissociated more by deep than shallow processing, consistent with predictions based on fuzzy trace theory that long-lasting gist traces strengthened by deep, meaning-based encoding will exert their effects predominantly under LTM conditions, when the availability of verbatim traces has been diminished. The stable false memory effect observed in our previous study (Flegal et al., 2010), where encoding instructions were not experimentally controlled, may have been a consequence of variability in depth of processing strategies engaged (and subsequently averaged) across participants.

However, the absence of a levels-of-processing effect on confidence ratings or "remember" responses to related lures appears at odds with gist-based explanations of false memory from the domain of LTM. If verbatim traces are strengthened by deep encoding, then their availability in STM would be expected to oppose false recollection at short delays, yet we found that phenomenological measurements were not affected by processing depth. Perhaps the simultaneous enhancement of gist traces through

deep encoding increases the probability of falsely attributing episodic details to related items (i.e., "content borrowing"; Lampinen, Meier, Arnal, & Leding, 2005), thereby counteracting the protective influence of verbatim memory in the short term. Even if that were the case, the deterioration of verbatim traces over time would predict encoding-related differences in subjective experience to emerge at longer delays, contrary to the pattern of results in the present study.

An alternative interpretation for selective dissociations between short-term and long-term memory distortions is offered by the activation-monitoring account of false memory (Gallo & Roediger, 2002; Robinson & Roediger, 1997; Roediger, McDermott, & Robinson, 1998). According to this theoretical framework, false memories arise from the associative activation of lure words related to studied items and the failure of monitoring processes to identify the source of memory signals. While activation is likely affected by delay, and by processing depth, the consequences of monitoring success or failure may be less time-dependent. That is, monitoring failures, whether they occur under STM or LTM conditions, may result in equally compelling false memory illusions. However, such an account cannot readily explain why, if the subjective experience of false recognition results from transient associative activation, phenomenological measurements would not measurably decline over time.

As Rose and Craik (2012) have proposed, the distinction between STM and LTM may be best understood from a processing approach. Rather than representing separable memory systems, it is likely that short-term and long-term testing rely on some shared, and some unique, memory processes. The degree of time-invariance observed in measures of memory performance and phenomenology will therefore

depend upon the amount of overlap between the encoding, maintenance, and retrieval processes required by the chosen STM and LTM tests. This reasoning is compatible with other unitary models of memory (e.g., Cowan, 1999; Nairne, 2002), which posit that STM and LTM are supported by common processes but allow for factors such as states of activation or the availability of different retrieval cues to vary over time.

The dissociation between effects of processing depth on false recognition at short-term and long-term testing observed in the present study suggests that our encoding manipulation revealed different processes operating under STM and LTM conditions (similar to the results reported by Rose et al., 2010). Such an interpretation is consistent with fuzzy trace theory, which states that the predominance of verbatim retrieval at short delays gives way to gist retrieval at longer delays. However, it is unclear how the involvement of different memory processes at STM and LTM can account for the stability observed in false memory phenomenology. Contrary to predictions that deep encoding would reduce false recollection by strengthening both short-lasting verbatim traces and long-lasting gist traces lacking in episodic detail, normalized rates of *remember* responses to related lures were not significantly affected by encoding strategy or delay. From a processing approach, this finding may point to a memory process contributing to subjective experience which is susceptible to distortion and, unlike verbatim and gist-based processing, common to STM and LTM tests.

References

Anastasi, J. S., Rhodes, M. G., & Burns, M. C (2000). Distinguishing between memory illusions and actual memories using phenomenological measurements and explicit warnings. *American Journal of Psychology*, *113*(1), 1-26.

Atkins, A. S., & Reuter-Lorenz, P. A (2008). False working memories? Semantic distortion in a mere 4 seconds. *Memory and Cognition*, *36*(1), 74-81.

Brainerd, C. J., & Reyna, V. F (2002). Fuzzy-Trace Theory and False Memory. *Current Directions in Psychological Science*, *11*(5), 164-169.

Coane, J. H., & McBride, D. M (2006). The role of test structure in creating false memories. *Memory and Cognition*, *34*(5), 1026-1036.

Coane, J. H., McBride, D. M., Raulerson, B. A., & Jordan, J. S (2007). False memory in a short-term memory task. *Experimental Psychology*, *54*(1), 62-70.

Cowan, N (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). Cambridge: Cambridge University Press.

Craik, F. I. M., & Lockhart, R. S (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684.

Deese, J (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-22.

Dewhurst, S. A., Bould, E., Knott, L. M., & Thorley, C (2009). The roles of encoding and retrieval processes in associative and categorical memory illusions. *Journal of Memory and Language*, *60*(1), 173-183.

Flegal, K. E., Atkins, A. S., & Reuter-Lorenz, P. A (2010). False memories seconds later: The rapid and compelling onset of illusory recognition. *Journal of Experimental Psychology: Learning Memory and Cognition*, *36*(5), 1331-1338.

Gallo, D. A., Roberts, M. J., & Seamon, J. G (1997). Remembering words not presented in lists: Can we avoid creating false memories *Psychonomic Bulletin & Review*, *4*(2), 271-276.

Gallo, D. A., & Roediger, H. L (2002). Variability among word lists in eliciting memory illusions: evidence for associative activation and monitoring. *Journal of Memory and Language, 47*(3), 469-497.

Gardiner, J. M (1988). Functional aspects of recollective experience. *Memory and Cognition, 16*(4), 309-313.

Geraci, L., & McCabe, D. P (2006). Examining the basis for illusory recollection: The role of remember/know instructions. *Psychonomic Bulletin & Review*, *13*(3), 466-473.

Lampinen, J. M., Meier, C. R., Arnal, J. D., & Leding, J. K (2005). Compelling untruths: Content borrowing and vivid false memories. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*(5), 954-963.

Lampinen, J. M., Neuschatz, J. S., & Payne, D. G (1998). Memory illusions and consciousness: Examining the phenomenology of true and false memories. *Current Psychology*, *16*(3-4), 181-224.

Lampinen, J. M., Ryals, D. B., & Smith, K (2008). Compelling untruths: The effect of retention interval on content borrowing and vivid false memories. *Memory*, *16*(2), 149-156.

Mather, M., Henkel, L. A., & Johnson, M. K (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory and Cognition*, *25*(6), 826-837.

McDermott, K. B (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, *35*(2), 212-230.

McDermott, K. B., & Roediger, H. L (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language*, *39*(3), 508-520.

Nairne, J. S (2002). Remembering over the short-term: the case against the standard model. *Annual Review of Psychology, 53*, 53-81.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/

Neuschatz, J. S., Payne, D. G., Lampinen, J. M., & Toglia, M. P (2001). Assessing the effectiveness of warnings and the phenomenological characteristics of false memories. *Memory*, *9*(1), 53-71.

Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language*, *35*(2), 261-285.

Rajaram, S (1993). Remembering and Knowing: Two means of access to the personal past. *Memory and Cognition*, *21*(1), 89-102.

Reyna, V. F., & Brainerd, C. J (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*(1), 1-75.

Rhodes, M. G., & Anastasi, J. S (2000). The effects of a levels-of-processing manipulation on false recall. *Psychonomic Bulletin & Review*, *7*(1), 158-162.

Robinson, K. J., & Roediger, H. L (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8*(3), 231-237.

Roediger, H. L., & McDermott, K. B (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning Memory and Cognition*, *21*(4), 803-814.

Roediger, H. L., McDermott, K. B., & Robinson, K. J (1998). The role of associative processes in producing false memories. In M. Conway, S. Gathercole, & C. Cornoldi (Eds.), *Theories of memory* (Vol. 2, pp. 187-245). Hove, England: Psychology Press.

Rose, N. S., & Craik, F. I. M (2012). A processing approach to the working memory/long-term memory distinction: Evidence from the levels-of-processing span task *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(4), 1019-1029.

Rose, N. S., Myerson, J., Roediger, H. L., & Hale, S (2010). Similarities and differences between working memory and long-term memory: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning Memory and Cognition*, *36*(2), 471-483.

Thapar, A., & McDermott, K. B (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory and Cognition*, *29*(3), 424-432.

Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A (1999). Recall accuracy and illusory memories: When more is less. *Memory*, *7*(2), 233-256.

Tulving, E (1985). Memory and Consciousness. *Canadian Psychology*, *26*(1), 1-12.

Yonelinas, A. P (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441-517.

Acknowledgements

<u>Table 1</u> Mean recognition proportions in Experiments 1 and 2.

	Experiment 1			
	<u>Shallow Encoding</u>		<u>Deep Encoding</u>	
	STM	LTM	STM	LTM
<u>Probe type</u>	"Yes"	"Yes"	"Yes"	"Yes"
Related lure	.18	.32	.16	.53
Unrelated lure	.02	.28	.05	.26
Target	.82	.53	.87	.84

	Experiment 2							
	<u>Shallow Encoding</u>				<u>Deep Encoding</u>			
	STM		LTM		STM		LTM	
<u>Probe type</u>	"Yes"	(R/K/G)	"Yes"	(R/K/G)	"Yes"	(R/K/G)	"Yes"	(R/K/G)
Related lure	.19	(.03/.06/.09)	.30	(.10/.10/.10)	.13	(.03/.06/.04)	.49	(.18/.16/.15)
Unrelated lure	.00	(.00/.00/.00)	.15	(.02/.05/.07)	.02	(.00/.00/.02)	.18	(.01/.06/.11)
Target	.85	(.57/.18/.09)	.52	(.25/.15/.13)	.92	(.71/.18/.03)	.90	(.71/.13/.06)

Table 2 Experiment 2 proportion of "Remember", "Know", and "Guess" responses out of total proportion "Yes" responses.

	Shallow Encoding											
	STM						LTM					
	"Remember"		"Know"		"Guess"		"Remember"		"Know"		"Guess"	
Probe type	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM
Related lure	.21	.08	.27	.07	.52	.08	.28	.06	.31	.06	.41	.08
Unrelated lure	.00	.00	.00	.00	.00	.00	.13	.06	.36	.09	.52	.09
Target	.63	.06	.25	.05	.12	.03	.45	.06	.29	.05	.26	.05

	Deep Encoding											
	STM						LTM					
	"Remember"		"Know"		"Guess"		"Remember"		"Know"		"Guess"	
Probe type	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM
Related lure	.29	.11	.45	.10	.25	.09	.36	.06	.28	.05	.36	.06
Unrelated lure	.00	.00	.00	.00	1.00	.00	.05	.04	.37	.08	.58	.09
Target	.77	.04	.19	.03	.04	.01	.80	.04	.14	.03	.06	.02

Figure 1 Experiment 1 design. During the study phase, a four-point judgment (shallow encoding or deep encoding, depending on task block) is made for each item. Subsequently, each four-word list is probed only once; either immediately following a 3-4 second filled retention interval (short-term memory), or in a surprise recognition test after all lists are encoded (long-term memory).
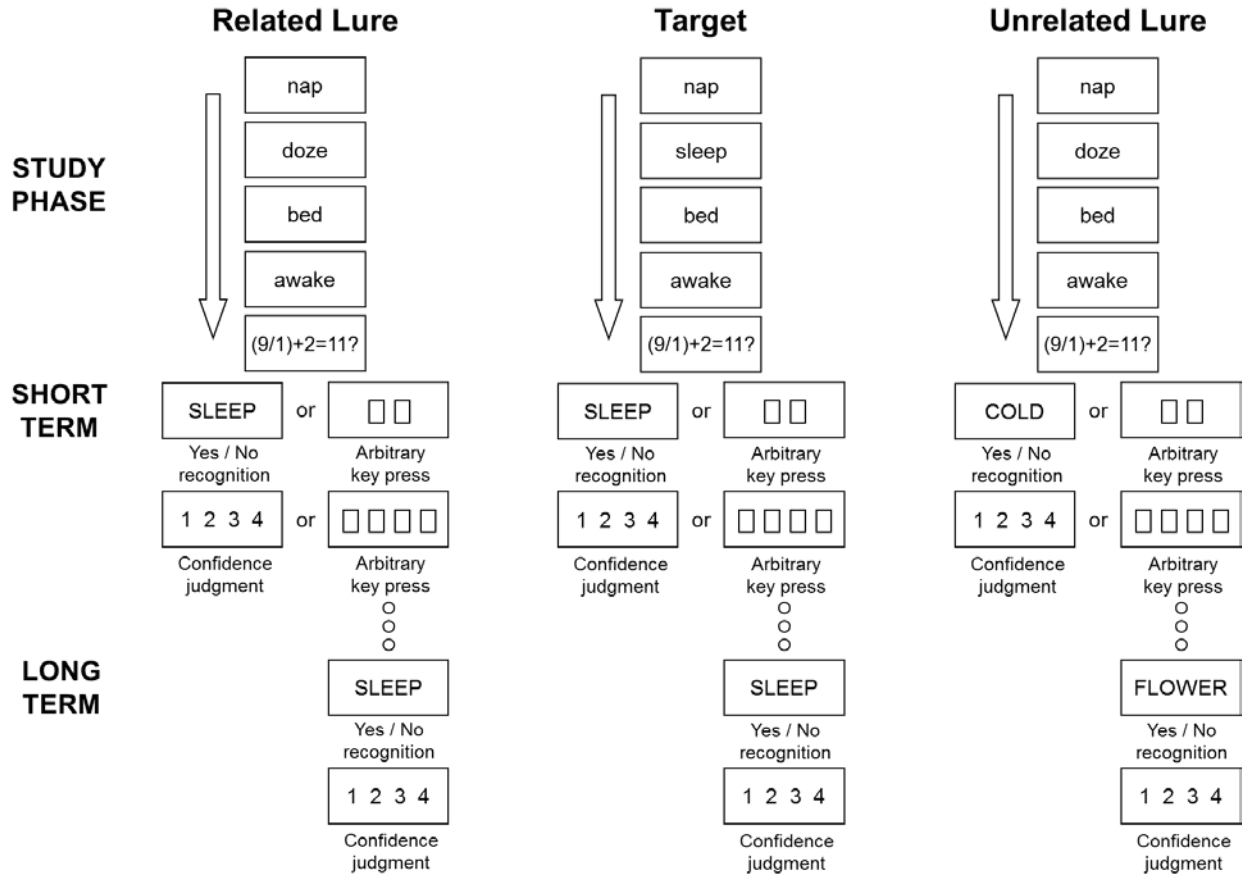
Figure 2 Experiment 1 mean confidence ratings by encoding condition, probe type, and delay (error bars = SEM; C = correct response).