Research Report

WHEN THINGS THAT WERE NEVER EXPERIENCED ARE EASIER TO "REMEMBER" THAN THINGS THAT WERE

C.J. Brainerd[1] and V.F. Reyna[2]

[1]College of Education and [2]Departments of Surgery and Medicine, University of Arizona

Abstract—*Can things that were never experienced be more readily accepted on recognition tests than things that were experienced? A current explanation of false memory predicts that this can happen when things that were never experienced provide superior access to the gist of events. This prediction was tested in three experiments in which the task was to accept all test items that were consistent with the substance of previously studied material, regardless of whether they had been studied. Acceptance rates were consistently higher for some never-studied items (those that provided superior access to gist memories) than for studied items. This effect varied predictably as a function of manipulations of the strength of gist memories and their accessibility. These results have implications for the use of exploratory memory-interrogation procedures in psychotherapy and the law.*

We report a surprising finding about false memory: On tests that parallel methods of memory interrogation that are common in psychotherapy and the law, things that were not experienced are more likely to be accepted than things that were. This finding is predicted by the theoretical principle that things that have not been experienced can provide privileged access to the semantic gist of events.

That memories undergo distortion over time, resulting eventually in high levels of false reports, is a notion that figured in the earliest theories of forgetting (Baldwin & Shaw, 1895; Philippe, 1897; Wulf, 1922). Recently, high levels of false reports have also been detected immediately after events were experienced and without any intervening misinformation or deception (e.g., Reyna & Kiernan, 1994; Roediger & McDermott, 1995; Schacter, Verfaellie, & Anes, 1997). This suggests that false memories can be by-products of how information is stored and retrieved in the first place, a possibility that was raised some years ago by Anisfeld and Knapp (1968).

A paradigm that induces remarkably high levels of immediate false reports under controlled conditions (Deese, 1959) is shown in Table 1. Subjects study a series of short word lists. The target words on the lists instantiate certain themes (e.g., furniture and medical themes in Table 1). A critical word that is an example of a list theme and is an associate of all the targets is omitted from each study list (*chair* and *doctor* in Table 1), but is included in test lists. On immediate recognition tests, for which the task is to accept only targets (e.g., *recliner, hospital*), acceptance rates for these critical distractors do not differ significantly from those for targets (Payne, Elie, Blackwell, & Neuschatz, 1996; Roediger & McDermott, 1995; Schacter, Verfaellie, & Pradere, 1996).

Fuzzy-trace theory's explanation of such data relies on a principle of differential access by distractors and targets to the types of memories that are stored at study (Reyna & Lloyd, 1997). Prior findings suggest that target presentations lead to the formation of dissociated memories of their surface forms (verbatim traces) and of various meanings that they cue (gist traces) (for a review, see Reyna & Brainerd, 1995). Lists like those in Table 1 have the important feature that particular meanings are repeatedly cued, resulting in strong gist memories of list themes (Reyna, in press; Schacter et al., 1997). On immediate tests, target acceptance can be based on verbatim traces of targets' surface forms or gist traces of cued meanings, but the verbatim basis predominates (Payne et al., 1996; Reyna & Kiernan, 1994). With distractors, gist is the predominant basis for false acceptance because verbatim traces of targets support correct rejections (e.g., "I heard *recliner*, not *chair*"; cf. Brainerd, Reyna, & Kneer, 1995; Clark & Gronlund, 1996). Distractors such as *chair* and *doctor* are especially good retrieval cues for strong gist memories of list themes because they are associates of all of the targets.

Thus, the posited memorial bases for target and distractor acceptance are different, the former being chiefly verbatim-based and the latter being gist-based. Because verbatim memories become inaccessible more rapidly than gist (Gernsbacher, 1985; Murphy & Shapiro, 1994), the otherwise puzzling finding that distractor acceptances can be more stable over time than target acceptances—the persistence effect—falls out as a prediction (Brainerd & Poole, 1997; Brainerd, Reyna, & Brandse, 1995). The same principle forecasts another crucial result: On immediate tests, acceptance rates for distractors will be even higher than those for targets if meaning is sufficient grounds for acceptance, for the simple reason that distractors are superior retrieval cues for these strong gist memories. There is a simple test, called meaning recognition, in which gist memories are the criterion for acceptance: Subjects are told to ignore whether an item was actually studied and to accept any item that preserves the meaning of studied material (Reyna & Kiernan, 1994).

This meaning-recognition procedure is also of considerable applied interest because it parallels forms of memory interrogation in psychotherapy and the law that focus on memory for substance. In psychotherapy, a number of practices encourage patients to go beyond what they explicitly remember and to recollect things that could or should have happened to them, based on beliefs, feelings, or intuitions about their experiences. Examples include biblio therapy, guided imagery, memory regression, memory work, and self-hypnosis (for surveys of such practices, see Kihlstrom, 1998; Poole, Lindsay, Memon, & Bull, 1995). In the law, trial testimony, like a standard recognition test, focuses on memory for actual experience. During the evidence-gathering phase of an investigation, however, police interviewers are trained to use interrogation techniques that encourage witnesses to assent to things that they cannot clearly remember (e.g., "Was the robber carrying a gun?") but that are consistent with the gist of remembered events, with the gist of their previous statements, or with the gist of statements made by other people (for a survey of such techniques, see Ceci & Bruck, 1995). In recent reversals of certain convictions, the courts have based their decisions on the tendency of such therapeutic and investigative procedures to taint trial testimony by causing witnesses to falsely assert that gist-

Address correspondence to C.J. Brainerd, College of Education, University of Arizona, Tucson, AZ 85721; e-mail: brainerd@u.arizona.edu.

Table 1. *Sample lists that produce high levels of immediate false reports*

List theme	Targets
Furniture words	table, sit, legs, seat, couch, desk, recliner, sofa, wood, cushion, swivel, stool, sitting, rocking, bench (critical missing word: *chair*)
Medical words	nurse, sick, lawyer, medicine, health, hospital, dentist, physician, ill, patient, office, stethoscope, surgeon, clinic, cure (critical missing word: *doctor*)

consistent events were experienced (e.g., "I *saw* the robber carrying a gun"; for extensive opinions, see *State v. Hungerford*, 1995; *State v. Morahan*, 1995).

In our experiments, we studied acceptance rates for distractors versus targets on meaning-recognition tests. Experiment 1 was concerned with whether the predicted phenomenon (higher acceptance rates for distractors than targets) could be demonstrated. The principle that predicts the phenomenon was then tested in Experiments 2 and 3. Experiment 2 dealt with whether the phenomenon reacts (as it should) to variations in the ability of distractors to access repeatedly cued gists. Experiment 3 dealt with whether the phenomenon reacts (as it should) to variations in how often gists are cued.

EXPERIMENT 1

Method

Subjects

Sixty undergraduates participated to fulfill a course requirement. Individuals were randomly assigned to either the standard or the meaning-recognition condition.

Materials and procedure

The materials were the 24 Roediger and McDermott (1995) lists, each of which has a critical distractor. Subjects were told that they would listen to several short lists of words, after which their memory would be tested. Next, each subject listened to an audio recording of 12 lists, which had been randomly selected from the pool of 24. Order of list presentation was random. Individual words were presented at a 3-s rate, with stronger associates presented before weaker associates (the order specified in Roediger and McDermott's appendix). There was a 10-s pause between consecutive lists.

Following List 12, subjects read a page of instructions about how to respond to the recognition test. Two minutes were allowed for reading. Half the subjects received standard instructions, and half received meaning-recognition instructions. Standard instructions described the types of items on the upcoming test (targets, distractors that preserved the themes of the lists, and unrelated distractors) and told subjects to accept targets and reject all distractors. Meaning-recognition instructions were the same, except, at the end, subjects were told to ignore whether or not an item had been studied and to accept all items that were consistent with list themes.

Table 2. *Mean proportions of acceptances (standard deviations in parentheses) in Experiments 1 and 2*

	Type of recognition	
Type of item	Standard	Meaning
Experiment 1		
Target	.61 (.08)	.68 (.09)
Distractor		
Critical	.63 (.13)	.88 (.08)
Unrelated critical	.19 (.08)	.28 (.14)
Unrelated exemplar	.16 (.11)	.25 (.15)
Experiment 2		
Target	.58 (.07)	.67 (.06)
Distractor		
Critical	.54 (.14)	.86 (.11)
List associate	.36 (.13)	.76 (.14)
Missing exemplar	.24 (.10)	.48 (.16)
Unrelated critical	.16 (.10)	.24 (.14)
Unrelated exemplar	.13 (.11)	.20 (.09)

Subjects then listened to an audio recording of a test list. All test lists consisted of 72 items: 36 targets (3 per study list) selected at random from Presentation Positions 3 through 13, the 12 critical distractors for the presented lists, the 12 critical distractors for the unpresented lists, and 12 additional distractors (1 per list) from Presentation Positions 3 through 13 of the unpresented lists. Individual words were presented in random order at a 4-s rate.

Results

Means and standard deviations for proportions of acceptances are shown by condition and type of item in Table 2. Standard recognition performance resembled that of previous studies in that acceptance rates for targets and critical distractors (.61 and .63) did not differ significantly and were higher than acceptance rates for either type of unrelated distractor (.18 overall). In meaning recognition, however, acceptance rates were much higher for critical distractors than for targets, as predicted. Acceptance proportions were .88 for critical distractors, .68 for targets, and .27 overall for unrelated distractors. A 2 (condition: standard vs. meaning recognition) $\times$ 4 (type of item: target vs. critical distractor vs. unrelated critical distractor vs. unrelated exemplar) analysis of variance (ANOVA) of acceptance proportions produced a Condition $\times$ Type of Item interaction, $F(2, 174) = 7.61$, $MSE = 0.03$, $p < .0001$. Planned comparisons showed different orders of acceptance rates in standard and meaning recognition:

standard recognition: critical distractors = targets > unrelated distractors

meaning recognition: critical distractors > targets > unrelated distractors

It might be thought that this pattern could be explained on the ground that different recognition instructions induce different

response criteria, in the signal detection theory sense (e.g., McNicol, 1972). Even if this hypothesis were true, however, the predicted result did not involve comparisons between different instructional conditions. In this design, standard recognition is merely a control condition to establish that the ordering of acceptance rates for the different types of items is the same as in prior studies. Given this baseline, the predicted result is that within the meaning-recognition condition, acceptance rates will be higher for critical distractors than for targets (while rates for both will still be higher than rates for unrelated distractors). Even the most lenient response criterion should not produce reversals of acceptance rates for targets and distractors.

EXPERIMENT 2

The finding that critical distractors are easier to accept than targets on meaning-recognition tests is predicted by the principle that these distractors provide superior access to strong gist memories. To explore whether confidence in this principle is warranted, we examined whether variations in the finding could be tied to experimental manipulations of (a) the ability of different distractors to access the relevant gist memories and (b) the strength of those gist memories. The first factor was investigated in Experiment 2 and the second in Experiment 3.

Method

Subjects

Sixty undergraduates participated to fulfill a course requirement. Individuals were randomly assigned to either the standard or the meaning-recognition condition.

Materials and procedure

Prior to the experiment, 74 undergraduates listened to recordings of the 24 Roediger and McDermott (1995) lists plus the critical distractors (i.e., each list consisted of 16 targets rather than 15). Subjects performed a two-step production task following each list: Write a single word that captures the meaning of most of the list words (henceforth: list-associate distractor), and write three other words that could have been on the list but were not (henceforth: missing-exemplar distractors). Because the targets and critical distractor on each of Roediger and McDermott's lists are the most typical exemplars of the list's theme, missing exemplars were necessarily less typical than targets. There were 12 lists for which more than half the subjects gave the same list-associate word. These lists were used in the present experiment. For some lists, the list associate was very similar to a list word (e.g., *medical* predominated for the second list in Table 1). In those instances, we replaced the corresponding list word with another word from the pool of missing exemplars (e.g., *operation* for the second list in Table 1).

Critical, list-associate, and missing-exemplar distractors should form an accessibility hierarchy, with repeatedly cued gists being most accessible by critical distractors (because they are associates of all the targets), next most accessible by list associates (because they are associates of many of the targets), and least accessible by missing-exemplar distractors (because they are associates of few targets). List-associate distractors, by virtue of their connections with multiple targets, should have higher acceptance rates than targets on meaning-recognition tests, though the effect ought to be smaller than for critical distractors. Missing-exemplar distractors should be poorer retrieval cues for gist memories than are targets and therefore should have lower acceptance rates on meaning-recognition tests.

The procedure was the same as for Experiment 1, except for the test lists. Test lists for individual subjects consisted of 60 targets (5 per list) selected at random from Presentation Positions 3 through 13, the 12 critical distractors for those lists, the 12 list-associate distractors for those lists, the 12 missing-exemplar distractors for those lists that had been most frequently mentioned on the production task (1 per list), the 12 critical distractors for the 12 unpresented lists, and 12 additional distractors (1 per list) from Presentation Positions 3 through 13 of the unpresented lists.

Results

Means and standard deviations for proportions of acceptances are shown by condition and type of item in the lower portion of Table 2. In standard recognition, acceptance rates for targets and critical distractors again failed to differ significantly (.57 vs. .54), acceptance rates were higher for list-associate distractors than for missing-exemplar distractors (.36 vs. .24), and acceptance rates were higher for missing-exemplar distractors than for either type of unrelated distractor (.24 vs. .15 overall). In meaning recognition, however, acceptance rates for critical distractors (.86) and list-associate distractors (.75) were both higher than acceptance rates for targets (.67). Acceptance rates for missing-exemplar distractors (.48) were lower than for targets but higher than for unrelated distractors (.22 overall).

Consistent with this pattern, a 2 (condition: standard vs. meaning recognition) $\times$ 6 (type of item: target vs. critical distractor vs. list-associate distractor vs. missing-exemplar distractor vs. unrelated critical distractor vs. unrelated exemplar) ANOVA of acceptance proportions produced a Condition $\times$ Type of Item interaction, $F(5, 290) = 20.24$, $MSE = 0.01$, $p < .0001$. Planned comparisons showed different orders of acceptance in standard and meaning recognition:

standard recognition: targets = critical distractors > list-associate distractors > missing-exemplar distractors > unrelated distractors

meaning recognition: critical distractors > list-associate distractors > targets > missing-exemplar distractors > unrelated distractors

EXPERIMENT 3

In Experiment 2, higher acceptance rates for distractors than for targets were tied to manipulations of distractors' ability to access strong gist memories. This phenomenon should also be affected by manipulations of the strength of to-be-accessed gists. Such a prediction is difficult to investigate with Roediger and McDermott's materials because lists' themes are cued so often at study that the strengths of the relevant gist memories are very high. The number of times that lists' themes are cued can be reduced by reducing list length, but this approach has the disadvantage that hit rates for targets will increase simply because length decreases (Robinson & Roediger, 1997).

Consequently, we switched to other lists that had the added advantages of more closely resembling materials that are common in the memory literature and of producing much lower levels of immediate false reports. Subjects studied a single, randomly ordered list of familiar words, some of which (e.g., *couch*, *horse*) were exemplars of familiar semantic categories. On recognition tests, there were two types of meaning-preserving distractors: category names for which

three exemplars had been studied (strong gist) and category names for which one exemplar had been studied (weak gist).

Method

Subjects

Eighty undergraduates participated to fulfill a course requirement. Individuals were randomly assigned to either the standard or the meaning-recognition condition.

Materials and procedure

We used categorized word pools (e.g., Battig & Montague, 1969; Murdock, 1976) to obtain a set of 48 categories, each of which could be named by a single familiar noun (e.g., *animal*, *fruit*). For each category, the four exemplars with the highest frequency on production norms (e.g., *apple*, *orange*, *pear*, and *banana* for *fruit*) were selected. Each subject listened to a list of 96 words composed of the following groups of items: (a) 36 category exemplars obtained by randomly selecting 3 exemplars apiece from 12 of the 48 categories (strong gist: cued three times), (b) 12 category exemplars obtained by randomly selecting 1 exemplar apiece from 12 of the remaining 36 categories (weak gist: cued once), and (c) 48 filler targets selected from the Paivio, Yuille, and Madigan (1968) norms so as to match the items in the first two groups in average difficulty. Words were presented in random order at a 3-s rate.

Next, subjects read either standard or meaning-recognition instructions for 2 min. Test lists for individual subjects consisted of 84 items: (a) 12 targets that were obtained by randomly selecting one of the three exemplars of each strong-gist category; (b) the 12 targets that were the exemplars of the weak-gist categories; (c) 12 randomly selected filler targets; (d) 12 distractors that were names of the strong-gist categories; (e) 12 distractors that were names of the weak-gist categories; (f) 12 distractors that were names of categories that had not been exemplified on the study list; and (g) 12 distractors, each of which was an exemplar of one of the categories in Group f.

Results

Means and standard deviations for numbers of acceptances (out of 12) are shown in Table 3. Standard recognition performance was consistent with the results of many previous studies using similar materials: The acceptance rate was higher for targets from strong-gist categories (.70) than for targets from weak-gist categories (.61) and filler targets (.56), the acceptance rate was higher for distractors that were the names of strong-gist categories than for distractors that were the names of weak-gist categories (.27 vs. .19), and the acceptance rate for unrelated distractors was lowest of all (.10). Unlike the critical distractors in Experiments 1 and 2, the two types of meaning-preserving distractors had false alarm rates (M = .23) that were much lower than the corresponding hit rates for targets (M = .66).

In meaning recognition, the prediction is that the names of strong-gist categories should be more likely than the names of weak-gist categories to show higher acceptance rates than targets. That prediction was confirmed. The acceptance rate for the names of strong-gist categories exceeded the acceptance rate for their instantiating targets (.92 vs. .82), but the reverse was true for the names of weak-gist categories (.62 vs. .71). Acceptance rates for both types of category names were higher than acceptance rates for unrelated distractors (.17).

Table 3. *Mean numbers of acceptances (standard deviations in parentheses) in Experiment 3*

	Type of recognition	
Type of item	Standard	Meaning
Target		
Strong gist	8.38 (1.63)	9.83 (1.39)
Weak gist	7.30 (1.77)	8.48 (1.60)
Filler	6.70 (2.03)	7.80 (1.77)
Distractor		
Strong-gist category name	3.20 (1.57)	11.08 (0.94)
Weak-gist category name	2.23 (1.39)	7.45 (1.74)
Unrelated category name	1.33 (1.16)	2.02 (1.31)
Unrelated category exemplar	1.13 (0.99)	1.95 (1.11)

Consistent with this pattern, a 2 (condition) × 7 (type of item) ANOVA of acceptance proportions produced a Condition × Type of Item interaction, $F(6, 486) = 68.04$, $MSE = 2.31$, $p < .0001$. Planned comparisons revealed different orders of acceptance rates for standard and meaning recognition:

standard recognition: strong-gist targets > weak-gist targets = filler targets > strong-gist category names > weak-gist category names > unrelated distractors

meaning recognition: strong-gist category names > strong-gist targets > weak-gist targets > filler targets = weak-gist category names > unrelated distractors

GENERAL DISCUSSION

Although gist memories arise from experience, there may be better ways to access them than to use experience itself as a retrieval cue. It follows from this principle that when memory tests encourage reliance on gist, things that were not experienced will be easier to accept on recognition tests than things that were. Three results from the meaning-recognition conditions of our experiments were consistent with this prediction. First, category-name distractors and the critical distractors for the Roediger and McDermott (1995) lists were accepted at higher rates than targets. Second, distractors' tendency to be accepted at higher rates than targets varied as a function of their ability to access gist memories. Third, the tendency for distractors to be accepted at higher rates than targets varied as a function of the strengths of the relevant gist memories.

Such findings were expected on the basis of fuzzy-trace theory's analysis of the different types of memories that are accessed by targets and distractors. However, are there other hypotheses that do not rely on the verbatim-gist distinction that might explain these data? For instance, it might be argued that the results could be explained by assuming that only verbatim traces of targets are retrieved on memory tests and that on meaning-recognition tests, distractors (e.g., *animal*) are accepted by retrieving verbatim traces (e.g., of *horse*) and noticing the meaning overlap. Targets would then be accepted either by accessing their own verbatim traces or by accessing verbatim traces of other targets (e.g., of *cow*) and noticing the meaning overlap. If we assume that distractors are stronger associates of targets than targets are of each other, acceptance rates should be higher for distractors than for targets. Of course, because the probability of a target producing retrieval of its

verbatim trace will be higher than the probability of a distractor producing retrieval of that same trace, it is necessary to assume that the former probability is not large enough to compensate for the distractor's superior ability to access verbatim traces of associated targets.

However, this verbatim-only hypothesis conflicts with the data. For example, this hypothesis predicts that, relative to baseline acceptance rates for unrelated distractors, acceptance rates for both targets and meaning-preserving distractors will be higher in meaning recognition than in standard recognition. However, when *d'* values for targets were calculated for our experiments, the values for standard and meaning-recognition conditions did not differ significantly. Also, the verbatim-only hypothesis obviously predicts that acceptance rates in the standard condition of Experiment 3 will be lower for strong-gist category names than for weak-gist category names because verbatim traces of three exemplars support correct rejection of the former but verbatim traces of only one exemplar support correct rejection of the latter. Our results were in the opposite direction. More generally, theories that posit that only verbatim traces are retrieved on memory tests are unable to explain why acceptance rates decline more rapidly over time for targets than for related distractors such as those that were investigated in our experiments (cf. Roediger, McDermott, & Robinson, 1997). In contrast, differential forgetting rates are a hallmark of the verbatim-gist distinction (Reyna & Lloyd, 1997).

Robinson and Roediger (1997) identified two main explanations of the phenomenon of high levels of immediate false reports for certain types of distractors on standard tests: fuzzy-trace theory and Underwood's (1965) implicit associative response (IAR) theory. According to the latter, target presentations at study provoke retrieval of semantically related distractors (IARs) that are associates of targets. Although a subjective "presentation" from an IAR would not be expected to strengthen distractor memory as much as an actual presentation strengthens target memory, the fact that there may be multiple IARs for individual distractors with lists like those in our experiments could result in higher strength of distractor memories (Robinson & Roediger, 1997). In IAR theory, acceptance rates in standard recognition are treated as measures of the relative strength of target and distractor memories (Underwood, 1965).

Consequently, IAR theory does not predict our findings. The equal acceptance rates for targets and critical distractors in the standard conditions of Experiments 1 and 2 should have led to equal acceptance rates in the meaning conditions, and the higher acceptance rates for targets in the standard condition of Experiment 3 should have led to higher acceptance rates in the meaning condition. It might be possible to preserve the IAR account of standard recognition by hypothesizing that meaning recognition engages fundamentally different types of memory processes. Not only is this hypothesis unparsimonious, but other evidence argues against it. We have reported a series of modeling studies in which this hypothesis of different memory processes was explicitly tested and rejected for standard versus meaning recognition (Brainerd & Reyna, in press; Brainerd, Stein, & Reyna, 1998).

Finally, theoretical interpretation aside, our findings supply a potential answer to an important applied question: Why, in psychotherapy and the law, can false memories be created by procedures that emphasize memory for substance? That is, why do these exploratory interrogation procedures later create assertions that never-experienced events were experienced? A key consideration, it now seems, is whether strong gists are operative, as they invariably are during therapy and witness interviews. In therapy, sessions revolve around powerful uniting themes (e.g., emotional or physical trauma), with the events of patients' lives being explored in relation to those themes. In witness interviews, questioning also revolves around powerful uniting themes (e.g., crimes that are under investigation), with witnesses' statements focusing on things that bear directly on those themes. Based on our results, it no longer seems remarkable that false reports could be common in these situations with procedures that emphasize memory for substance. When strong gists are operative, things that were not experienced can seem more memorable than actual experience.

Acknowledgments—This work was supported by a grant to the authors from the National Science Foundation (SBR-9730143). We thank Kathleen McDermott, David Payne, Henry Roediger, and Johnathan Schooler for helpful comments.

REFERENCES

Anisfeld, M., & Knapp, M. (1968). Association, synonymity and directionality in false recognition. *Journal of Experimental Psychology, 77*, 171–179.

Baldwin, J.M., & Shaw, W.J. (1895). Memory for square size. *Psychological Review, 2*, 236–239.

Battig, W.F., & Montague, W.E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs, 80*(3, Pt. 2).

Brainerd, C.J., & Poole, D.A. (1997). Long-term survival of children's false memories: A review. *Learning and Individual Differences, 9*, 125–152.

Brainerd, C.J., & Reyna, V.F. (in press). Fuzzy-trace theory and children's false memories. *Journal of Experimental Child Psychology*.

Brainerd, C.J., Reyna, V.F., & Brandse, E. (1995). Are children's false memories more persistent than their true memories? *Psychological Science, 6*, 359–364.

Brainerd, C.J., Reyna, V.F., & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. *Journal of Memory and Language, 34*, 157–185.

Brainerd, C.J., Stein, L.M., & Reyna, V.F. (1998). On the development of conscious and unconscious memory. *Developmental Psychology, 34*, 342–357.

Ceci, S.J., & Bruck, M. (1995). *Jeopardy in the courtroom*. Washington, DC: American Psychological Association.

Clark, S.E., & Gronlund, S.D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3*, 37–60.

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

Gernsbacher, M.A. (1985). Surface information loss in comprehension. *Cognitive Psychology, 17*, 324–363.

Kihlstrom, J.F. (1998). Exhumed memory. In S.J. Lynn & N.P. Spanos (Eds.), *Truth in memory* (pp. 3–31). New York: Guilford Press.

McNicol, D. (1972). *A primer of signal detection theory*. London: Allen & Unwin.

Murdock, B.B., Jr. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General, 105*, 191–216.

Murphy, G.L., & Shapiro, A.M. (1994). Forgetting of verbatim information in discourse. *Memory & Cognition, 22*, 85–94.

Paivio, A., Yuille, J.C., & Madigan, S.A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monographs, 76*(1, Pt. 2).

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*, 261–285.

Philippe, J. (1897). Sur les transpositions de nos images mentales. *Review of Philosophy, 43*, 481–493.

Poole, D.A., Lindsay, D.S., Memon, A., & Bull, R. (1995). Psychotherapy and the recovery of memories of childhood sexual abuse: U.S. and British practitioners' opinions, practices, and experiences. *Journal of Consulting and Clinical Psychology, 63*, 426–437.

Reyna, V.F. (in press). Fuzzy-trace theory and false memory. In M. Intons-Peterson & D. Best (Eds.), *Challenges and controversies: Memory distortions and their prevention*. New York: Wiley.

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

Reyna, V.F., & Kiernan, B. (1994). The development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology, 30*, 178–191.

Reyna, V.F., & Lloyd, F. (1997). Theories of false memory in children and adults. *Learning and Individual Differences, 9*, 95–123.

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

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

Roediger, H.L., III, McDermott, K.B., & Robinson, K.J. (1997). The role of associative processes in creating false memories. In M.A. Conway, S.E. Gathercole, & C. Cornoldi (Eds.), *Theories of memory. II* (pp. 187–245). Hove, England: Psychological Press.

Schacter, D.L., Verfaellie, M., & Anes, M.D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology, 11*, 331–342.

Schacter, D.L., Verfaellie, M., & Pradere, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language, 35*, 319–334.

State v. Hungerford, Supreme Court, New Hampshire, 94-S-045 – 94-S-047 (1995).

State v. Morahan, Supreme Court, New Hampshire, 93-S-1734 – 93-S-1936 (1995).

Underwood, B.J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*, 122–129.

Wulf, F. (1922). Beitrage zur psychologie der gestalt: IV. Uber die veranderung von vorstellungen. *Psychologische Forschung, 1*, 333–375.

(RECEIVED 6/20/97; REVISION ACCEPTED 4/30/98)