Fuzzy-Trace Theory and False Memory

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

Department of Surgery (C.J.B., V.F.R.), Department of Special Education, Rehabilitation, and School Psychology (C.J.B.), and Department of Medicine (V.F.R.), University of Arizona, Tucson, Arizona

Abstract

A key problem confronting theories of false memory is that false-memory phenomena are so diverse: Some are characteristic of controlled laboratory tasks, others of everyday life; some occur for traumatic events with legal consequences, others for innocuous events; some are characteristic of one developmental level, others of another developmental level. Fuzzy-trace theory explains false memories via a small set of principles that implement a single representational distinction. Those principles generate new predictions, some of which are counterintuitive.

Keywords

fuzzy-trace theory; false memory; explanation; prediction

Fuzzy-trace theory (FTT) is not so much a model of false memory as a model of the interface between memory and higher reasoning processes. In origin, it was a response to findings about how the validity of solutions to reasoning problems (e.g., decision making, deductive inference, quantitative judgment) is related to memory for background facts that determine which solutions are valid (e.g., the premises in deductive-inference problems). Contrary to then-current information processing models, the data suggested that the key memory support for reasoning lies in having the right take on the background facts—getting the gist—and that, in contrast, remembering their exact content is not usually crucial (for a review, see Reyna & Brainerd, 1995). A surprising finding was that reasoning accuracy was largely independent of memory accuracy.

Research on human false memories intensified during the mid-1990s, and since then, FTT's assumptions have been used to gain leverage on such phenomena. The theory's explanations and predictions of false-memory effects are the topics of the present article. First, we sketch general criteria for a successful theory of false memory. Next, we summarize principles that FTT's account of false memory relies upon. Last, we describe a series of novel predictions about false memory that fall out of those principles.

CONSTRAINTS ON THEORIES OF FALSE MEMORY

Successful theories perform two principal tasks, explanation and prediction. We postpone the issue of prediction until later. Explanation, telling why the data are as they are, has two aspects, adequacy and parsimony. Adequacy refers to explanatory breadth: Other things being equal, explaining more facts is preferable to explaining fewer. Parsimony refers to explanatory frugality: Other things being equal, explaining known facts with fewer assumptions is preferable to explaining them with more assumptions. There is a trade-off between adequacy and parsimony such that increased adequacy can be purchased in the coin of reduced parsimony, which means that successful theories achieve a balance between empirical breadth and assumptive frugality.

Thus, explanatory progress consists of improving adequacy, parsimony, or both. In false-memory research, the most pressing problem at the moment is to improve adequacy. The reason is apparent in Table 1, which exhibits 10 false-memory effects, arranged in approximate historical order, each of which has an associated literature. The diversity of these phenomena is great: Some are produced by controlled laboratory tasks, others by real-world experiences; some involve innocuous events, others emotionally charged events with legal consequences; some are characteristic of adults, others of children. Therefore, although it is very desirable to be able to explain, say, false memories that arise when undergraduates study word lists (Examples 3 and 4), the compass of a successful theory will have to be far broader, so that the theory will be applicable to topics as varied as comprehension of narratives, psychotherapy, forensic interviews, and human development.

Although the diversity of the phenomena in Table 1 seems daunting, there is a common thread: Falsely remembered information is consistent with the gist of subjects' experience (e.g., an undergraduate falsely remembers reading *doctor* after reading several medical words, a high school student falsely remembers hearing the statement *The coffee is hotter than the cocoa* after hearing *The coffee is hotter than the tea* and *The tea is hotter than the cocoa*, a child-abuse victim falsely remembers being touched on the right breast rather than correctly remembering being touched on the left breast, an adult crime witness who knows that the prime

Published by Blackwell Publishing Inc.

Table 1. *Ten demonstrations of false memory*

Example	Phenomenon
1	After children observe a series of objects that ascend in magnitude (e.g., a series of rods that increase in length), their recall of the objects' appearance (e.g., the rods' lengths) is distorted upward.
2	Over repeated retellings, recall of the details of folk stories is distorted in ways that preserve their themes.
3	After studying a short word list (e.g., *nurse, hospital, sick, cure, ill*, . . .) whose members share a familiar semantic associate (*doctor*), subjects falsely recall (and recognize) the associate when asked to recall (or recognize) only list words.
4	After studying a long word list containing a few exemplars of a familiar semantic concept (e.g., *elm, maple, pine*), subjects falsely recognize unpresented exemplars of that concept (e.g., *oak*).
5	After reading a narrative, subjects falsely recognize unpresented sentences that connect the meaning of several narrative sentences.
6	Witnesses to crimes falsely recognize events that are consistent with the nature of the crimes, especially following suggestion.
7	Witnesses to crimes falsely recognize faces of innocent people that have the same gender and ethnicity as police suspects.
8	During interviews that focus on possible sexual abuse, children falsely recall incidents of abuse.
9	During psychotherapy, clients who believe that they may be victims of abuse falsely recall being sexually abused as children.
10	After reading metaphorical prose, subjects falsely recognize unstated metaphors that are consistent with the stated metaphors.

police suspect is a Hispanic male identifies a photograph of an innocent Hispanic male). It is this frequent feature of false-memory phenomena—that otherwise disparate falsely remembered objects and events fit with the gist of experience—that allows them to be explained with a common set of principles.

EXPLANATORY PRINCIPLES

It might be thought, considering the diversity of false-memory phenomena, that a comprehensive theory would have to be quite complex. Research has shown, however, that this need not be the case. In FTT's account of false memory, most of the explanatory load is borne by the following five principles, all of which implement a single representational distinction—namely, the distinction between verbatim and gist traces of experience.

Principle 1: Parallel Storage of Verbatim and Gist Traces

Subjects process the surface form and meaning content of experience in parallel, storing dissociated representations of the former (verbatim traces) and the latter (gist traces). Verbatim traces are episodically instantiated representations of the surface forms of experienced items, including contextual cues, and gist traces are episodic interpretations of concepts (meanings, relations, patterns) that have been retrieved as a result of encoding items' surface forms. Although it was once supposed that meaning content is extracted from previously stored verbatim memories, it is now known that meaning processing begins as early as 20 to 30 ms after the onset of a stimulus, well before the processing of the item's surface form is completed. Gist traces can be stored at various levels of specificity (e.g., reading *spaniel* may cause the concepts "hunting dog," "farm animal," and "pet" to be stored).

Principle 2: Dissociated Retrieval of Verbatim and Gist Traces

Memory performance is based on the retrieval of both verbatim and gist traces. The mix of verbatim and gist retrieval, which determines the level of false memory for gist-consistent information, turns on factors such as (a) retrieval cues, (b) the relative accessibility of verbatim and gist memories, and (c) forgetting. Concerning retrieval cues, items that were experienced (e.g., *spaniel*) are better retrieval cues for verbatim traces than items that were not, and nonexperienced items that preserve the meaning of experience (e.g., *collie*) are usually better retrieval cues for gist traces than for verbatim traces. However, retrieval can be slanted in one direction or the other by instructions that tell subjects how to use retrieval cues. Concerning relative accessibility, verbatim retrieval is favored when verbatim traces are strong relative to gist traces (as when the same item, e.g., *spaniel*, has been repeatedly presented), and gist retrieval is favored when gist traces are strong relative to verbatim traces (as when different exemplars of a meaning are presented but no exemplar is repeated, e.g., *hound, poodle, retriever, spaniel*). Concerning forgetting, a

familiar finding is that over time, memory for surface form declines more rapidly than memory for meaning, suggesting that verbatim traces become inaccessible more rapidly than gist.

Principle 3: Dual-Opponent Processes[2] in False Memory

Verbatim and gist retrieval both support true memory for experienced items, either because the corresponding experiences are specifically recollected (verbatim retrieval) or because items' meanings are familiar. However, verbatim and gist retrieval have opposite effects on false memory for items that preserve the meaning of experience (e.g., *collie* when *spaniel* was presented). Gist retrieval supports false memory because items' meanings seem familiar, but verbatim retrieval suppresses false memory by neutralizing meaning familiarity, either at the level of individual items ("No, I distinctly remember hearing *spaniel*, not *collie*") or at the level of general cognitive strategies ("I won't accept any word unless I have an auditory image of its presentation"). An important exception to this rule occurs when false items have been presented as suggestions (e.g., Example 6 in Table 1). In such circumstances, both gist retrieval and retrieval of verbatim traces of suggestions support false memory, while retrieval of verbatim traces of original experiences suppresses false memory.[3]

Principle 4: Developmental Variability

Acquisition, retention, and retrieval of both verbatim and gist memories improve as the child develops into an adult. Gist memory improves during development because both the ability to process the meaning of individual items and the ability to connect meaning across different items improves. These abilities decline during late adulthood, but declines are more marked for verbatim memory.

Principle 5: Verbatim and Gist Processing Both Cause Vivid Remembering

Retrieval of verbatim traces supports a vivid form of remembering, sometimes called recollection, in which subjects consciously reexperience items' occurrence in specific contexts. Retrieval of gist traces usually supports a more generic form of remembering, sometimes called familiarity, in which nonexperienced items are perceived to resemble experienced items but their occurrence is not explicitly recalled. However, when gist traces are especially strong, they can support high levels of phantom recollective experience for certain types of nonexperienced items—namely, items that are good cues for the gist of experience.

PREDICTIONS ABOUT FALSE MEMORY

The other task of a theory is to forecast new phenomena. Some predictions are more probative than others, and the most interesting ones are those that seem counterintuitive, either from the perspective of currently accepted theories or from the perspective of common sense. Predictions of this sort have been emphasized in applications of FTT to false memory. These predictions have often been contrasted with those of one-process approaches, such as constructivism (Bransford & Franks, 1971) or the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), which assume that true and false memories are based on a common memory code (constructivism) or a common retrieval process (source monitoring).

Dissociations and Associations Between True and False Memories

In memory research, two phenomena are said to be positively associated if higher frequencies of one produce higher frequencies of the other and lower frequencies of one produce lower frequencies of the other. Otherwise, the phenomena are said to be dissociated. Positive associations between true and false memories are basal predictions of one-process approaches (Reyna & Brainerd, 1995). In contrast, FTT predicts that true-false dissociations will be observed under certain conditions, and associations will be observed under others. These predictions fall out of Principles 1 through 3, according to which subjects retrieve dissociated verbatim and gist traces, with true responses being supported by both kinds of traces and false responses being supported by gist traces only. Thus, dissociations should be observed in situations that foster reliance on verbatim traces as the basis for true responses (e.g., when memory for memorable material is tested immediately after its presentation), so that true and false responses will be based on different representations. Data confirming this prediction have been reported (Reyna & Kiernan, 1994, Experiment 1, and 1995, Experiment 1; Roediger, Watson, McDermott, & Gallo, 2001). In contrast, positive associations between true and false memory should be observed in situations that foster reliance on gist traces as the basis for true responses (e.g., when memory testing is delayed for several days and subjects are instructed to rely on the meaning content of experience), so that true and false responses will be based on the same representations. Con-

firmatory data have been reported for this prediction as well (Reyna & Kiernan, 1994, Experiment 2, and 1995, Experiment 2).

False Persistence

Common sense suggests that true memories will be far more stable than false ones because the events that figure in the former were actually experienced, whereas those that figure in the latter were not. This idea occupies a prominent position in the law, where it is known as the consistency principle of testimonial credibility: Events that witnesses report consistently over time are regarded as more likely to be true than events that they report inconsistently. However, FTT predicts, on the basis of Principles 2 and 3, that false memories can be highly persistent and that, surprisingly, they can be more persistent than true memories under certain conditions. Persistence of false memories is predicted because they arise from particularly stable memory representations—namely, gist traces. The conditions under which false memories are expected to be more persistent than true memories are ones in which the events in the false memories are especially good retrieval cues for strong gist memories (e.g., Examples 3 and 4 in Table 1), and initial true-memory responses are based on less stable verbatim traces. Several experiments have shown high levels of false-memory persistence, including greater persistence of false than true memories (Brainerd, Reyna, & Brandse, 1995; Payne, Elie, Blackwell, & Neuschatz, 1996; Toglia, Neuschatz, & Goodwin, 1999).

Creation of False Memories by Mere Testing

In the law, particularly in eyewitness-identification cases and cases in which evidence is provided by children, it is acknowledged that false memories can be implanted by investigative interviews that suggest to witnesses events that support charges against defendants. However, it is also assumed that as long as interviewers do not offer suggestions and merely provide recall prompts (e.g., "Tell me about the robber.") and recognition probes (e.g., "Did the robber have a gun in his hand?"), such interviews do not elevate levels of false reporting during subsequent interviews and sworn testimony and that, indeed, their principal effect is to inoculate true memories against forgetting. However, FTT predicts, on the basis of Principles 2 and 3, that neutral, nonsuggestive memory questions can substantially elevate later false memory when subjects respond to the questions by retrieving gist traces. False-memory levels ought to rise under such conditions because subjects become practiced at processing the very types of representations that support false-memory responses. Experiments have shown steady increases in false-memory responses over a series of nonsuggestive recognition or recall tests (Brainerd & Reyna, 1996; Payne et al., 1996).

Developmental Increases in False Memory

A familiar generalization, in the law as well as the psychological literature, is that children, particularly young children, are especially prone to false memories. Although there is much empirical support for this generalization (for a review, see Ceci & Bruck, 1993), FTT predicts, on the basis of Principles 3 and 4, that only certain false memories conform to this generalization and that, further, some important forms of adult false memory will be largely absent in children. According to Principle 4, false memories will conform to this generalization if young children are able to store and retrieve the relevant gist traces; in this case, false memories will decrease with developmental improvements in the ability of verbatim retrieval to suppress false-memory responses. By the same principle, false memories will violate the generalization if young children are unable to store or retrieve the relevant gist traces, so that false memories will increase with developmental improvements in these abilities.

An example is provided by a simple ability that is surprisingly slow to develop in children—namely, the formation of gist memories that involve spontaneously connecting meaning across several distinct events. (Examples 3 and 4 in Table 1, which involve the gist memories "a bunch of medical words" and "a bunch of tree names," are illustrative.) Because this ability develops slowly, FTT predicts that false-memory responses that depend on connecting meaning across distinct events will also develop slowly. Confirmatory data have been reported (Brainerd, Reyna, & Forrest, in press).

Phantom Recollection

FTT predicts (Principle 5) that certain false memories will be accompanied by high levels of phantom recollection (illusory vivid mental reinstatement of events' "occurrence"). Confirmatory data have been reported for tasks like those referred to in Examples 3 through 5 in Table 1 (e.g., Payne et al., 1996). A further prediction that follows from Principle 5 is that phantom recollection should be dissociated from true recollection (i.e., vivid mental experience that accompanies true-memory responses) by manipulations that selectively affect gist and verbatim memory, respectively. Confirmatory data have also been reported for this prediction (Brainerd,

Wright, Reyna, & Mojardin, 2001; Payne et al., 1996).

Repetition Has Different Effects on True and False Memory

Repeated presentation of target material (i.e., the event to be remembered) is one of the oldest manipulations in memory research. By Principles 2 and 3, repetition should increase false-memory responses if gist traces are not yet available, but should decrease false-memory responses if gist traces are already available (by strengthening the verbatim traces that can be used to suppress such responses; Reyna & Lloyd, 1997). Because gist traces are more stable than verbatim traces, repetition has disproportionate effects on verbatim traces once gist traces have been stored. Seamon and his associates (e.g., Seamon et al., 2002) have proposed that with standard experimental procedures for presenting stimuli, the differential effects of repetition on verbatim and gist traces should lead to an inverted-U relation between repetition and false-memory responses—that is, such responses should increase with initial repetitions but then decrease with subsequent repetitions as verbatim traces are strengthened. Experiments conducted by Seamon et al. (2002) confirmed this prediction (see Fig. 1).

Similarity Can Be Distinctive

A seemingly self-evident claim about false-memory responses is that they will increase as their similarity to experienced material increases (e.g., for Example 4 in Table 1, *oak* will produce more false recognition than, say, *fern*). However, according to FTT, this relation should hold consistently for meaning overlap but not for surface overlap (Reyna & Lloyd, 1997). This prediction follows because increasing surface overlap ought to have two opposing effects. On the one hand, high surface resemblance between true and false items should make it more difficult to discriminate the two when the latter are presented on memory tests, thereby increasing false responding. On the other hand, because verbatim traces are representations of surface information, increasing the surface resemblance between true and false items should make the latter better retrieval cues for verbatim traces of the former, thereby decreasing false responding. Findings that various surface-similarity manipulations either failed to increase or decreased false responding are consistent with this prediction (see the review by Reyna & Lloyd, 1997).

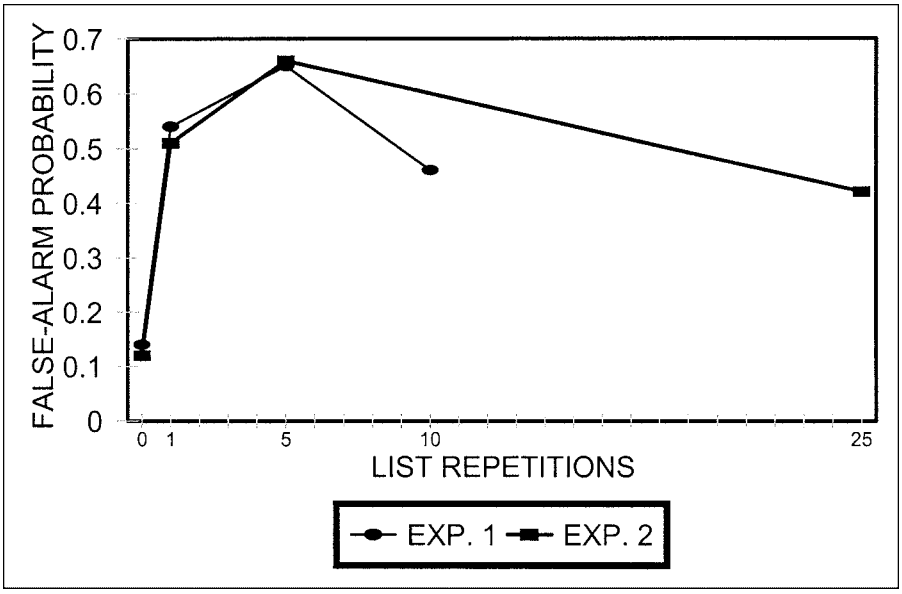

Fig. 1. False recognition (false alarm probability) of meaning-preserving items as a function of list repetition in two experiments by Seamon et al. (2002). Example 3 in Table 1 illustrates the task used. Target lists were presented either 0, 1, 5, or 10 times (Experiment 1) or 0, 1, 5, or 25 times (Experiment 2).

CONCLUSION

FTT has been applied to false memories in a range of laboratory tasks and real-world situations. The diversity of these phenomena notwithstanding, a small set of principles that implement a core representational distinction has proved useful in explaining false memory and generating new predictions about it. False-memory research is still in its infancy, however, and theoretical principles could change considerably down the road.

Recommended Reading

Brainerd, C.J., & Reyna, V.F. (2001). Fuzzy-trace theory: Dual-processes in reasoning, memory, and cognitive neuroscience. *Advances in Child Development and Behavior*, *28*, 49–100.

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, *51*, 481–538.

Schacter, D.L., Norman, K.A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology*, *51*, 289–318.

Acknowledgments—This research was supported by a National Science Foundation grant (SBR-9730143), by National In-

stitutes of Health grants (NIH31620 and P50AT00008), and by a Schering Plough/Integrated Therapeutics Group grant.

Notes

1. Address correspondence to C.J. Brainerd, Departments of Surgery and Special Education, Rehabilitation, and School Psychology, University of Arizona, Tucson, AZ 85721; e-mail: brainerd@u.arizona.edu.

2. *Dual-opponent processes* means that two distinct processes are involved in false memory, but they have opposite effects on observed levels of false-memory responses.

3. In the interest of brevity, we ignore this special case, restricting our attention to the standard situation in which verbatim traces that support false-memory responses are not available and such responses therefore arise from gist processing.

References

Brainerd, C.J., & Reyna, V.F. (1996). Mere memory testing creates false memories in children. *Developmental Psychology*, *32*, 467–476.

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., & Forrest, T.J. (in press). Are young children susceptible to the false-memory illusion? *Child Development*.

Brainerd, C.J., Wright, R., Reyna, V.F., & Mojardin, A.H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 307–327.

Bransford, J.D., & Franks, J.J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, *3*, 193–209.

Ceci, S.J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, *113*, 403–439.

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.

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.

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., & Kiernan, B. (1995). Children's memory and interpretation of psychological metaphors. *Metaphor and Symbolic Activity*, *10*, 309–331.

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

Roediger, H.L., III, Watson, J.M., McDermott, K.B., & Gallo, D.A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, *8*, 385–405.

Seamon, J.G., Luo, C.R., Schwartz, M.A., Jones, K.J., Lee, D.M., & Jones, S.J. (2002). Repetition can have similar and different effects on accurate and false recognition. *Journal of Memory and Language*, *46*, 323–340.

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