Representations, Predictions, and Remembrances in CHARM: A Reply to Lindsay (1991)

Janet Metcalfe
Dartmouth College

Lindsay (1991), in his comment *CHARMed, but not Convinced: Comment on Metcalfe (1990),* acknowledged that distributed models of human memory using the construct of a composite memory trace, such as the Composite Holographic Associative Recall Model (CHARM), are able to account for most of the findings within the eyewitness-testimony paradigm. Despite this success, Lindsay found CHARM to be of limited usefulness as a model of eyewitness suggestibility. The reasons stated for this peculiar conclusion are discussed in this rebuttal.

Lindsay (1991) noted that research on eyewitness testimony has been fraught with controversy and seemingly contradictory results. Loftus and her colleagues (e.g., Loftus, Miller, & Burns, 1978) have found that misleading suggestions give rise to impaired memory on event details. But other researchers (McCloskey & Zaragoza, 1985) using a Modified Test have found no apparent memory decrement. In addition, a variety of other experimental variants of the basic paradigm have produced seemingly conflicting findings. Such a situation cries out for rigorous formal modeling to allow integration and reconciliation of the results. Models may also provide new predictions and questions about the nuances of human memory, nuances that may be especially important in such a socially relevant paradigm. Lindsay appreciated many of the interesting characteristics of the Composite Holographic Associative Recall Model (CHARM), as it was applied to the domain of eyewitness testimony. For example, he said that this research "is valuable for a number of reasons. CHARM's interactive nature, use of composite storage, and success at simulating a variety of memory phenomena... make it an appealing model. At a more specific level, this work helps to integrate the suggestibility literature into the broader context of contemporary theoretical approaches to memory and cognition. This is important because discussions of suggestibility have sometimes been theoretically naive. Most important, CHARM demonstrates a formal mechanism by which suggestions could impair ability to remember event details without affecting performance on the Modified Test" (Lindsay, 1991, p. 102). These and other laudatory comments, though, belie his dissatisfaction with the model.

Lindsay's reservations centered around four major points. First, he thought that the prototype, or family-resemblance, type of representation that is used to represent category names in the model is incorrect and that category names must rather be represented as nodes in a hierarchical structure. Second, he thought that the model makes a prediction that disagrees with his intuitions about what would happen in a particular

experiment, were that experiment to be conducted. In particular, he thought that the model predicts that there should be no difference between category names and exemplars, but that such a difference would be found were the experiment to be conducted. Third, he thought that the model cannot properly be called a model of memory at all, because it has not been applied to memory-for-source characteristics of the to-be-remembered events. He considered that memory for source or for the "conditions of encounter" must be handled if a model is to count as a model of memory. And fourth, Lindsay evidently thought that it is the programmer's intelligence, rather than the inherent properties of the model, that gives rise to the theoretical predictions consistent with the data. In the paragraphs that follow, I discuss and, I hope, allay those concerns that are attributable to misconceptions and acknowledge those that represent genuine limitations in the model as it is currently formulated.

Lindsay's first objection to the CHARM model was that, in it, category names are represented as vectors that are roughly the central tendency (or prototype) of the category, a representation that he claimed to be incorrect. He stated that category names must be represented as superordinate units in a hierarchy: "In studies of human suggestibility the controls have been superordinate terms (e.g., "tool"), not prototypes (e.g., ?)... A superordinate term does not reduce to a blend of its subordinates" (Lindsay, 1991, p. 102). I do not argue that a family resemblance, or prototype, form of representation (in vector form) is the only conceivable form of representation for category names such as *tool*. The representation preferred by Lindsay is an alternative, and there are others. The question is, though, which form of representation in which situation explains the data? Rosch and Mervis (1975) reported multidimensional scaling solutions of people's similarity ratings of category members (including *prototypic members) and category names.* "In all cases the category names and the most prototypical items appeared to be the most central in the scaling solution regardless of the number of dimensions or the rotation used" (p. 583). These empirical results are mirrored by the representations used in CHARM. Their use allows the model to account for the memory data to which it is addressed. Lindsay presented no evidence whatever favoring the representation he preferred, but merely asserted that it must be used and that the CHARM model should be discredited for using an alternative. Its use of a

---

form of representation that is both widely accepted (see Smith & Medin, 1981; Wittgenstein, 1958) and that allows explanation of the data is not an adequate reason for rejecting a model.

Lindsay's second reservation about the model was that it makes a "counterintuitive prediction." But counterintuitive predictions are *attractive* features of models. Such predictions provide opportunities to test models on new grounds and perhaps discover something unexpected about human memory. As a strategy, finding counterintuitive predictions and *testing* them is reasonable. But Lindsay did not do this. Rather, he simply asserted that the model must be wrong given that it did not conform to (his) intuitions. Such overreliance on all-too-faulty, untested intuitions is one of the pitfalls that explicit models allow researchers to avoid.

In any case, though, the model does not make the prediction that Lindsay claimed, namely, that misled performance would exceed control performance, if the category name were pitted against the original event, *to the same extent* as control performance would exceed misled performance, if an unpresented category exemplar were pitted against the original event. Perhaps it would be "odd indeed" to find such a result, but it is not a prediction of CHARM. Lindsay miscomputed the predicted relative probabilities in the experimental and control conditions, because he based these computations on an example in which no category structure was represented. This case provides a mathematical limit and is interesting from that perspective, but it bears no relation to any real-world situation. There are no categories that have no category structure, so predictions based on that assumption are vacuous. He claimed that, for CHARM, there is little difference between exemplars and category names, but this reveals a misunderstanding of the effects of category structure in the model.[1] (This is not a trivial point, because understanding the effect of categorical structure is necessary for comprehension of why the model makes the correct predictions regarding Belli's, 1989, and Tversky & Tuchin's, 1989, recognition data, which are a major portion of the eyewitness-testimony data modeled in the original article.)

Lindsay's third critique was that CHARM is not a model of memory at all, but rather just a model of association. He pointed out that it does model adequately results within the eyewitness-testimony paradigm: the work of Loftus et al. (1978), in which impairment was found as a result of misleading information; the results of McCloskey and Zaragoza (1985), in which—with different alternatives at test—no impairment was found; Belli's (1989) and Tversky and Tuchin's (1989) data from recognition tests; the results of Chandler (1989), who used a control condition different from that of the previous researchers; Loftus's (1977) results on color biasing; and Zaragoza, McCloskey, and Jamis's (1987) paradigm using a recall test. But Lindsay considered that information about the source, context, or conditions of encounter of the to-be-remembered event, rather than just the event itself, is a necessary and defining characteristic of memory. And so he said that the model, which has not yet been applied to source judgments, is not really a model of memory. However, Lindsay and Johnson's (1989) own experiments investigating memory for source indicated its remembrance is nonobligatory. When, in their experiment, subjects were asked for simple recognition of the events (but not the source), the results were different in pattern from when they were also asked for source judgments. It was clear that subjects could use source knowledge to improve memory, but they did not automatically do so (as one would expect if source attribution were a necessary and defining feature of all memory tasks). Especially in the light of these results, it is not obvious that judgment of source should be considered the criterion in determining whether a particular phenomenon is or is not a memory phenomenon. Of course, it will be interesting to apply the model to source judgments, as another manifestation of human memory abilities. It has been applied to a variety of other paradigms in addition to the eyewitness-testimony situation, including, for example, interference paradigms, elaboration and encoding specificity effects, and recognition failure of recallable words (see Metcalfe, in press; Metcalfe & Murdock, 1981; Metcalfe Eich, 1982, 1985). These are classic memory phenomena, and a model that explains them, as well as the major memory data within the eyewitness testimony paradigm, is clearly a model of human memory.

Finally, Lindsay thought that the results of the model are due to the intelligence of the programmer (in selecting how objects are represented and which representations are to be associated), rather than to intrinsic properties of the model, but this is not so. All applications of the model were faithful representations of the experimental situations in the language of the model. The simulation results are the unbiased consequences of the joint constraints of the nature of the model (i.e., the operations for association formation, storage, retrieval, and decision) and the procedures and events used in the experiments.[2]

---

[1] Lindsay's footnote alluding to the effects of implementing similarity structure in CHARM is unclear. First, the model does not induce "distortion of event details," so claiming that such a construct (and the implications one might infer from it) is inherent to the model is incorrect; second, it is not obvious which responses should be considered correct and which in error, in the situation Lindsay outlined, so the experiment described seems doomed to be at best confusing; and third, the prediction Lindsay made based on his intuition, rather than on derivations or simulations of the model, seems plausible. But it is plausible from any perspective.

[2] Lindsay echoed several specious sentiments that recur occasionally and so deserve brief comment. First, CHARM does *not* assume that the features exist in the stimulus itself. Perceptual processing before entry in the episodic memory system has always been assumed. Second, the number of features is not unbounded as has previously been falsely claimed and as Lindsay reiterated. The model is finite. However, there may be a very large number of features. If we mean by features something like neural units, or subsets of neurons, then the number could be very large indeed, and no cogent argument has yet been raised against this possibility. Although Lindsay stated that it is difficult to envisage "an efficient system in which all items are represented by the same ordered set of features" (Lindsay, 1991, p. 104), exactly this kind of representation is commonplace in the nervous system. The retina provides a familiar example, although other examples exist. Finally, the operations of convolution, or correlation, need not be especially slow. They can be done in parallel in a single step (regardless of the size of the vectors). Indeed, hardware now exists in the form of an analog computer chip that computes correlation in one step.

References

Belli, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General, 118,* 72–85.

Chandler, C. C. (1989). Specific retroactive interference in modified recognition tests: Evidence for an unknown cause of interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 256–265.

Lindsay, D. S. (1991). CHARMed but not convinced: Comment on Metcalfe (1990). *Journal of Experimental Psychology: General, 120,* 101–105.

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition, 17,* 349–358.

Loftus, E. F. (1977). Shifting human color memory. *Memory & Cognition, 5,* 696–699.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 19–31.

McCloskey, M., & Zaragoza, M. S. (1985). Misleading postevent information and memory for events: Arguments and evidence against the memory impairment hypothesis. *Journal of Experimental Psychology: General, 114,* 1–16.

Metcalfe, J. (1990). Composite holographic associative recall model (CHARM) and blended memories in eyewitness testimony. *Journal of Experimental Psychology: General, 119,* 145–160.

Metcalfe, J. (in press). Recognition failure and the composite memory trace in CHARM. *Psychological Review.*

Metcalfe, J., & Murdock, B. B. (1981). An encoding and retrieval model of free recall. *Journal of Verbal Learning and Verbal Behavior, 20,* 161–189.

Metcalfe Eich, J. (1982). A composite holographic associative recall model. *Psychological Review, 89,* 627–661.

Metcalfe Eich, J. (1985). Levels of processing, encoding specificity, elaboration, and CHARM. *Psychological Review, 92,* 1–38.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Tversky, B., & Tuchin, M. (1989). A reconciliation of the evidence on eyewitness testimony: Comments on McCloskey and Zaragoza. *Journal of Experimental Psychology: General, 118,* 86–91.

Wittgenstein, L. (1958). *Philosophical investigations* (3rd ed., G. E. M. Anscombe, Trans.). New York: Macmillan.

Zaragoza, M. S., McCloskey, M., & Jamis, M. (1987). Misleading postevent information and recall of the original event: Further evidence against the memory impairment hypothesis. *Journal of Experimental Psychology: General, 13,* 36–44.