Journal of Experimental Psychology: General
1991, Vol. 120, No. 2, 203–210

0096-3445/91/$3.00

COMMENTS

Composite Models Never (well, hardly ever) Compromise: Reply to Schooler and Tanaka (1991)

Janet Metcalfe
Dartmouth College

Robert A. Bjork
University of California, Los Angeles

With respect to the influence of postevent information, Schooler and Tanaka (1991) made a useful distinction between composite recollections—in which subjects retrieve "items from both the original and the postevent sources" (p. 97)—and compromise recollections—in which subjects retrieve "at least one feature that cannot be exclusively associated with either the original or the postevent sources, but which reflects some compromise between [the] two" (p. 97). Schooler and Tanaka argued that only the latter constitutes good evidence for blend-memory representations of the CHARM-type. As it turns out, Schooler and Tanaka's intuitions (and Metcalfe & Bjork's, initially) are faulty. Compromise recall—defined as a preference for an intervening alternative over either of the actually presented alternatives—is not normally a prediction of CHARM and may not be a prediction of composite-trace models in general. Only under specialized conditions—a systematic displacement of the test alternatives or a systematic shift attributable to assimilation to prior semantic knowledge—will computer simulations of CHARM produce unimodal compromise recollection. Equally surprising is the fact that separate-trace models, under a different set of conditions, can predict compromise recollection.

As Schooler and Tanaka (1991) have pointed out, the notion of blended memories is slippery, and the implications may not always be obvious. It turns out that their intuitions, as well as those of Metcalfe, that so-called compromise memories are the most compelling evidence for the construct of a composite trace are incorrect. Before developing that argument further, it is necessary to specify the empirical results that should be considered evidence of *compromise recollection.*

Schooler and Tanaka's (1991) definition of compromise—that subjects retrieve at least one feature that cannot be exclusively associated with either the original or the postevent source—does not determine what empirical results should be considered evidence for compromise. Surely, a single subject's choosing blue-green in the prototypic Loftus (1977) experiment cannot be taken as such evidence. Such a subject may simply be recollecting a misencoding of the original or misleading event or might even be guessing after not having encoded either event. A stronger empirical requirement is to say that subjects' recollection of the original event must show a unimodal shift along some dimension in the direction of the misleading event. That is, if—in the case in which the original and misleading information can be viewed as points on some dimension such as color or size—subjects pick or recall an intermediate value more frequently than they do either the original or the postevent values, then one might want to argue that a "compromise" has occurred. Such a requirement is consistent with what Metcalfe (1990) refers to as "positive blends," is consistent with the blue-car–green-car case cited as an example of compromise memories by Schooler and Tanaka, and, we think, is consistent with what they mean by compromise recollections.

In any case, such positive blends, which we will also refer to as compromise recollections, appear to be the most compelling evidence for composite trace models. This appearance is deceiving. When, provoked by Schooler and Tanaka's (1991) commentary, we tried to nail down this unimodal-shift prediction, we found that it is *not* a prediction of CHARM (Composite Holographic Associative Recall Model). Indeed, it is only with the greatest difficulty (as elaborated shortly) that we are ever able to get the model to produce such a result.

The data suggesting that compromises occur were presented by Loftus (1977). In her experiment, subjects saw a slide in which a green car was evident behind the scene of an accident. Later, in the Misled condition, the experimenters suggested that the car had been blue. In the Control condition, no suggestion about the color of the car was given. At time of test, subjects were presented a series of colors and asked to choose the one that best matched the color of the car. The basic finding of the experiment was that in the Control condition, subjects chose the appropriate green color with the greatest frequency, whereas in the experimental (Misled) condition, there was a distinct and unimodal shift toward the blue end of the color spectrum.

Metcalfe (1990) simulated a situation such as that given by Loftus (1977) with CHARM, which is a distributed model in

This research was supported by Grant 1R29MH48066-01 from the National Institute of Mental Health to Janet Metcalfe. We thank Elizabeth Bjork for her comments on this article.

Correspondence concerning this article should be addressed to Janet Metcalfe, Department of Psychology, Dartmouth College, Hanover, New Hampshire 03755.

which items—represented as vectors—are associated by convolution, stored by being added into a composite memory trace, and retrieved by the operation of correlation. Although the CHARM model was used for the previous simulations of the eyewitness paradigm and for the simulations that will be presented shortly, the results obtained in the eyewitness testimony paradigm are not unique to that model. Nearly all distributed models that have been applied to human memory (e.g., Anderson, 1977; Kortge, 1990; McClelland & Rumelhart, 1986; Murdock, 1982; Rumelhart & McClelland, 1986) and also models that rely on parallel or superimposed retrieval of traces (e.g., Hintzman, 1986) would produce the same favorable results on the previously modeled data. To our knowledge, they would also generate the same predictions outlined below.

The Loftus (1977) green-car–blue-car experiment was simulated as follows in the original Metcalfe (1990) article. First, a lexicon of random vectors was set up. A color continuum was then constructed to mimic the gradation from yellow to purple, including the blue-to-green range. Two random vectors were chosen as the end points of the continuum. Features of the intervening items were then replaced to make them progressively more like one or the other end point, in a graded fashion. Each item differed on three features from each adjacent item, as shown in Figure 1, except that the location of replaced features was random. Green was considered to be two items removed from blue. As such, *blue* differed from *green* on six features. A random vector, labeled *car* was convolved with the vector for *green*. In the Misled condition, *car* was also convolved with *blue*, and the results were added, along with some irrelevant convolutions, into the composite memory trace, itself a vector. To mimic retrieval, *car* was correlated with the composite trace and the vector that resulted was matched to all the lexical items (though, in fact, only those within the color range were systematically nonzero). The item exhibiting the highest dot product against the retrieved item, hereafter called *resonance*, was considered to be the chosen color.

The intermediate color, *blue-green*, was chosen with a high frequency. This high frequency of choice of an unpresented "compromise" item provides evidence favoring integrated storage, as that construct is specifically instantiated in the model. The question that we may ask, though, is this: Was this intermediate item chosen more frequently than the originals (indicating compromise recall) or less frequently than those alternatives? The results from Simulation 5 of Metcalfe (1990) were ambiguous. The intermediate blue-green item was numerically chosen less frequently than either the presented blue or green items, but the differences among these three were small and might have been attributable to random error.

Simulations of the Composite-Trace Model

The first simulation below was carried out in the same manner as Simulation 5 in Metcalfe (1990), but rather than using a lexical item two steps removed from the original green item as the misleading blue suggestion, an item six steps removed was used. This change provided more opportunities for the shape of the distribution to manifest itself. The entire simulation, for each of the experimental and control conditions, was replicated 1,000 times.

Figure 2 shows that all of the intermediate color representations between the two encoded items had the same resonance to the retrieved item, whereas those of the representations of colors outside the range declined, depending on how far removed they were. The distribution of choices (Figure 3) is clearly bimodal.

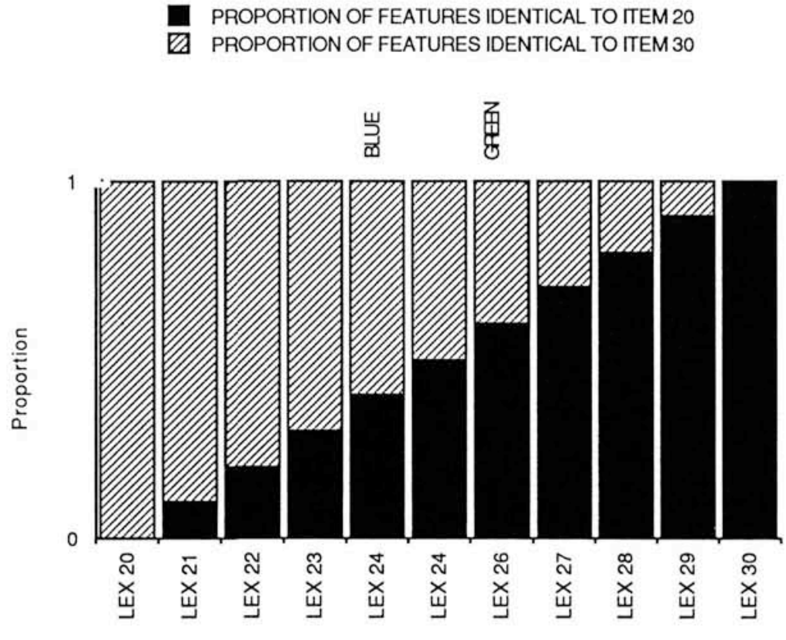

Figure 1. Presumed overlap of features in the memory vectors representing individual colors.

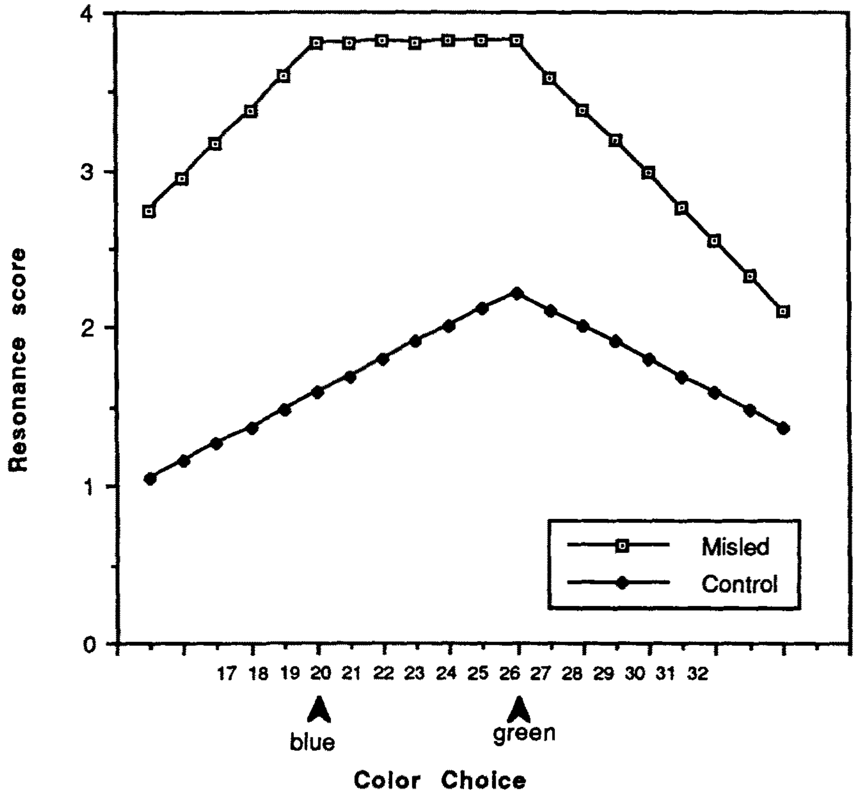

Figure 2. Resonance scores in the simulations of Loftus's (1977) color-shifting experiment given that blue and green are assumed to be six items apart and each item differs from its neighbors by three features.

We conducted intermediate simulations investigating variants of the degree of similarity among the items. We also varied the number of lexical items that intervened between the presented green and the suggested blue items. The results of all simulations were qualitatively the same. In what follows, we present only the extreme case that seemed most likely to show unimodal distributions, if such were a prediction of the model.This limiting case, which we will call Simulation 2, was like the first simulation, except that the separation between the original green and the misleading blue items was as small as possible—only two rather than six lexical items removed. In addition, each item differed from its neighbors by only one feature. We thought that the distribution might become unimodal in the CHARM model if the events were sufficiently similar to one another, and if there was only one choice between them. We ran this simulation through three runs of 1,000 replications each to be sure that the slight drop in the choice data (as was shown in all previous simulations, including Simulation 5 in Metcalfe, 1990) was not just a random error. It was suspected that the lower choice probability in the original simulation was probably due to chance factors; this (false) inference was at the heart of the compulsion evident in this simulation.

The simulated choice data are shown in Figure 4. Although the intermediate item is chosen much more frequently than any other nonpresented item (and particularly, it is chosen more than the nonpresented items immediately adjacent to the presented items, which bear the same similarity to their neighboring presented item as does the intervening item), it is chosen less frequently than the presented items.

Why the bimodal choice of distributions? As the distributions of resonance scores given above indicate, the degree of match to the retrieved item is high and about the same for the original and misleading items and for all of the items that intervene. This near identity in the resonance scores occurs because of the manner in which the continuous stimuli were constructed. Any feature in one of the intervening items that mismatched one of the critical items (say blue) necessarily matched the other (green). Thus, when an approximation to the superposition of blue and green was retrieved, if a given feature did not match on the green signal components it necessarily matched on the blue—providing a perfect trade-off and near identical resonance scores over the entire intervening range.

The probability of choice, however, depends not only on the degree of match, or the resonance of a given lexical item itself, but also on the degree of match of the particular retrieved vector to other lexical alternatives (i.e., not only on the item's resonance score, but also on the resonance scores of its neighbors). Consider, first, a presented item, say, the original green item in the Misled condition. Suppose a particular retrieved item matches it quite closely. It will, on average, also match its immediate neighbor to the left very closely, but will not match its immediate neighbor to the right as closely,

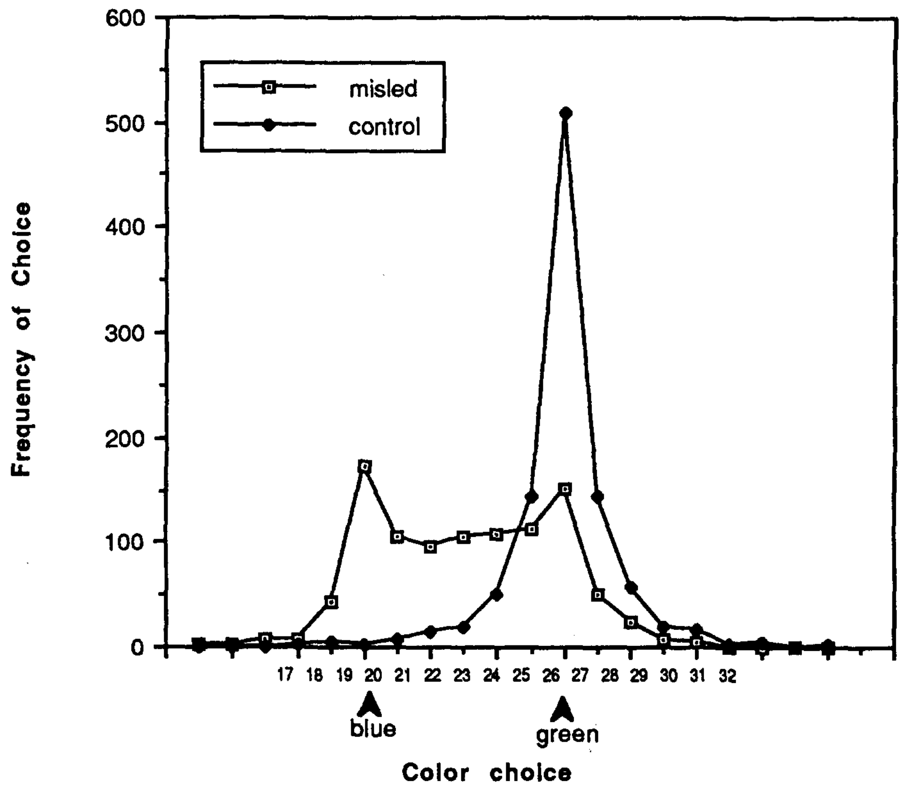

Figure 3. Color-choice data predicted in the simulations of Loftus's (1977) color-shifting experiment given that blue and green are separated by six hypothetical items.

as the resonance distributions show. Thus, there will be considerable competition from the neighbor to the left but little from the neighbor to the right. Contrast this situation to a central item, a blue-green item. Given the same degree of match of the retrieved item to this item as was shown in the first situation, one now finds that there is considerable competition not only from its neighbor to the left but also from its neighbor to the right. Thus, the probability that this central item will actually be chosen is decreased because of the increased competition from its neighbors. If one considers not only the immediately nearest neighbor, but also more remote neighbors (weighted appropriately), the result is still the same. For example, if one considers the competition for choice to be based on the three nearest neighbors on either side, one finds that the middle blue-green suffers from the most interference (from the high resonances of its competitors), the bluish green and greenish blue suffer slightly less interference, and the pure blue and the pure green suffer still less. Thus, despite the high resonances of intermediate items, the probability of their choice reflects not only their resonance or their goodness-of-match to the retrieved item, but also the competitive effects of the also-high goodness-of-match of their neighbors. The reason for the bowing shown here, as an aside, is roughly analogous to the reason for the bowing shown as a function of position in Estes's (1972) perturbation model of serial recall; namely, anchoring (or lack of confusion) at the ends of the continuum gives those items an advantage.

A Composite-Trace Explanation of Unimodal Shifts: The Displacement Hypothesis

The Loftus (1977) results showed a unimodal shift toward the blue end of the color spectrum. Is there *any* way that the CHARM model could produce the unimodal results? The answer is "yes," but only if there was some systematic bias in the experiment. Suppose that the nominal original green color given at test was slightly more yellowish than was the green color presented at study. This slight change in a direction outside the range could have been the same in both the Misled condition and in the Control condition, of course. Such a situation could arise for a variety of reasons. Perhaps there was a slight but real difference in the colors. Perhaps the lighting was slightly different between the slide sequence and the viewing of the Munsell chips. Perhaps subjects' adaptation level to the ambient light was a little different at study and at test. Perhaps these factors were the same, but the surrounding visual events in the slide produced some induced color alteration that was not apparent in the testing conditions. In terms of the simulation reported below, we supposed that the color encoded initially for the green car corresponded to one lexical item, but instead of having that exact item as an alternative, a near match (slightly more yellowish), was presented at the time of test. This is a small modification, but one that changes entirely the choice distributions.

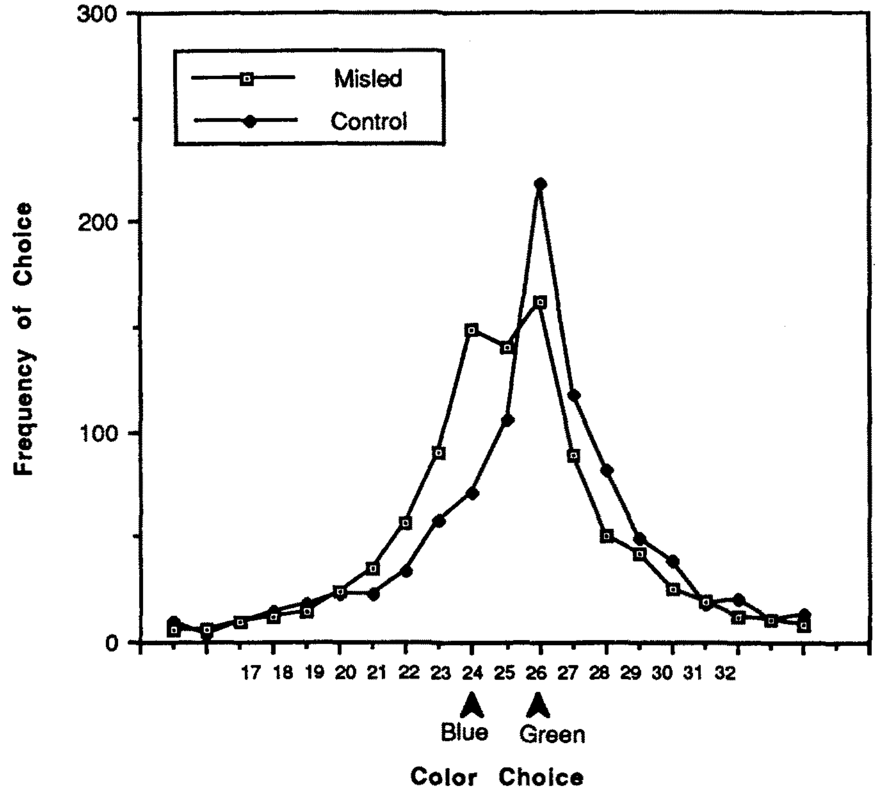

Figure 4. Color-choice data predicted in the simulation of Loftus's (1977) color-shifting experiment given that each item differs from its neighbors by one feature.

The simulation was set up in exactly the same manner, aside from different random starting values, as was Simulation 1 reported above. In the Misled condition, Item 26 was encoded as the green associated with the car vector, and Item 20 was given as the misleading blue alternative. In the Control condition, only Item 26 (green) was convolved with *car*. The alternatives from which the simulation chose the best match to the retrieved item were Items 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, and so on. Critically, Item 26—the true green— was not presented, but in its stead was Item 27 (a slightly yellowish green).

The results (Figure 5) indicate a nice peak at Item 27 for the Control condition. This peak occurs because Item 27 is highly similar to the actually encoded color—Item 26—and considerably more similar than is any other item. The results also show a unimodal blue shift in the Misled condition.

The model can also produce the pattern shown in Figure 5 if the memory representation itself, rather than the test color, drifted due to factors that were not purely episodic. Suppose, for example, that people had some a priori schematic representation of the typical color of cars. Over time, perhaps because of rehearsal of a prototypical color, the representation might drift toward this default color. Suppose further, that this default color was more toward the blue end of the spectrum than toward the yellow end of the spectrum. Casual observation of the color distributions of cars suggests that this direction is correct—yellow and orange cars are rare. This schematic bias due to nonepisodic factors would have the same effect in the model as the shifting of the test color toward yellow. There are two reasons why we cannot exclude this possibility: (a) As discussed below, Belli (1988) has shown that such schematic biasing does occur. (b) Loftus did not counterbalance the colors in the blue-shift experiment. Showing a blue car and implying a green car in the misleading sentence may not, according to such a modified version of CHARM, produce a unimodal shift.

A Separate-Trace Explanation of Unimodal Shifts: The Summed Distributions Hypothesis

Although intuitively the finding of a unimodal shift seems incompatible with the notion that traces are stored separately and may be sampled on an either–or basis, once again we find that those intuitions are incorrect. The problem is that the choice distributions plotted by Loftus (1977) represent the group mind, not necessarily the individual minds (memories) of the subjects in her experiment. That is, a given subject made a single choice in her experiment, and from that single choice one can conclude very little, as mentioned at the outset of this article.

As an illustration, assume that subjects in the Misled condition in the Loftus (1977) experiment sampled only their representation of the original event, or only their representa-

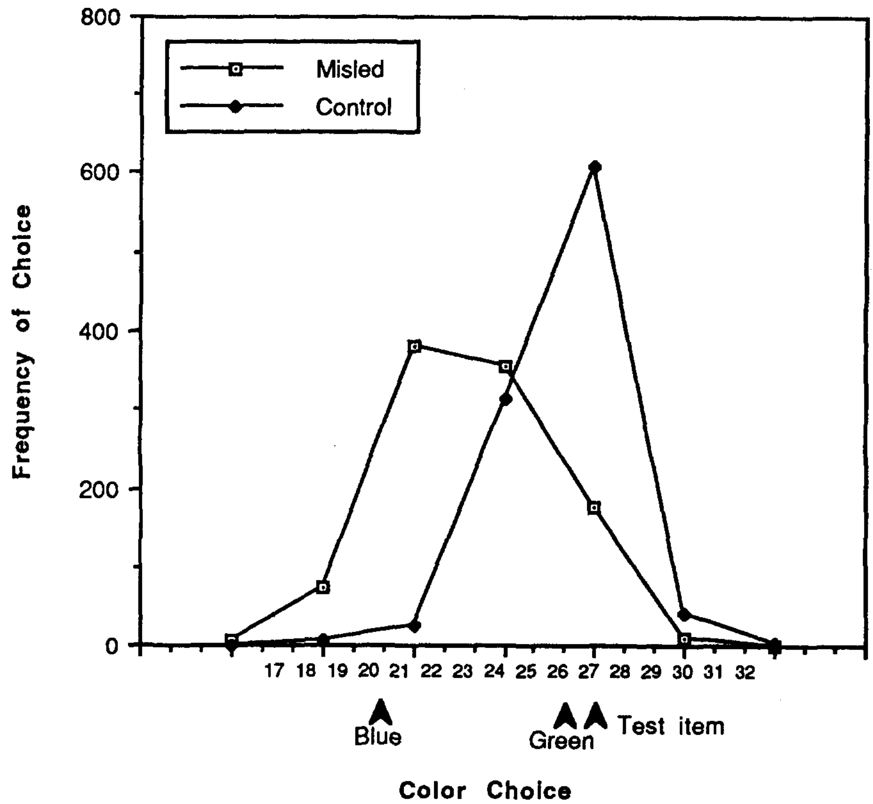

Figure 5. Predicted color shift when the test *green* is shifted one step in the yellow direction from the original green color and the alternatives are widely spaced.

tion of the misleading postevent sentence, but not both. Assume further that there was variation across subjects in how the original green color of the car was encoded and that there was also variation (in all likelihood, *greater* variation) in how the word *blue* in the misleading sentence was encoded (imaged). Two such hypothetical distributions are shown in Figure 6. The green distribution is a bell-shaped approximation of Loftus's Control condition (green only) results; the blue is a higher variance distribution of the same type with a mean in the blue range.

If we now assume, at the time of the final test, that some subjects referred to their memory representation of the original event and that the remaining subjects referred to their representation of the misleading sentence, the predicted group data are a unimodal distribution with a mean between the blue and green distributions. The particular distribution shown assumes that two of Loftus's (1977) subjects referred to the more recent sentence for every one subject who referred to the original event. The unconnected points in Figure 6 are the actual choice frequencies obtained by Loftus in her Misled condition.

The main point of Figure 6 is not to show a separate-trace fit of Loftus's (1977) data but to illustrate that the separate-trace-access assumptions outlined above *can* predict a unimodal shift. We do not want to leave the impression, however, that such an outcome is a parameter-free prediction of the separate-trace model. If the two summed distributions are too far apart (in d' terms), their sum will be bimodal not unimodal, and the nature of the shift is sensitive to the forms of the distributions and, of course, to their mixture ratio.

Belli's (1988) Data

Schooler and Tanaka (1991) considered the study of Belli (1988) to be critical evidence against the composite-trace hypothesis. Citing individual data points, they argued that "although misinformation produced rather few 'blend' errors, it substantially increased the percentage of subjects who selected the suggested color, uninfluenced by the original color. . . .Taken together, Belli's results clearly show that subjects were influenced by the misinformation, but provide little support for compromise recollections" (p. 98). Belli's data are hard to interpret because, as he demonstrated, his subjects had a default color for the green pitcher that was the critical item in his studies. The default color was yellow. Thus, control subjects, left to their own devices, show a drift in choices over time toward yellow. When shown no pitcher at all, they chose a yellow tone as the most probable color for the pitcher. To interpret whether there was biasing as a result of having been given misinformation, one needs somehow to factor out the effect of this schematic presupposition that the pitcher was yellow.

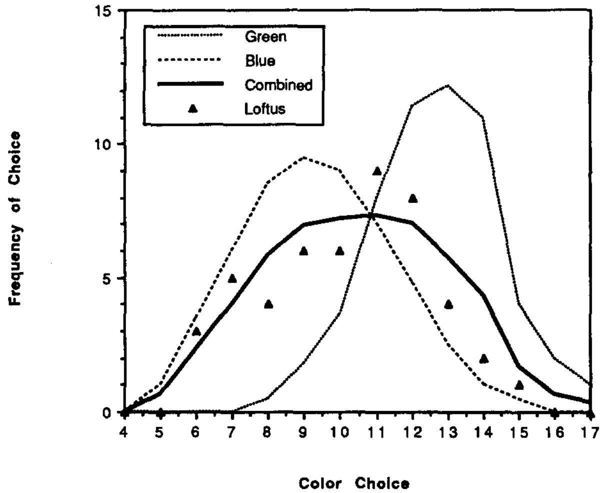

Figure 6. Prediction of the separate-trace model assuming that two thirds of the subjects accessed their *blue* representation and one-third of the subjects accessed their *green* representation; Loftus's (1977) data are shown as unconnected points.

To do so, we simply subtracted the color choices for the no-presentation condition (which peaked on yellow) from the green-only control condition, from the green-presentation-but-suggested-blue condition, and from the green-presentation-but-suggested-yellow condition. The results are noisy, and we certainly would not care to claim that they are either unimodal or bimodal. To see whether there was, at least, a bias effect, we smoothed the data by plotting each point as the average of itself and its immediate neighbor on each side. The results, given in Figure 7, show a clear effect of episodic biasing in just the directions that would be expected: toward the blue end of the spectrum if the suggestion was blue and toward the yellow end if the suggestion was yellow. These

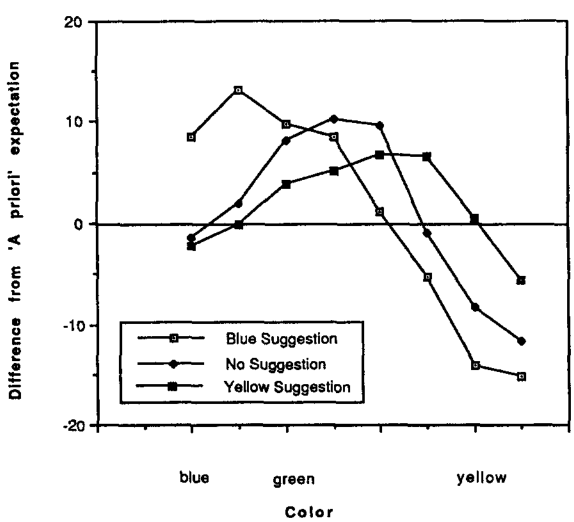

Figure 7. Belli's (1988) smoothed data showing the influence of episodic suggestions, with prior schematic knowledge subtracted out.

data do not distinguish between the separate- versus composite-trace models detailed in the last section of the article, but they certainly do not provide any counterevidence falsifying composite-trace models, as Schooler and Tanaka (1991) have claimed. The main value of these data appears to be as a reminder that both episodically presented events and a priori schematic expectations can influence later memory.

A Final Caveat

Distinguishing between separate-trace and composite-trace models is not as simple as the supposition that one model predicts compromises while the other does not. Rather, both models can be made to predict compromises but the conditions under which they will do so are different. The separate-trace model predicts compromise-like data when the original and misleading suggestion distributions show considerable overlap. The composite-trace model, on the other hand, does not predict a unimodal shift no matter how close the representations are in similarity, so long as the exact original and misleading colors are used at time of test. Another difference is that the composite-trace model predicts unimodal distributions when the choices at times of test are shifted slightly outside the range of the blended continuum, or if subjects' schematic expectations mimic such a shift. This shift manipulation should result in unimodal results even with a fairly broad difference in the color choices.

In conclusion, we face a situation in which our intuitions were doubly wrong. The composite-trace representation yields the compromise unimodal shift only under certain very restricted assumptions, whereas it is not difficult to derive unimodal shifts in group data from separate trace representations. It is not our goal here to advocate either the summed distribution idea in the separate-trace model or the displacement hypothesis in the composite-trace model. Our goal, rather, is to illustrate that the composite–compromise issue is in need of rigorous quantitative analysis; at the level of qualitative verbal arguments, we expect there will be little progress.

References

Anderson, J. A. (1977). Neural models with cognitive implications. In D. LaBerge & S. Samuels (Eds.), *Basic processes in reading* (pp. 27–90). Hillsdale, NJ: Erlbaum.

Belli, R. F. (1988). Color blend retrievals: Compromise memories or deliberate compromise responses? *Memory & Cognition, 16,* 314–326.

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–190). Washington, DC: Winston.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93,* 411–428.

Kortge, C. A. (1990). Episodic memory in connectionist networks. In *Proceedings of the Twelfth Annual Conference of the Cognitive Society* (pp. 764–771). Cambridge, MA: Kaufman.

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

McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel distributed processing* (Vol. 1). Cambridge, MA: MIT Press.

Metcalfe, J. (1990). Composite Holographic Associative Recall Model (CHARM) and blended memories in eyewitness testimony. *Journal of Experimental Psychology: General, 119,* 145–160.

Murdock, B. B. (1982). A theory of the storage and retrieval of item and associative information. *Psychological Review, 89,* 609–626.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing* (Vol. 2). Cambridge, MA: MIT Press.

Schooler, J. W., & Tanaka, J. W. (1991). Composites, compromises, and CHARM: What is the evidence for blend memory representations? *Journal of Experimental Psychology: General, 120,* 96–100.

Received September 24, 1990
Revision received November 16, 1990
Accepted December 5, 1990 ■