Blend Errors During Cued Recall

Leigh Erik Nystrom **James L. McClelland**

Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213
ln0c+@andrew.cmu.edu
jlm+@andrew.cmu.edu

Abstract

Connectionist models of memory account for recall behavior using processes which simultaneously access multiple memory traces and interactively construct the recalled information. This also allows the models to account for prototyping phenomena, but seems to predict retrieval of composite or "blended" information during ordinary recall. By contrast, models that simulate recall as a probabilistic selection of a single trace would not predict recall blend errors. To examine memory blending during recall, three experiments were performed; in each, subjects read sentences, some sharing words with one other sentence. They later recalled the sentences given partial-sentence cues. In all experiments subjects made blend errors, recalling one word from each of two similar sentences more often than one word from each of two dissimilar sentences, as predicted by multiple-trace models. The frequency of blend errors was relatively low, but a good account of this and other aspects of the results was provided by a multiple-trace model based on an Interactive Activation network as applied to memory retrieval in McClelland (1981).

Introduction

Much of the current interest in connectionist models of memory stems from their ability to capture not only recall and recognition behavior, but also the prototyping and generalization found in concept formation experiments (Knapp & Anderson, 1984; McClelland, 1981; McClelland & Rumelhart, 1985). In all of these models, memory retrieval processes can lead to prototyping because they involve an activation and synthesis of multiple memory traces to produce a pattern of activation that may not necessarily correspond to a single trace as originally stored.

For example, McClelland (1981) outlined a connectionist model of memory wherein each memory trace consists of one "instance" unit representing the trace as a whole along with "property" units for each of that trace's properties (see Figure 1). Property units within a trace reinforce each others' activation through the centralized instance unit, while inhibiting the activation of competing property units from other traces. During retrieval, traces in memory become active to the extent that they include the properties given as a recall cue, and to the extent that their properties reinforce one another. The inhibition between the property and instance units of different traces also influences the final state of activation as traces compete for activation of shared properties. Properties that ultimately remain active constitute the recalled information, regardless of whether they all occurred in the same trace, although mutual reinforcement of same-trace properties often leads to the retrieval of a single trace anyway. It would seem that this sort of "multiple-trace" access model would predict abundant errors during everyday recall -- "blend" errors, consisting of mixtures of more than one trace.

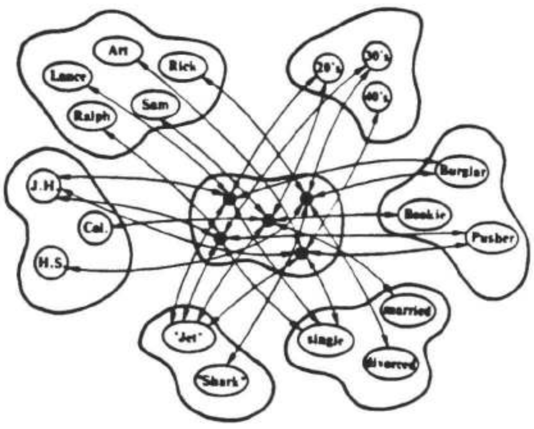

Figure 1: Multiple-trace model from McClelland (1981)

Not all theories of memory retrieval assume that recall involves synthesis of multiple memory traces. Some models view the recall process as a search through memory to find a single trace; most memory models before the 1980's had this character. Without the joint contributions of multiple traces, single-trace models do not automatically lead to prototyping or generalization -- and as a side effect, memory blends would be avoided. Shiffrin's SAM model of associative memory (Raaijmakers & Shiffrin, 1981; Gillund & Shiffrin, 1984) provides one example of a more recent single-trace model. In the SAM model, cued recall proceeds as a probabilistic selection of a single trace from the collection of all traces. The probability of selecting a given trace is a function of the baseline "strength" of the trace as well as the degree of match between the cue's properties and the trace's properties. Once selected, all properties of the trace are accessible, and no other traces will be accessed. Another single-trace model, Anderson's ACT theory (Anderson, 1976; 1983), represents related information as being subsumed under a single "trace" node; the goal of recall is the selection of one of the possible trace nodes, as a function of the activation of its subsumed

Credits

This research was supported in part by an NSF Graduate Fellowship awarded to the first author, and NSF Grant BNS 88-12048 awarded to the second author.

nodes. As with Shiffrin's model, once a single trace is selected, all of its constituent information can be retrieved without interference from other traces.

So the question arises: Do blend errors actually exist? Or more specifically, do they occur as several memory traces combine at the time of recall? While memory blends have been studied in the context of eyewitness testimony (Loftus, 1977; Loftus, Miller, & Burns, 1978), the procedure used to induce blending in those and other experiments has left open the possibility that subjects integrated multiple traces at the time of encoding rather than retrieval. Both single- and multiple-trace models can easily account for blending in such circumstances -- any traces blended during encoding must necessarily be retrieved as blended information. Therefore, in order to distinguish single- from multiple-trace models, the focus of investigation was directed solely to blending of similar memory traces <u>at the time of retrieval</u>. The present experiments sought to find evidence for or against the existence of blend errors using sentence materials based on the classic studies of Anderson (1976), with methods designed to minimize the possibility that subjects might integrate sentences during encoding.

Imagine the situation wherein two sentences share three out of five content words, and the three shared words are used as a recall cue. This cue would match the two overlapping sentences equally well. If a subject were asked to recall the other two content words from one of the sentences, they could correctly answer with the two remaining words from <u>either</u> overlapping sentence. For example:

#1:*"The doctor gave the plumber the coat in the lobby."*
#2:*"The doctor gave the plumber the watch in the kitchen."*
cue:"The doctor gave the plumber the ____ in the ____."

Assuming that each sentence is encoded as a single trace, with its constituent words as properties, a multiple-trace model would predict difficulty for a subject's recall, because the ambiguous cue would equally activate the traces for both of the sentences and their constituent properties. As a consequence, subjects may be prone to making a "cross-over" intrusion error by responding <u>*coat*</u> + <u>*kitchen*</u>, or <u>*watch*</u> + <u>*lobby*</u>, mixing words from each overlapping sentence. By contrast, a single-trace model would predict that either of the two traces matching the ambiguous cue would be accessed in a probabilistic search with equal likelihood, yet the words from only one of these two traces would be retrieved. Therefore, a single-trace model would predict no blending of these highly similar memory traces despite the ambiguous cues.

In accordance with the logic of this example, the following experiments used a list of sentences in which half of the sentences shared several content words with one other sentence, while the other half served as completely dissimilar controls. Subjects studied the sentences individually under the pretense of a cover task, unaware that they would need to later recall any sentences. They therefore had no reason to continually rehearse sentences or actively integrate the traces of overlapping sentences, thereby allowing the experiment to plausibly investigate blending during retrieval alone. After subjects learned the list of sentences, a delay task ensured some degree of temporal separation between encoding and retrieval. Finally, subjects were given cues which consisted of the words corresponding to the positions shared by overlapping sentences. They were asked to recall the two missing target words from a single sentence studied earlier in the experiment.

In all three of the experiments, two general results were of primary interest, each following a hypothesis of multiple-trace models concerning recall blends. The first hypothesis was that subjects should blend similar memory traces, as evidenced by more cross-over intrusion errors between the similar overlapping sentences than between the dissimilar control sentences. If no difference was discovered, the null hypothesis of single-trace models would be supported. The second hypothesis was that subjects should be less likely with overlapping than with control sentences to correctly recall a complete sentence after having already accessed its trace. This result would be expected because while the overlapping sentence traces would continually influence each other during the independent retrieval of the trace properties, the control sentence traces would be immune from such post-access interference. This hypothesis was tested by examining the conditional probability that a subject would correctly recall both target words given that they had already correctly recalled one of the words. If no difference was found in this conditional probability as a function of sentence similarity, then single-trace models would be supported. Such models account for interference effects during the search for a trace, but once one is accessed, other traces are assumed to be of no further relevance.

Experiment 1

Thirty-two experimental sentences were randomly created for every subject. Each sentence included five word-positions which were filled by random selection (without replacement) from a prearranged list of words of the appropriate semantic type. Care was taken to avoid words closely-related across semantic types, such as "barber" and "barbershop". Two sentences from each set of four were selected to become overlapping sentences, the other two left as matched control sentences. Three of the word-positions in each set were randomly chosen for the overlap; any three of the five were equally likely to be chosen for overlap. The three words in the overlap positions of one of the two overlapping sentences were then replaced with the corresponding words from the other sentence, thereby creating a pair of sentences differing by only two words. The order of presentation of the 32 sentences was randomized for each subject with the constraint that a minimum of 12 sentences intervened between any two matched overlapping or matched control sentences from a matched set of four.

The random generation of sentences created many unlikely combinations of sentence constituents, enabling the experiment to use a convincing cover story: subjects were told that the experiment was designed to explore how people read sentences that "sound strange," those in which the content words do not seem to belong together in real life. They therefore had no cause to believe that they would later need to remember anything about the sentences. Subjects read each sentence and made a judgement about the overall "plausibility" of the sentence, followed by judgements of how appropriate each of the five main words seemed, given the content of the rest of the sentence. First the sentence appeared along with a prompt to rate it as a whole, then each of the five words appeared sequentially underneath the sentence along with a prompt to rate it. The purpose of these ratings was both to make the cover story more convincing and to ensure that the subjects paid attention to every sentence and its primary words, forming memory traces through processing them at a "deep," semantic level.

After a 5-minute delay task, subjects were given recall cues containing all words from a previously-presented sentence except the two words in the non-overlapping word-positions (replaced by blanks of a fixed length). Presenting any one sentence from an overlapping pair in this manner therefore acted as a cue for either of the two paired sentences. Subjects were told that they may have seen two sentences that fit a cue equally well, but that in those cases, they would have to remember both of them, one at a time. However, they were never told that a particular test cue matched two sentences until they had first already recalled one sentence. It was stressed that in choosing each response, they should take care to recall two words from within the same sentence from the first phase. Therefore, with overlapping sentence cues, after the subject first recalled two words from one sentence, they were asked to recall two more words belonging to the other matching sentence. In this manner, all 16 of the overlapping sentences from the first phase were tested, along with 8 randomly-selected control sentences, all in a random order.[1]

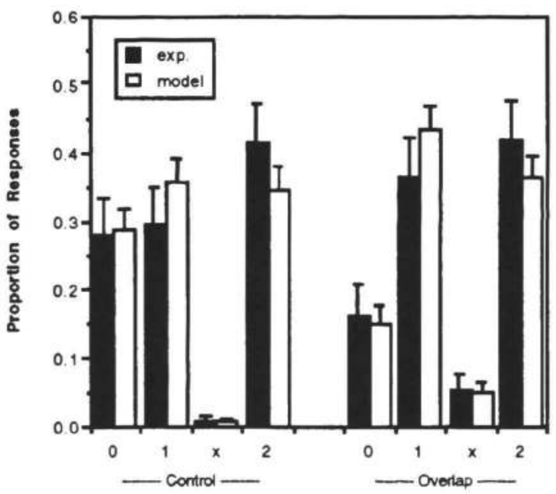

Figure 2: Experiment 1 Data

The overall recall performance of the 35 subjects can be seen in Figure 2. Subject data can be seen in the solid columns (error bars indicate 95% confidence intervals), illustrating the proportion of responses for both control and first overlap conditions as a function of the number of words correctly recalled. Results of the first hypothesis test supported a multiple-trace model: subjects made cross-over errors more often (using a sign test across subjects, $p < .01$) when tested in the overlap conditions (5.4% with the first, 2.1% with the second) than in the control condition (0.7%). The difference between the rates in the second overlap versus control conditions was not significant ($p > .05$) while the difference between the first and second overlap conditions was significant ($p < .05$). The greater frequency of intrusions between overlapping sentences was not simply a result of a greater tendency to recall words from overlapping sentences, because words from overlapping sentences intruded into recall of a matched control sentence only 1.4% of the time; words from control sentences intruded into recall of a matched overlapping sentence 2.7% of the time.

Results of the second hypothesis test also supported multiple-trace models: the conditional probability of recalling two words from the same sentence given the correct recall of one word was lower in the overlap condition ($p = .50$, first overlap; $p = .36$ second overlap) than with controls ($p = .58$). Using a Wilcoxon signed-ranks tests across subjects, the conditional probability in the control condition was significantly higher than in either the first overlap ($z = 1.81$, $p < .05$) or the second overlap conditions ($z = 4.11$, $p < .0001$); the conditional probability in the first overlap condition was significantly higher than in the second overlap condition ($z = 3.10$, $p < .001$).

Experiment 2

It was difficult to know how to best interpret the data from the two overlap conditions in Experiment 1. The fact that the first sentence recalled in the overlap condition is likely to be the stronger of the two, while the second may be contaminated by the earlier recall of the first, makes quantitative comparison of the overlap and control conditions difficult. Experiment 2 was designed to generate clearer data, by using unambiguous overlapping-sentence test cues and thereby allowing only one correct answer in all cases. This was accomplished by reducing the number of content words shared among overlapping sentences from three to two, while continuing to use cues with only two blanks. For example:

#1: *"The doctor gave the plumber the coat in the lobby."*
#2: *"The doctor gave the lawyer the watch in the kitchen."*
cue: "The doctor gave the plumber the ____ in the ____."

Note that this cue only matches the first overlapping sentence completely, while offering a partial match to the second sentence. With this experiment, we expected to replicate the findings of Experiment 1, perhaps with less overall blending due to the lessened overlap between sentences.

Thirty-six sentences were randomly created for every subject, 24 overlapping and 12 control. One of the overlapping sentences in each overlapping pair was also used as a mate for a control sentence, since it did not share any content words with the control sentence. The order of presentation of the 36 sentences was randomized for each

[1]Post-test questionnaires indicated that most subjects were unaware of sentence overlapping during the first phase of the experiment. Only two subjects reported that they had recalled or integrated overlapping sentences at the time of encoding; excluding their data from the analyses had no significant effect. These results held across all three experiments.

subject by presenting 12 control sentences and 12 overlapping sentences (the "overlap-test" sentences, one from each pair of overlapping sentences) as the first 24 sentences, followed by the rest of the overlapping sentences (the "overlap-distractors"). This allowed the overlap-distractor sentences to provide retroactive inhibition, both for the overlapping-test sentences and the control sentences. The order was then randomized with the further constraint that a minimum of 12 sentences intervened between overlap-test or control sentences and their paired overlap-distractor sentence. The 24 test cues contained all but two words from an overlap-test or control sentence, and were presented in a completely random order.

Overall, the 38 subjects performed worse than those in Experiment 1 (see Figure 3.) Perhaps the lower overall level of performance in this experiment indicates that subjects had too difficult a time recalling the earlier of the two overlapping sentences due to the recency of the overlap-distractor sentence. In fact, fully 9% of subjects' answers with overlap-target cues consisted of recalling both words from its paired overlap-distractor sentence -- a sort of double cross-over intrusion -- despite the fact that the cue unambiguously matched the overlap-target sentence.

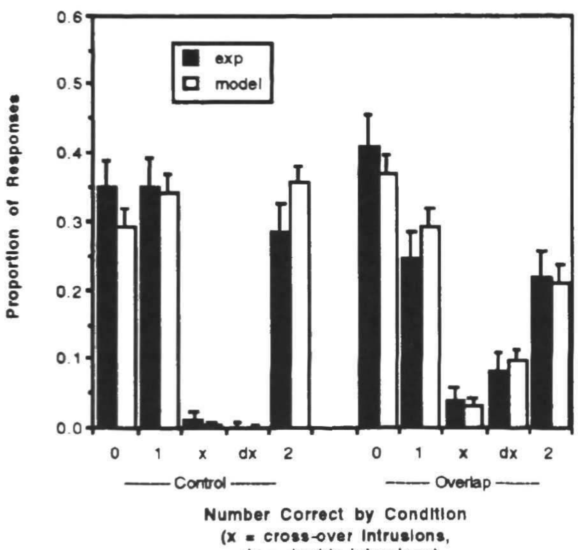

Figure 3: Experiment 2 Data

Results of the first hypothesis test showed that cross-over errors were more frequent between matched overlapping sentences (18 occurrences, or 4.0%) than between overlap-distractors and their matched controls (6 occurrences, 1.3%). As in Experiment 1, the difference between the cross-over error rates of the overlapping and the control sentences was significant (sign test, $p < .05$), as predicted by multiple-trace models.

The test of the second hypothesis supported single-trace models, however. The conditional recall probability was essentially the same in both conditions: .43 with overlapping-test sentences, and .44 with control sentences ($z = 1.07, p > .05$). This pattern of results would be expected if there were no post-access influences from similar memory traces.

Experiment 3

One potential critique of Experiments 1 and 2 concerns the evaluation procedure used in the study phase of the experiments. Subjects had been asked to rate the plausibility of a whole sentence, and then of each individual content word; this latter requirement could conceivably have caused subjects to encode a separate individual memory trace for each content word as well as for the whole sentence. If this was the case, both multiple- and single-trace models might explain the conflicting pattern of results by proposing that subjects were occasionally able to exploit associations between traces for individual words along with more complete traces for the sentences containing them. Therefore, Experiment 3 was performed, using the same materials and procedure as Experiment 1, but with the elimination of subjects' ratings of individual words during the study phase.

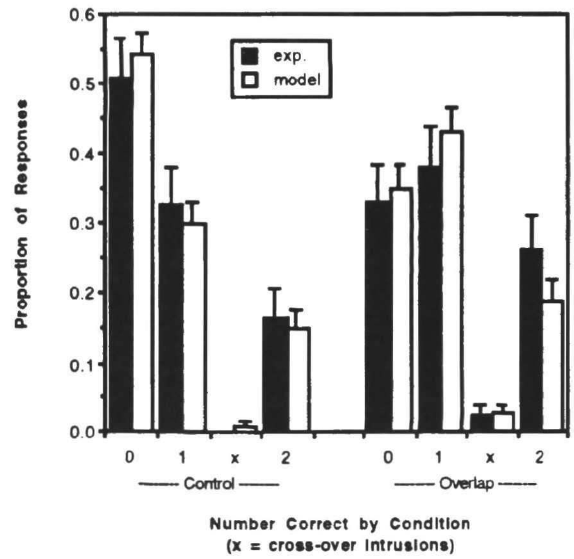

Figure 4: Experiment 3 Data

Overall, the 37 subjects performed worse than the subjects in the other experiments, no doubt because the new procedure led subjects to encode the stimuli less rigorously (see Figure 4). No cross-over errors were obtained between control sentences. Consequently, cross-over errors were significantly more frequent (sign test, $p < .01$) in the first overlap condition (2.4%) than in the control condition. They were also more frequent ($p < .05$) in the second overlap condition (1.7%) than in the control condition. There was no significant difference ($p > .05$) between the cross-over error rates in first and second overlap conditions. These differences again provided support for a multiple-trace model.

The conditional recall probabilities did not significantly differ ($z = -1.16, p > .05$) between the first overlap condition ($p = .35$) and the control condition ($p = .31$). On the other hand, the difference between the conditional probability in the control and second overlap ($p = .16$) conditions was significant ($z = 2.78, p < .005$), as was the difference between the first and second overlap conditions ($z = 2.93, p < .005$).

Models

In all of the experiments, cross-over errors were more frequent between overlap sentences than controls, supporting the first hypothesis of multiple-trace models that similar sentence traces should blend during recall. On the other hand, the pattern of conditional recall probabilities, indicating the presence or absence of post-access interference between sentences, did not consistently support the second hypothesis of multiple-trace models. However, one important concept has been missing from the discussion of the models' predictions thus far: encoding failure. Subjects did not have perfect memory for every sentence in any of the experiments, considering the rate of 50% or more recall failure. In any model of memory, an item cannot be recalled if it was improperly encoded in the first place. Yet the models as discussed so far have simply assumed that subjects completely encoded all sentence information into well-integrated traces. Therefore, one needs to examine how both multiple- and single-trace models perform when a significant proportion of the properties in traces are either missing or too weak to be recalled.

How would a multiple-trace model deal with encoding failure? Because of the simultaneous access of all relevant traces in a multiple-trace model, whenever one of the two target words in an overlapping sentence had not been adequately encoded, the corresponding word from the other overlapping sentence trace should automatically be retrieved instead, as long as it had been properly encoded itself. However, multiple-trace models were already predicting blend errors when encoding failure was not considered; wouldn't they now predict entirely too many blend errors relative to the fairly low frequency obtained in the experiments?

In order to discover whether or not a multiple-trace model could exhibit behavior conforming to that of our subjects, we devised an interactive activation model based on the simple connectionist model discussed in McClelland (1981). This model was adapted for the current experiments by assuming that the learning of each sentence created a instance unit for the sentence trace as a whole, linked with bidirectional excitatory connections to five property units, one for each major content word in the sentence. All of the instance units together formed a "pool" of units, as did all of the property units corresponding to a given sentence position. All units within a pool were linked by bidirectional inhibitory connections (using negative weights), reflecting the fact that the units within them represent mutually-exclusive information that should not be recalled simultaneously.

Because the overlapping sentences shared words, the corresponding units in the property unit pools would also be shared. Simulation of recall proceeded by activating the property units representing the recall cue words. In the case of an overlapping sentence, activation from the cued units would be sent along excitatory connections to the instance units for both overlapping sentences. The two instance units would compete for activation, due to their inhibitory connections; they would also feed back activation to the cue words, as well as the target words. In the pools containing the target words, further competition would result from the activation of the mutually-inhibitory units. Eventually, the activation levels of all units will stop fluctuating, after the network has "settled", and the units with the highest activations can be chosen as the response. Note that if the cue had been words from a control sentence then the correct answer would be retrieved without competition, because activation would flow to only one word per pool.

Actually, as outlined above, the model would not be able to settle the competition between overlapping sentence words, because their inputs and resulting activations would always be equal. Based on this, the model led us to expect that our experiments would generate more blend errors than were actually obtained. However, there is a general problem with the interactive activation model that has recently been uncovered in other applications which appears to reflect a failure to take account of the important role of inherent variability in processing (Hinton & Sejnowski, 1986). Once this variability is introduced to the model, several difficulties are resolved (McClelland, 1991). In the present application, adding variability causes the model to tend to favor one of two complete sentence traces rather than a blend; the blend states represent less-optimal points in the "goodness" landscape of network states (c.f. Rumelhart et. al., 1986), and variability allows the network to escape such local minima. McClelland (1991) indicates that variability may be introduced in a variety of ways. In this case, we simply injected a small amount of normally distributed random noise into the input to each unit at each update.

The addition of intrinsic variability allowed the model to produce approximately correct blend rates, but tended to produce too large a difference in the probability of completely correct recall between the control and overlap conditions. (This is equivalent to over-predicting the difference in the conditional probability of recalling both words). The addition of the assumption of occasional encoding failures allowed the model to overcome this problem. To capture encoding failure, we assumed that each link between a property unit and its associated instance unit had a 20% chance of being absent entirely. In the experimental data, subjects' errors indicated that there were some associations between words from completely different sentences; these associations were captured by inserting randomly determined, non-negative weights ($M = 0.2$, $SD = 0.25$) between all word and instance units. The weights on connections within pools of word units were all set to -1.0, while those within the instance unit pool were stronger (-2.1), to counterbalance the activation coming from the three word units and ensure adequate competition between sentences.

The model networks for each experiment reflected the materials used in that experiment. For Experiments 1 and 3, the network therefore contained of a pool of 32 instance units, two pools of 32 target word units, and three pools of 24 word units (16 control words + 8 overlap words per pool). For Experiment 2, the network consisted of 36 instance units, two pools of 24 overlapping cue units (12 overlap + 12 control), one pool of 36 non-overlapping cue units, and two pools of 36 target units. Because the differing results from Experiments 1 and 3 were thought to have been a result of less rigorous encoding in Experiment 3, the only difference between the networks for Experiments 1 and 3 rested in increasing the model's encoding failure rate for Experiment 3 from 20% to 35%. To test recall, external activation was input to the three cue word units for each of the test sentences. The response of the computer subject was simply taken as the unit with the highest activation above a response threshold (.1) in each target word pool after 100 time cycles, enough time to allow the network to settle into an equilibrium and form a reasonable response hypothesis. A total of 100 computer subjects were run for each computer simulation; with each computer subject, a new random set of stimulus and random association weights was generated. The performance of each network was fit to the data as shown in Figures 2 through 4. In so doing, we ensured that the model not only had the correct overall recall rates, but also the appropriate performance with respect to both hypothesis tests – blend rates and post-access interference rates.

In the data fits illustrated in Figures 2 to 4 (in which error bars reflect 95% confidence intervals), the model's performance was not statistically different from that of our human subjects (for Exp.1, $\chi^2(7) = 7.0$, $p = .42$; for Exp.2, $\chi^2(9) = 15.9$, $p = .07$; for Exp. 3, $\chi^2(7) = 11.2$, $p = .13$). However, the fit comes close to failing significantly in two of the three cases. We were able to obtain even closer data fits by increasing the inhibition between instance units in Experiments 1 and 3 (to -2.5) and decreasing it for Experiment 2 (to -2.0). Thus, with the use of an additional free parameter one can achieve a nearly perfect correspondence between the performance of the model and experimental subjects (for Exp.1, $\chi^2(7) = 1.8$, $p = .97$; for Exp. 2, $\chi^2(9) = 2.5$, $p = .98$; for Exp.3, $\chi^2(7) = 3.0$, $p = .89$). It is not entirely clear why that particular parameter should differ between experiments, though it might reflect subjects' ability to strategically increase competitive inhibition between equivalently-cued overlapping traces in Experiments 1 and 3.

Given the success of our particular multiple-trace model, it is worth considering how a single-trace model would behave given similar assumptions about encoding failure. A single-trace model like Shiffrin's SAM model might also fit the data if multiple access attempts were allowed in cases where an initial trace selection resulted in incomplete recall. Successive attempts proceed just like the first, with the probability of selecting a given trace based on that trace's similarity to the cue. Since overlapping sentence traces equally match the cue, they have an equally good chance for selection. So if a first attempt accessed an overlapping sentence trace containing only one of the two target words, and a second attempt accessed the other overlapping sentence, the subject could make a cross-over error. It seems plausible that an account could be given for all aspects of the data based on such a possibility. However, this kind of extension of the single-trace model in some sense amounts to turning it into a multiple-trace account. The primary difference between this and our account is simply that the multiple traces are accessed in succession in one case and simultaneously in the other.

Conclusion

In the end, what has been learned here about the nature of memory blending during recall? The present experiments and simulations have changed the way we approach multiple-trace models in general, as we have abandoned the original notion of unintended blending of complete traces in favor of the somewhat less dramatic notion of filling in missing parts of incompletely encoded traces.

What remains to be studied is how other instantiations of a multiple-trace model, such as the more superpositional matrix- or convolution-based models (for example, Knapp & Anderson, 1984, or Metcalfe, 1990), would manage the task of simulating the specific patterns of recall performance found in our experiments. It will be of interest to discover whether these types of models could produce as few blend errors as were obtained in our studies, given their tendency to blur the distinctions between traces. Conversely, it will be interesting to examine whether the prototyping and generalization properties of these other models can be captured with the model we have presented here. These complementary studies may take us some distance toward understanding what governs the human ability to generalize well and yet preserve relatively distinct access to particular prior events.

References

Anderson, J. R. (1976). *Language, Memory, and Thought.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. (1983). *The Architecture of Cognition.* Cambridge, MA: Harvard University Press.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psych. Review*, **91**, 1-65.

Hinton, G. E., & Sejnowski, T. J. (1986). Learning and relearning in Boltzmann machines. In Rumelhart, D. E. McClelland, J. L. & the PDP Research Group. *Parallel Distributed Processing*, Vol. 1. Cambridge, MA: MIT Press.

Knapp, A., & Anderson, J. A. (1984). A signal averaging model for concept formation. *JEP: Learning, Memory, and Cognition*, **10**, 616-637.

Loftus, E. F. (1977). Shifting human color memory. *Memory & Cognition*, **6**, 696-699.

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

McClelland, J. L. (1981). Retrieving general and specific information from stored knowledge of specifics. *Proceedings of the Third Annual Meeting of the Cognitive Science Society*, 170-172.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *JEP: General*, **114**, 159-188.

McClelland, J. L., & Rumelhart, D. E. (1988). *Explorations in parallel distributed processing.* Cambridge, MA: MIT Press/Bradford Books.

McClelland, J. L. (1991). Stochastic interactive processes and the effect of context on perception. *Cog. Psych.*, **23**, 1-44.

Metcalfe, J. (1990). Composite holographic associative recall model (CHARM) and blended memories in eyewitness testimony. *JEP: General*, **119**, 145-160.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psych. Review*, **88**, 93-134.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In McClelland, J. L., Rumelhart, D. E. & the PDP Research Group, *Parallel Distributed Processing*, Vol. 2. Cambridge, MA: MIT Press.