In: Proceedings of the 23rd Annual Conference of the Cognitive Science Society, London: Erlbaum, 2001, pp. 510-515

Episode Blending as Result of Analogical Problem Solving

Boicho Kokinov (bkokinov@nbu.bg)[12]
Neda Zareva-Toncheva (nzareva@cogs.nbu.bg)[2]
[1]Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. G. Bonchev Street, bl.8
Sofia 1113, Bulgaria

[2]Central and East European Center for Cognitive Science, Department of Cognitive Science and Psychology,
New Bulgarian University, 21 Montevideo Street
Sofia 1635, Bulgaria

Abstract

We know that misinformation presented in interrogating questions or in advertising produces blendings, that even imagining a possible episode might produce blending as well, however, we do not know whether reasoning and problem solving can produce the same effect. On the other hand, models of analogy-making assume "perfect memory" for old episodes. The AMBR model of analogical problem solving has mechanisms for interaction between memory and reasoning which explain partial memory and memory distortions and has predicted blending effects which are due to the reasoning process. Such predictions have no parallel in any other model we know of. There has been no experimental support for these predictions so far. The current paper describes an experiment explicitly designed to test these predictions. It consists of three sessions: 1) solving three problems, 2) solving two more target problems by analogy with some of the problems in the first session, and 3) reproduction of the three problems in the first session. The results demonstrate that the degree of blending in the recalled stories depends on the target problem solved in the second session.

Motivation

There is a considerable amount of research demonstrating various types of memory distortions, such as schematisation, blending, and false memory illusions (Bartlett, 1932; Loftus, 1977; Loftus, 1979; Loftus, Feldman, & Dashiell, 1995; Loftus, Miller, & Burns, 1978; Loftus & Palmer, 1974; Moscovitch, 1995; Neisser, 1998; Nystrom & McClelland, 1992; Reinitz, Lammers, & Cochran, 1992; Schacter, 1995). These findings, however, have been established in "pure memory tasks" conditions. Very rarely researchers have tried to integrate memory tasks with other kinds of cognitive tasks in order to explore whether there will be interactions between them. Thus for example, we know that misinformation presented in interrogating questions or in advertising produces blendings, that even imagining a possible episode might produce blending as well, however, we do not know whether reasoning and problem solving can produce the same effect.

At the theoretical end these findings are typically explained by the constructive nature of human memory, however, very few models exist that lay out the specific memory mechanisms that could explain the memory distortion effects. The key concept has always been the postulation of distributed representations of some type in human memory. These include Hintzman's (1988) multi-trace model, Metcalfe's (1990) CHARM model, and McClelland's (1995) PDP-type of model. The last two models explicitly deal with memory blending effects.

On the other hand, models of human analogical reasoning which necessarily include a memory component (since they have to explain how analogous episodes are retrieved) simply ignore both the experimental findings about memory distortions and the theoretical ideas about the constructive nature of human memory. They typically assume "perfect" memory for past episodes and even some kind of nice organization of memory that would allow the retrieval of the relevant episode (Forbus, Gentner, & Law, 1995; Hummel & Holyoak, 1997; Kolodner, 1984; Thagard, Holyoak, Nelson, & Gochfeld, 1990; Warton, Holyoak, & Lange, 1996). All these models assume centralized representation of episodes which means that episodes are either retrieved as a whole (and than mapped onto the target problem description) or they fail to be retrieved. No blending of episodes may occur, no false memories can arise, no partial retrieval can happen. Surprisingly, this is true even for the LISA model (Hummel & Holyoak, 1997) which is based on distributed representations, since the representations are truly distributed only in working memory, but the episode representation in LTM is highly centralized, although distributed in some sense (in the same sense in which the representation is distributed in the model described in the current paper). A list of all units representing a single episode is assumed and if the episode wins the competition between episodes then all it corresponding units are switched from dormant to active state.

The main motivation for recent developments in the AMBR model of analogical problem solving (Kokinov, 1998; Kokinov & Petrov, 2000, 2001; Petrov & Kokinov, 1998, 1999) was to propose mechanisms for interaction between memory and reasoning which will allow for explaining memory distortions and blendings among other things. AMBR has predicted blending effects which are due to the reasoning process. Such predictions have no parallel in any other model we know of. However, we do not know of any experimental support for them either. That is why an experiment was designed to test these predictions. The current paper describes this experiment and its outcomes.

AMBR Model and its Predictions

The AMBR model of analogical problem solving was introduced in (Kokinov, 1988) and then further developed over the years (Kokinov, 1994a, Kokinov, 1998, Kokinov & Petrov, 2001). It is based on a general cognitive architecture called DUAL (Kokinov, 1994b,c) which relies on emergent computations produced by a society of micro-agents (Minsky, 1986; Hofstadter, 1995). The latest developments (Kokinov, 1998, Kokinov & Petrov, 2000, 2001; Petrov & Kokinov, 1998, 1999) are directed towards building a more psychologically plausible memory for episodes and thus allowing for memory distortions to take place. Moreover, a prediction is made that the reasoning process during the analogical problem solving may itself produce memory distortions.

Episode Representation

Episode representation in AMBR is highly decentralized, which means that each episode is represented by a large coalition of micro-agents each of them representing some aspect of the situation. We may call this representation "distributed at a higher level" by analogy to the connectionist representations which are distributed over a set of simple features. In reasoning, and in analogy-making in particular, the relations are even more important than single features and therefore they have to be explicitly represented. The agents here are more complex than the connectionist units (Kokinov, 1994a, b, c) and roughly speaking they represent a whole proposition (such as "there is a coffee pot", "the coffee-pot is made of metal", "the coffee-pot is on the plate", "the plate is hot", "the water is in the coffee-pot", "proposition 1 causes proposition 2", etc.). Thus the episode is represented by a set of such propositions (by a coalition of the corresponding micro-agents), but there is nowhere in memory a list of all propositions (all micro-agents) involved in a given episode. That is why we call this representation decentralized. There is one agent which refers to the unique point in time and space – when and where the event has happened, and all agents from a given coalition have links directed to it. In this way the system can differentiate among propositions belonging to different episodes, however, there are no links from this agent to the agents in the coalition, i.e. it does not know any of the members of the coalition. Therefore this agent cannot contribute to the retrieval or construction process.

Each agent has a level of activation that changes dynamically. The activation level determines the degree to which the information contained in that agent is available and the degree to which this agent participates in the computational process (Kokinov, 1994b,c, Petrov & Kokinov, 1999). Working memory is considered to be the active part of long term memory, i.e. the set of all active agents at a given moment. There is a process of spreading activation as well as a process of decay which guarantees that an agent that do not receive activation will soon quit WM.

Episode "Retrieval" or "Construction"

Episode "retrieval" corresponds to the process of activation of the agents representing various aspects of the event and bringing them into WM. This means that typically the recall is only partial since there is no way to guarantee the activation of all members of a coalition. Some coalitions will be stronger (having stronger links between the agents) and therefore the corresponding episode will tend to be reproduced more fully, other coalitions are weaker and only few aspects of the episode are reproduced.

However, many agents belonging to other coalitions will also turn out to be activated – agents representing some general knowledge (concepts, facts, rules, etc.), agents representing aspects of other episodes. Thus the set of agents happened to be in WM will produce a description of the episode which is partial, but also containing intrusions from general knowledge and other episodes. Intrusions from other episodes are in fact blendings between two or more episodes. In fact, the representation of an old episode is not just retrieved from LTM, but is actively constructed in WM. The process of spreading activation is an automatic one but it depends on the current state of WM, the goals of the system and its input from perception. Thus the reasoning process, which runs in parallel to the memory process, interferes with this process of episode construction and in fact even guides it to a certain extend.

Mapping Guidance in Episode Construction

In the context of analogical problem solving the construction of the old "retrieved" episode is guided by the mapping process between this episode and the current target problem. The mapping process in AMBR does not start after the old episode is retrieved as in all other models of analogy-making, but runs in parallel to it. This makes it possible the already established partial mapping to guide the episode construction in such a way that the old episode is reconstructed in directions which allow better alignment between base and target. For example, the intrusions from general knowledge and from other episodes will be not arbitrary, but will correspond to elements of the target description which do not have corresponding elements in the base description (they are either missing in the encoding of the episode or are simply not activated at the moment). The precise mechanisms for this episode extension are described elsewhere (Kokinov & Petrov, 2000). A simulation experiment with AMBR has demonstrated that the parallel run of reasoning and "retrieval" in this model yields the retrieval of structurally similar episodes that would otherwise be not retrieved (Petrov & Kokinov, 1998).

AMBR's Prediction

AMBR's prediction relevant to the current paper is that intrusions from other episodes will happen more often if the currently constructed representation of the most active episode is missing elements which are important for the mapping process with the target problem. This emphasizes not any missing element, but elements which are crucial for the mapping. This is in contrast to a model that is very similar in flavor (Nystrom & McClelland, 1992; McClelland, 1995) which, however, is not sensitive to the structure but fills in any missing information. The latter model explains data from a memory experiment on sentence recall and is not intended to explore the relation between memory and reasoning. Structure might be unimportant in this case. In contrast, AMBR makes stronger predictions about the relevance and structural consistency of the intruded elements.

In this case the analogy-making process itself produces blending under certain conditions. Therefore the more partial the mapping between the target and the base problem is the more intrusions will occur and thus higher degree of blending between episodes will be observed. This is the prediction tested in the following experiment.

Experiment

The main idea of the experiment is the following one. Ask the participants to solve several base problems and as a side effect to hold them in their long-term memory. Half of the participants will then solve one target problem and the other half another target problem. After that we will ask the participants to retell us the base problems as accurate and complete as possible. We will measure the degree of blending between the problems and expect that it will depend on which target problem has been solved.

Method

Design

The experiment consists of three sessions:

- Session 1: solving three base problems (A, B, C);
- Session 2: solving one of two target problems (T1 – partially analogous to A and partially analogous to B, or T2 – partially analogous to B and partially analogous to C);
- Session 3: recalling the problems from session 1.

The whole trick is that target problem T1 partially maps to both A and B, while target problem T2 partially maps to both B and C. That is, in order to solve problem T1 one needs to make a double analogy (with A and B) and use two different principles which are provided in A and B respectively. This requires that A and B are both partially "retrieved" in WM and partially mapped to T1. We may describe this situation as blending the two episodes A and B and constructing a new "old" episode AB that is then remembered. Thus later on in session 3 we would expect that participants who solved problem T1 will tend to blend episodes A and B. Reversibly, participants who solved target problem T2 in the second session will tend to produce more blends of B and C in session 3.

Thus we use a between group design. The independent variable is the type of target problem solved in session 2 and its relation to the base problems. The dependent variables have to measure the blending occurring in the retold stories in session 3. We use two types of measures:

- a binary variable – "yes/no" expert judgments of blended memories;
- degree of blending (a value between 0 and 1) – measured as the degree of mixture between statements related to each of the base problems A, B, and C.

Material

The problems used in the experiment both in session 1 and in session 2 have been designed to fulfill several criteria:

- they should be solved by some general principles that can then be applied to another problem;
- if the principles used for solving the base problems A, B, and C in session 1 are called PA, PB, and PC respectively, then the target problems in session 2 should be solved by a combination of two principles (T1 by PA and PB, and T2 by PB and PC).

In *problem A* we used a criminal story about an attempt to come in for money of a wrong person who actually killed the legatee and dressed and acted like her. She imitated successfully even the gestures of the dead relative, including her habit to arch her right eyebrow when asking questions. She practiced these gestures for a long period of time in front of a mirror. Finally, however, she was recognized as a fake legatee. The question is how she was recognized. And the correct solution relies on principle PA: "Left and right are reversed in an mirror image". Thus, the lady arched her wrong eyebrow.

Problem B was an expanded version of the radiation problem. The principle underlying the correct solution was PB: the convergence of several weaker X-rays in one point form a stronger X-ray.

Problem C involved baking 3 flat loafs in a small baking tin which can hold only 2 loafs. The question was what is the minimal time period required to bake the 3 loafs turning each of them on both sides. And the principle underlying the correct solution was PC which outlines a turning schema: you first bake one side of two loafs, then you turn one of them and replace the second loaf with the third one, and finally you bake the remaining sides of the second and the third loafs.

All three problems were told as folk tales with some superficial similarities in the plot (having kings, princesses, wise man, etc.), still they were quite different.

Target problem T1 was about a five-headed dragon that has to be killed, but he can be killed only if at the very same moment his 2nd head from the left to right is cut off and his heart is destroyed by a strong laser beam. However, there were several obstacles: you cannot look at the eyes of the dragon directly because you will become blind, and also there were only 3 weak laser beams available. Thus the participants had to apply principle PA and to cut off the 2nd right head (instead of the 2nd left one) staying backwards and looking at the dragon into a mirror and principle PB to use a converging configuration of three weak laser beams.

Target problem T2 involved killing another dragon where again his heart should be destroyed by a strong laser beam and you have only three small ones, however, here the dragon was behind a moat full of lava which could be passed only by using three magic stones. The obstacle is that you can use each stone only once to step on each of its two sides. The solution involved principle PB from above, and principle PC used in problem C – the particular scheme of turning the stones and loafs respectively.

Procedure

During the first session the participants were told that they will undergo a series of experiments on human thinking and they will have to solve various problems. The problems were given one by one without an explicit time restriction. After the participants produced a written solution the experimenter read it aloud and if this was not the targeted solution she encouraged them to find an alternative solution, if this did not help a hint was given, and finally if the target solution was not found it was provided by the experimenter. The aim was all participants in session 1 to solve the base problems correctly and to acquire the basic principles PA, PB, and PC.

The second session followed after a period of 3 to 7 days. During this second session the group was split and half of the subjects solved problem T1, and the other half – T2. The second session was run again individually and the thinking aloud method was used, the speech of the participants was recorded. No hints were provided here.

The third session followed immediately after the second session. During this session the participants were asked to retell as accurate and complete as possible the problems from the first session. The stories were reproduced orally and tape-recorded.

Participants

48 undergraduate students participated in the experiment, but only 33 went through all sessions. 16 were female and 17 – male.

Results and Discussion

The records were transcribed and the protocols of the third session were used as the main data set. Each story was segmented into short phrases which express independent and understandable statements. The texts of the problems A, B, and C were also segmented into separate statements and their appearance in the body of the protocol was encoded. For example, we separated the text of problem C into 22 statements – C1-C22 and whenever a phrase (or its semantic equivalent) occurred in the narration of the subject the corresponding Ck was inserted in the protocol encoding.

Quite often when reproducing one of the problems participants inserted statements from one of the other two problems. This was exactly what we were counting. Thus for measuring the AB blending (blending between problem A and B) we counted how many As and how many Bs we have in the reproduction. The degree of blending was calculated as the ratio: number of As over number of Bs (when the As are less than the Bs), or reverse – number of Bs over number of As (when the As are more than the Bs). To put it differently we measured the percentage of intrusions in the text arising from another problem. Thus if the number of As or Bs are zero than no blending has occurred (degree of blending is 0), and when the number of As and Bs are equal then an absolute blend has been produced (with degree of blending equal to 1). The results are shown in Table 1 and Figure 1. As we can see the results are coherent with our hypothesis: we have higher degree of blending of type AB in group 1 (solving T1 in the second session) and higher degree of blending of type BC in group 2 (solving T2 in the second session). Just to remind that T1 required double analogy with A and B, and T2 required a double analogy with B and C. At the same time there is no difference whatsoever in the degree of blending of type AC which should be expected since none of the target problems required combining base A and base C. The performed analysis of variance showed a significant 2-way interaction between the groups (target problems solved) and the type of blending occurred ($F(2,62)=4.41$, $p<0.016$). The difference in the degree of AB blending is significant, but the difference in the degree of BC blending is not significant. Analyzing our data we found out that very few of the participants in group 2 were able to solve target problem T2 (in fact only 4 out of 16) and therefore they have not done the double analogical mapping with B and C. Thus we cannot expect the blending effect in this case. In contrast, problem T1 was solved by 11 out of the 17 participants. Evidently the second target problem was too difficult. We looked into the solved/unsolved difference and found support for this interpretation. The mean degree of BC blending in group 2 is: 0.148 for the subjects who solved T2 and 0.025 for the subjects who did not solved it, i.e. it is 6 times higher for the subjects who solved the target problem T2. The difference is significant at p-level 0.028 (measured by one-tailed t test, $t=2.45$)

Table 1: Mean degrees of blending in each group.

	# of subjects	AB blending	BC blending	AC blending
group 1 (solving T1~AB)	17	0.102	0.043	0.036
group 2 (solving T2~BC)	16	0.022	0.056	0.035

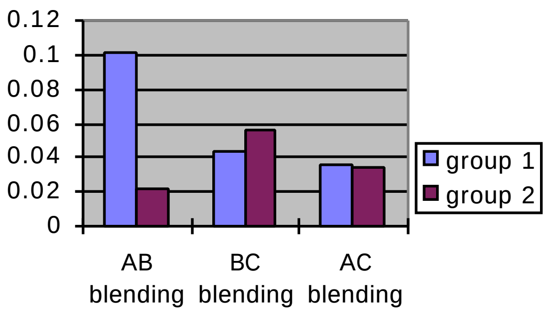

Figure 1: Degree of blending between the base problems reproduction depending on experimental group. There is a significant 2-way interaction between groups and type of blending.

Since we used a very formal method of measuring blending – the number of intrusions as registered in the protocols – we were curious to compare this to a more qualitative judgment done by human experts who may recognize whether there is a real blending or just some general knowledge intrusions or superficial mixture. Two independent judges had to read each protocol (without knowing neither about our hypothesis nor which group this protocol comes from). The experts had to judge whether there was a blending between some old problems. There was a high degree of agreement between the experts (about 10% disagreement where a third expert was called for judgment). The frequencies of blending of type AB, BC, and AC are presented in Table 2 and their percentage in Figure 2.

Table 2: Number of blendings as judged by experts.

	# of subjects	AB blending	BC blending	AC blending
group 1 (solving T1~AB)	17	11	6	5
group 2 (solving T2~BC)	16	5	8	5

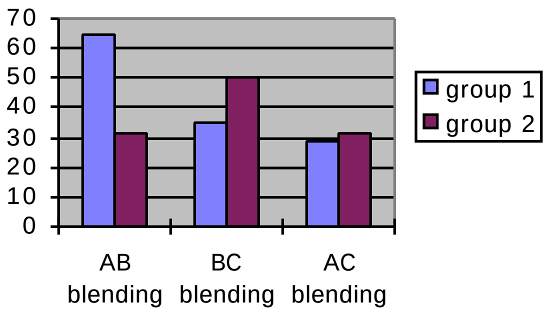

Figure 2: Percentage of blendings as judged by experts.

The results from expert judgments are generally coherent with the measurements of the degree of blending and again we have more AB blends in group 1, and more BC blends in group 2. The number of AC blends is almost equal in both groups. The log-linear analysis declared all the differences insignificant, however.

Conclusions

AMBR mechanisms of memory access and mapping work in parallel and interact with each other and thus predict mutual influence between these processes. One specific prediction is that when the target problem maps only partially with two different bases and a double analogy is required to solve it, then both bases are partially activated and elements from both episodes are brought into WM. In that way a blend between the two old episodes is produced and remembered. That is why a higher degree of blending is expected after subjects having solved such targets requiring double analogies.

This prediction has been experimentally tested in a series of three sessions in which the participants had first to learn the bases (by solving the problems in session 1), then to solve one additional target problem (session 2), and finally to retell the base stories (session 3). It turns out that higher degree of blending between two episodes is observed in the group where the target problem in session 2 required usage of a double analogy with these two episodes.

We consider these results as a first step in a series of experiments which have to test this hypothesis. We are currently running several more experiments varying the material, the design of the experiment and the timing of the sessions.

Acknowledgments

We would like to thank the AMBR research team for their continuous support and stimulating environment.

References

Bartlett, F. (1932). *Remembering.* Cambridge: Cambridge University Press.

Forbus K., Gentner D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science,* 19, 141-205.

Hintzman, D. (1988). Judgements of frequency and recognition memory in a multiple-trace model. *Psychological Review, 95*, 528-551.

Hofstadter, D. and the Fluid Analogies Research Group (1995). *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thoughts*. New York: Basic Books.

Hummel, J. & Holyoak, K. (1997). Distributed representation of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427-466.

Kokinov, B. (1988). Associative memory-based reasoning: How to represent and retrieve cases. In T. O'Shea and V. Sgurev (Eds.), *Artificial intelligence III: Methodology, systems, applications*. Amsterdam: Elsevier.

Kokinov, B. (1994a). A hybrid model of reasoning by analogy. In K. Holyoak & J. Barnden (Eds.), *Advances in connectionist and neural computation theory: Vol. 2. Analogical connections* (pp. 247-318). Norwood, NJ: Ablex

Kokinov, B. (1994b). The DUAL cognitive architecture: A hybrid multi-agent approach. *Proceedings of the Eleventh European Conference of Artificial Intelligence*. London: John Wiley & Sons, Ltd.

Kokinov, B. (1994c). The context-sensitive cognitive architecture DUAL. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kokinov, B. (1998). Analogy is like cognition: Dynamic, emergent, and context-sensitive. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences.* Sofia, Bulgaria: NBU Press.

Kokinov, B. & Petrov, A. (2000). Dynamic Extension of Episode Representation in Analogy-Making in AMBR. In: *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Kokinov, B. & Petrov, A. (2001). Integration of Memory and Reasoning in Analogy-Making: The AMBR Model. In: Gentner, D., Holyoak, K., Kokinov, B. (eds.) *The Analogical Mind: Perspectives from Cognitive Science,* Cambridge, MA: MIT Press.

Kolodner, J. (1984). *Retrieval and organizational strategies in conceptual memory.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Loftus, E. (1977). Shifting human color memory. *Memory and Cognition, 5*, 696-699.

Loftus, E. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press

Loftus, E., Feldman, J., & Dashiell, R. (1995). The reality of illusory memories. In D. Schacter (ed.), *Memory distortions: How minds, brains, and societies reconstruct the past*. Cambridge, MA: Harvard University Press.

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

Loftus, E. & Palmer, J. (1974). Reconstruction of automibile destruction: An example of interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 585-589

McClelland, J. (1995). Constructive memory and memory distortions: A parallel distributed processing approach. In D. Schacter (Ed.), *Memory distortions: How minds, brains, and societies reconstruct the past*. Cambridge, MA: Harvard University Press.

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

Minsky, M. (1986). *The society of mind*. New York: Simon and Schuster.

Moscovitch, M. (1995). Confabulation. In D. Schacter (Ed.), *Memory distortions: How minds, brains, and societies reconstruct the past*. Cambridge, MA: Harvard University Press.

Neisser, U. (1998). Stories, selves, and schemata: A review of ecological findings. In M. Conway, S. Gathercole, & C. Cornoldi (Eds.), *Theories of memory* (Vol. 2). Hove, UK: Psychology Press.

Nystrom, L. & McClelland, J. (1992). Trace synthesis in cued recall. *Journal of Memory and Language, 31*, 591-614

Petrov, A. & Kokinov, B. (1998). Mapping and access in analogy-making: Independent or interactive? A simulation experiment with AMBR. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences*. (pp. 124-134) Sofia, Bulgaria: NBU Press.

Petrov, A., & Kokinov, B. (1999). Processing Symbols at Variable Speed in DUAL: Connectionist Activation as Power Supply. In: Dean, T. (ed.) Proceedings of the 16th International Joint Conference on Artificial Intelligence. (pp. 846-851) San Francisco, CA: Morgan Kaufman.

Reinitz, M., Lammers, W., & Cochran, B. (1992). Memory-conjunction errors: miscombination of stored stimulus features can produce illusions of memory. *Memory & Cognition, 20*, 1-11.

Schacter, D. (1995b). Memory distortion: History and current state. In D. Schacter (Ed.), *Memory distortions: How minds, brains, and societies reconstruct the past*. Cambridge, MA: Harvard University Press.

Thagard, P., Holyoak, K., Nelson, G., & Gochfeld, D. (1990). Analog retrieval by constraint satisfaction. *Artificial Intelligence, 46*, 259-310.

Wharton, C., Holyoak, K., & Lange, T. (1996). Remote analogical reminding. *Memory and Cognition, 24* (5), 629-643.