Simulation of Episode Blending in the AMBR Model

Maurice Grinberg (mgrinberg@nbu.bg)[1]
Boicho Kokinov (bkokinov@nbu.bg)[1,2]

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

[2]Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. G. Bonchev Street, bl.8
Sofia 1113, Bulgaria

Abstract

This paper presents a series of simulation experiments related to the interaction of memory and analogy-making in the AMBR model. This interaction makes it possible to demonstrate blending between superficially dissimilar episodes as a result of the established analogical mapping between them and of superficially and structurally dissimilar episodes as a result of a double analogy with a third episode. Both simulation experiments model the blending effect of analogy-making. The conditions for the emergence of such blending are explored on the basis of a proposed specific analogy-like retrieval mechanism.

Introduction

Much evidence has been accumulated that human memory is not always (or maybe even rarely) accurate (Loftus, 1979/1996; Loftus, Feldman, & Dashiell, 1995; Neisser, 1998; Roediger & McDermott, 1995; 2000; Roediger, 1996; Moscovitch, 1995; Reinitz, Lammers, & Cochran, 1992; Schacter, 1995, 1999; Schacter, Koutstaal, Norman, 1997, Schacter, Norman, Koutstaal, 1998, Koriat, Goldsmith, Pansky, 2000; Goldsmith, Koriat, & Weinberg-Eliezer, 2002). Cases in which we mistakenly remember some details of an event that are not accurate (or even a whole event that has never happened) are called false memories or illusory memories. There are two main sources of memory distortions: general knowledge (schemata) and blending of episodes. Schematization has been studied since Bartlett (1932), and blending of episodes since the early research of Loftus (1977; 1979/1996). This paper focuses on the mechanisms of blending of episodes. Blending of episodes is supposedly the main reason for misinformation effects and eyewitness testimony failure (Loftus, Feldman, & Dashiell, 1995; Loftus, Miller, & Burns, 1978; Loftus & Palmer, 1974), for "reality monitoring errors" (Johnson & Raye, 1981), for cross-modality confusion errors (Lane & Zaragoza, 1995).

Although experimental research on memory distortions revealed a lot of interesting phenomena the underlying mechanisms remain largely unexplored and the theories are still vague. It remains unclear whether the blends are result of the encoding process which produces similar and undistinguishable memory traces for similar events, or it is a failure of the retrieval mechanism. Most existing models of memory blending (Hintzman, 1988; Metcalfe, 1990; McClelland, 1995, Schacter et al, 1998) are based on similar general assumptions:

- similar episodes correspond to memory representations with overlapping features;
- overlapping representations are mixed up at retrieval because they are difficult to be distinguished from one another.

These models explain the most popular findings in this area, namely, that episodes which are literally similar (Gentner, 1989) tend to be blended later on. Thus, for example, Elizabeth Loftus and her colleagues (Loftus, 1977, 1979/1996, Loftus et al., 1995) demonstrated that when two episodes (e.g. an observed scene and a heard story, a real and an imagined event) differ in a few features than our memory mixes them up: intrusions of features of one episode into the other causes the two episodes to be blended and people report falsely participating objects (hammer vs. screwdriver) or their properties (blond hairs vs. dark hairs). James McClelland and his colleagues (Nystrom & McClelland, 1992; McClelland, 1995) demonstrated "trace synthesis errors" when subjects were presented with highly similar sentences. When the sentences were overlapping (sharing most of the words except two nouns) there were intruders from the alternative similar sentence.

However, these models are challenged by recently emerging findings that blending can also occur between episodes which are only structurally similar, or even between dissimilar episodes. Thus Kokinov (1998) reported a case of blending between superficially dissimilar but structurally similar episodes. The participants solved the radiation problem and a week later they solved an analogous but superficially dissimilar light bulb problem (Holyoak & Koh, 1985). When asked to retell the radiation problem they tended to blend the base and the target – 64% of the participants mistakenly used laser rays (taken from the light bulb problem) instead of X-rays in the description of the radiation problem. The two problems are quite different at the superficial level and could not be represented by very similar feature vectors as suggested by traditional memory models. Thus these models could not easily explain the blending effect obtained.

Based on the main principles of the DUAL cognitive architecture (Kokinov, 1994b, 1994c) and the AMBR model of analogy-making (Kokinov, 1988, 1994a) it was predicted that blending can occur *even between highly dissimilar episodes* (Kokinov, 1998, Kokinov & Petrov, 2001). A sufficient condition would be that the two episodes have participated in a double analogy with a third episode. As result of this analogy new correspondence links are established between the two episodes and the common target. These links serve later on as a means for conveying activation from one of them to the other and thus blending between them may occur.

Traditional memory models cannot explain blending between episodes which are not literally similar since they do not pay attention to the representation of structure and structural similarity between episodes and do not change the representations as result of reactivation and reasoning.

The current paper presents a simulation study with AMBR in an attempt to suggest the specific change in representations, the mechanisms of retrieval and blending, as well as the contribution of various factors to the degree of blending. The factors to be explored are 1) the structure of the representation of an episode and the connectivity within it; 2) the connectivity of this representations with other representations, 3) the availability of the requested information in the best matched episode (encoded/non-encoded, activated/non-activated elements).

Recollection and Blending in AMBR

The AMBR model is based on a general cognitive architecture called DUAL (Kokinov, 1994b,c). This architecture is a hybrid one integrating symbolic representation of structure and connectionist representation of context and relevance. It is based on decentralized representations and emergent computations produced by a society of micro-agents.

An episode is represented in DUAL and AMBR by a coalition of micro-agents. Each of the agents represents symbolically some single aspect of the episode (e.g., a simple statement). The active agents at a particular moment of time comprise the WM. The elements of a coalition which are currently in WM represent the current partial view on the episode.

AMBR views the recollection as a constructive process. It requires the re-activation of a coalition of agents that collectively represents an old episode. However, various members of the coalition will be connected to the rest of it to various degrees. Thus some tightly connected agents will be easily reactivated, while others will be rarely or never recollected. The results will also depend on which elements of the coalition will be initially activated. The activation of an element is, however, not enough for reporting this element since being part of WM does not necessarily mean being aware of it, neither being able to verbalize it. We suggest to view the process of verbally reporting an episode in response to a query or a collection of target cues as a process very similar to analogy-making: the query is considered as a target, the old episodes as possible bases, the task is to find the best match (an old episode that is superficially and structurally as similar with the target as possible), to map the base onto the target and to transfer the missing facts from it. In this way the target query plays the role both of retrieval cue and of partial guidance for which missing facts to be recollected. This process of recollection is performed by the same mechanisms that perform analogy-making in AMBR and involves spreading activation, mapping hypotheses building, competition between episodes, between mapping hypotheses, and between analogical inferences.

Let us suppose that there are two similar episodes in LTM, corresponding to highly similar events that happened in two different moments of time: these two episodes will compete for the mapping with the target cue. If one of these episodes becomes an "undisputed winner", i.e. all mapping hypotheses between the target cues and the elements of this old episode are winners, then we will have a case of accurate memory for the event (although it might be partial memory if some of the elements of the coalition are not activated). However, if the winners come partially from one coalition and partially from another one then we will have a case of episode blending since analogical transfer will be made from both coalitions. There are many factors which determine which of these cases will take place and the simulation experiments explore some of them. Evidently, the result will depend on the activation level of the mapping hypotheses which depend in turn on the supporting hypotheses and on the activation level of the elements of the coalition representing the old episode. These activation levels depend on several factors:

- the degree of familiarity of the particular episode, i.e. the more familiar the episode is, the stronger the connections toward it from outside the episode are and the stronger the links among the elements within the episode are:
  - the degree of connectivity of the whole coalition to other memory elements (e.g. agents representing concepts) – the more and stronger links toward this particular episode are coming from the conceptual agents or agents from other episodes, the greater the chances of the coalition to win the competition with other coalitions;
  - the degree of connectivity within the coalition (the tighter the connections the less the probability for intrusions from other episodes, the weaker the links toward a particular episode element are, the greater the chances this particular element to be "replaced" by an intruder from a different episode);
- the information encoded in the representation of the event: even if the coalition is strong and highly active, if there is missing information about a specific aspect of the event, this is a direct invitation for intruders from other coalitions to be mapped on this specific element of the

target cues (either agents representing pieces of general knowledge or similar aspects of another episode);
• the structural similarity between episodes (even episodes that are not literally similar, e.g. the participating objects are dissimilar, might be blended if they have similar structure and an analogy has been made between them earlier on; the established analogy will be reflected by additional links between the episodes which will in turn facilitate their activation and possible blending);
• the reasoning history in which the two episodes took part (if they participated in a double analogy with a common target episode and new corresponding links are established between them as result of it, they might be blended again).

Simulation of Blending

The first simulation focuses on the case of blending of two superficially dissimilar episodes between which an analogy has been established earlier. This analogy is reflected by the correspondence relations encoded in LTM between the mapped elements from the two episodes. The two episodes taking part in this simulation experiment are called WTO ("heating Water in a Teapot in an Oven") and MTP ("heating Milk in a Teapot on a hot Plate") in LTM. Some of the main agents in these coalitions are shown in Table 1 and the connectivity between the agents – in Figure 1.

Table 1: Main elements in the retrieved episode WTO and the alternative episode MTP.

Analogical episodes in long-term memory	
WTO: "Water in Teapot in a Oven"	**MTP**: "Milk in a Teapot on a hot Plate"
water	milk
in(water, teapot)	in(milk, teapot)
teapot	teapot
color-of(teapot, black)	color-of(teapot, green)
in(teapot, oven)	on(teapot, hot plate)
temp.-of(oven, high-temp)	temp.-of(plate, high-temp)

The retrieval query is formulated as a coalition of target cues (Table 2). Some of the target cues are identical with elements of the episodes, for example, water and teapot. They are represented by separate agents connected to the corresponding agents by *c-coref* links (see Kokinov, 1994a). Thus there is not direct overlap of the coalitions, but they are connected with links instead. For example, the instance of teapot in the WT episode and the instance of teapot in the WTO episode are connected via *c-coref* link representing the fact that this is one and the same teapot, while the instance of teapot in MTP is not linked since it is a different teapot. Other retrieval cues represent the questions we have in mind (like what was the color of the teapot, where was the teapot put, what kind of heating device was used, etc.). These cues are represented by instances of more general objects, properties, or relations (e.g. heating source, color, spatial relation, etc.). The response of the system will be based on the winners in the mapping competition for these retrieval cues. If all the winners are elements of the WTO episode than the recall is correct, while if some of the winners are from WTO and others from WTP, then blending between these two episodes has been produced.

In the first run of this simulation both episodes in LTM were considered to be highly familiar and consolidated – with many and strong associative links within each of the episodes and links from the concepts to their corresponding instances in each of the episodes. The system's response is shown on Figure 1. Each element from the target cue coalition (WT) was mapped to an element from WTO episode. In this case, the whole WTO episode was recalled and no blending occurred.

Table 2: Representation of the target cues - "recall how you heated water in a teapot". Only the main agents are presented.

Target cues **WT**: "Recall the WTO episode"	
Agent	**Propositional representation**
water	(inst-of water-WT water)
tpot	(inst-of tpot-WT teapot)
in	(in water-WT tpot-WT)
heat-source	(inst-of heat-source hdevice-WT)
spatial-rel.	(spat.rel tpot-WT hdevice-WT)
high-T	(inst-of high-T-WT high-temp)
T-of	(temperature-of water-WT high-T-WT)
color-of	(color-of tpot-WT color-WT)
color	(inst-of color-WT color)

In a second run the same LTM episodes were used but their connectivity was modified. Thus the WTO episode ("heating Water in a Teapot in a Oven") was considered as highly atypical and less consolidated than the more typical episode MTP ("heating milk in the teapot on a plate"). This was modeled by strengthening the links from concepts like "heating" and "teapot" to the MTP episode and weakening the links to WTO. Also, the WTO coalition was made looser – weaker associative links among the elements of the coalition. As result of these changes the activation of some elements of WTO (those which were not directly activated by the *c-coref* links from the target cues) became lower. At the same time the activation of the MTP episode became higher since it received higher activation from general concepts as well as from the WTO episode via the "corr-to" links (which are a result of previously established analogy between the two episodes).

This turned out to be crucial and via the strong links within the coalition the rest of the elements received enough activation to become the winners in the competition. The net result is that "hot-plate" and "on (teapot, hot plate)" in episode MTP became more active than "oven" and "in (teapot, oven)" and eventually became retrieved. Similarly, the retrieved colour of "teapot" is "green" (from MTP) instead of "black" (from WTO).

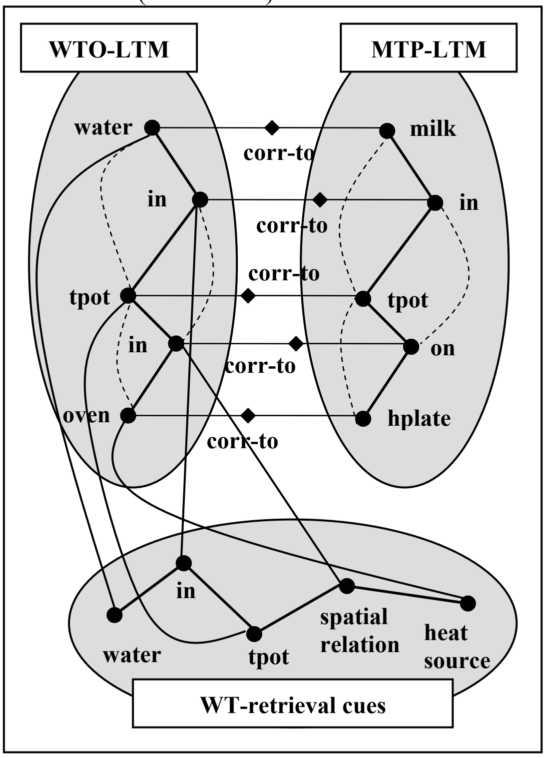

Figure 1: Retrieval of episode WTO from memory without blending. Both episodes (WTO and MTP) have a large number of associative links among their agents, denoted by dotted lines. The analogical correspondences existing in long term memory are also shown by lines connecting episodes WTO and MTP.

The results from the simulation are presented in Figure 2 and Table 3. This is a clear case of blending: the elements specifically described in the target cue (WT) like water and teapot were mapped onto elements of WTO, while the unspecified elements of WT (i.e. the target questions) were mapped onto elements of MTP. In this way the system produced a blended version of the two episodes "the water was in a *green* teapot put *on* a *hot plate*", where the elements in italic are false memories.

The second simulation tries to produce blending between superficially and structurally dissimilar episodes as result of double analogy with a third one.

The double analogy resulted in established new correspondence relations encoded in LTM between the mapped elements from the two episodes and the third episode. The three episodes in LTM taking part in the second simulation experiment are called MCC ("Milk in a Cup made of China"), GP ("Glass on a hot Plate is broken") and BF ("wooden Bowl on fire is burnt out"). The main agents in these coalitions that take part in the double analogy are shown in Table 4.

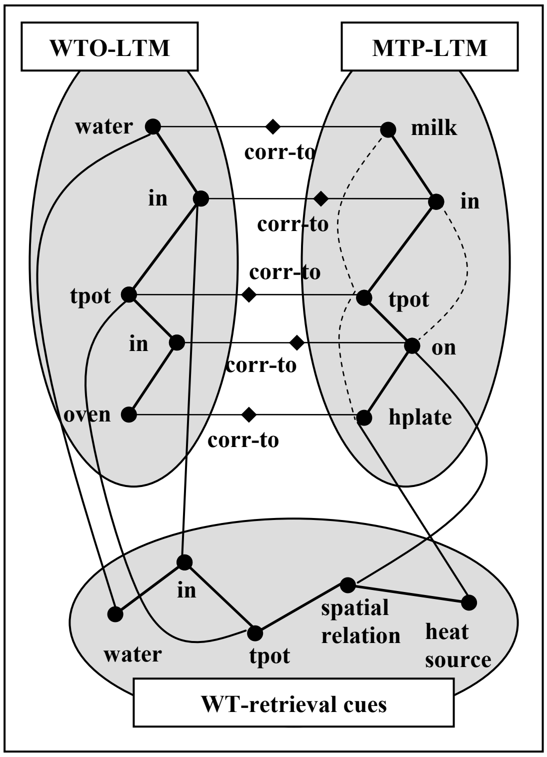

Figure 2: Partial retrieval of episode WTO from memory with blending. Only the episode MTP has a large number of associative links among its agents, denoted by dotted lines. The analogical correspondences existing in long term memory are also shown by lines connecting episodes WTO and MTP.

The results from the simulation are presented in Table 5. The episode to be retrieved "milk in a china cup" has been augmented by the "fact" that it was broken. This intruder has been taken from the episode GP ("glass on a hot plate is broken). In this simulation, the same query method has been used as in the previous simulation.

Table 3: Retrieved situation elements as corresponding to the retrieval cues. The signs "?" denote elements, whose specific fillers were not given in the retrieval cues, but had to be filled from memory. The elements in grey from episodes WTO and MTP have not been retrieved.

Elements retrieved from episodes WTO and MTP		
WT (recall **WTO** episode)	**WTO**	**MTP**
water	water	milk
in(water, teapot)	in(water, teapot)	in(milk, teapot)
teapot	teapot	teapot
color-of(teapot, ?)	color-of(teapot, black)	color-of(teapot, green)
heat-source(?)	oven	hot plate
spatial relation(teapot, heat-source(?))	in(teapot, oven)	on(teapot, hot plate)

Table 4: Main elements in the retrieved episode MCC and the episodes GP and BG. Parts of GP and MCC are analogical to different parts of BF.

Episodes in long term memory		
GP: "Glass on a hot Plate is broken"	**BF**: "Wooden Bowl on a Fire is burnt" episode partly analogical to MCC and GP	**MCC**: "Milk in a China Cup" (episode to be retrieved)
	water	milk
	in(water, bowl)	in(milk, cup)
glass	bowl	cup
on(glass, hot plate)	on(bowl, hot plate)	
hot plate	hot plate	
cause(is-broken)	cause(is-burnt)	
is-broken	is-burnt	
yesterday	today	yesterday
in(episode, room-1)	in(episode, room-2)	in(episode, room-1)
room-1	room-2	room-1

The target cue included a query about the state of the cup ("?phys-relation(cup)") and the element "is-broken" from the GP episode has been mapped to it. The remaining retrieval cues were mapped to the elements of the episode to be retrieved (MCC).

Table 5: Retrieved situation elements as corresponding to the retrieval cues. The signs "?" denote elements, whose specific fillers were not given in the retrieval cues, but had to be filled from memory.

Elements retrieved from episodes MCC and GP		
Retrieval cues for MCC	Retrieved	From episode
milk	milk	MCC
in(milk, cup)	in(milk, cup)	MCC
cup	cup	MCC
phys-rel.(cup)(?)	is-broken	GP (blend)
yesterday	yesterday	MCC
room-1	room-1	MCC

Finally, the following blended version was obtained form the two episodes "the milk is in the china cup, *which is broken*", where the elements in italic are false memories (see Table 5). The reason for the blending is the following: the elements of the episode MCC are activated by the retrieval cue (*c-coref* links); they activate the elements of episode BF which activates the episode GP. Another important source of activation and blending of the episodes MCC and GP is the same place and day of occurrence ("yesterday in room-1").

Conclusions

Two simulation experiments have been performed. The first experiment demonstrated the ability of AMBR to produce blending between superficially dissimilar episodes if an analogy has been previously established between them which resulted in additional "correspondence" links between their elements. If the "to be retrieved" episode is not very typical and familiar (resulting in weak links towards it and within the coalition that represents it) the correspondence links can be used for conveying enough activation to the alternative episode and parts of it to become winners in the competition. This results in blending between the two episodes. Thus connectivity (number and strength of the links within the coalition representing the episode which reflects the semantics of the domain and its familiarity) and typicality (strength of the links from general concepts to their instances) turn out to be crucial factors for blending. This is a prediction that needs to be tested in a psychological experiment.

The second simulation dealt with an even more interesting case: blending occurred between the episode to be retrieved and a second episode, both of which were mapped to different parts of a third episode in a double analogy task. The blended episodes were superficially and structurally dissimilar. This type of blending is an even bolder prediction that has to be experimentally tested. This prediction has been recently tested and there are initial data that confirm it (Kokinov & Zareva-Toncheva, 2001; Zareva-Toncheva & Kokinov, 2003). Moreover, blending occurs in these experiments only after the third problem has been solved, i.e. if the "integrated" problem has not been solved by the participants little or no blending happens. This fact tells us that blending in this case is not an encoding effect, but rather that subsequent reactivation and mapping results in some kind of representational change which is responsible for the blending effect.

The results of the simulations showed that varying the factors: encoding, familiarity and typicality of the episode elements makes it possible to describe different types and levels of blending in memory retrieval tasks in AMBR.

Acknowledgments

We would like to thank Alexander Petrov, Neda Zareva-Toncheva and the AMBR research team for their continuous support and stimulating discussions.

References

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

Gentner, D. (1989). The Mechanisms of Analogical Learning. In: Vosniadou, S. & Ortony, A. (eds.) *Similarity and Analogical Reasoning*. New York: Cambridge University Press.

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

Holyoak, K., Koh, K. (1987). Surface and Structural Similarity in Analogical Transfer. *Memory and Cognition*, 15, pp. 332-340

Johnson, M., Raye, C. (1981). Reality Monitoring. Psychological Review, 88, pp. 67-85

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.* Sofia, Bulgaria: NBU Press.

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.

Kokinov, B. & Zareva-Toncheva, N. (2001). Episode Blending as Result of Analogical Problem Solving. In: *Proceedings of the 23nd Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Koriat, A., Goldsmith, M. (1996). Monitoring and Control Processes in the Strategic Regulation of Memory Accuracy. Psychological Review, 103: 490-517.

Koriat, A., Goldsmith, M., Pansky, A. (2000). Toward a Psychology of Memory Accuracy. Annual Review of Psychology, 51: 481-537.

Lane, S., Zaragoza, M. (1995). The recollective experience of cross-modality confusion errors. Memory and Cognition, 23, pp. 607-610.

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

Loftus, E. (1979/1996). *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.

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

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.

Roediger. H., McDermott, K. (1995). Creating False Memories: Remembering Words not Presented in Lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21: 803-814.

Roediger. H., McDermott, K. (2000). Tricks of Memory. *Current Directions in Psychological Science*, 9: 123-127.

Roediger, H. (1996). Memory Illusions. *Journal of Memory and Language,* 35: 76-100.

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

Schacter, D. (1999). The Seven Sins of Memory. *American Psychologist*, 54: 182-203.

Schacter, D., Koutstaal, W., Norman, K. (1997). False Memory and Aging. *Trends in Cognitive Science*, 1: 229-236.

Schacter, D., Norman, K., Koutstaal, W. (1998). The Cognitive Neuroscience of Constructive Memory. *Annual Review of Psychology*, 49: 289-318.

Zareva-Toncheva N. & Kokinov, B. (2003). Blending of Non-Similar Episodes as a Result of Analogical Mapping with a Third One. *Proceedings of the 25th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.