

The Story with Reminding: Memory Retrieval is Influenced by Analogical Similarity

Charles M. Wharton¹, Keith J. Holyoak¹, Paul E. Downing¹,
Trent E. Lange², and Thomas D. Wickens¹

¹Dept. of Psychology and ²Dept. of Computer Science, University of California, Los Angeles, CA 90024

Abstract*

AI models of reminding (ARCS, MAC/FAC) that predict that memory access is influenced by analogical similarity are tested. In Experiment 1, subjects initially studied a set of 12 target stories. Later, subjects read 10 other cue stories and were asked to write down the stories they were reminded of from the first set. Cue stories were associated with either an analogous and disanalogous target (competition condition), an analogous target (singleton condition), or a disanalogous target (singleton condition). An effect of analogical similarity was found only in the competition condition. Experiment 2 used the same design but targets and cues were simple subject-verb-object sentences. Cue sentences shared similar nouns and verbs with target sentences. Materials were constructed such that associated nouns either consistently mapped or cross-mapped between cues and targets. Consistent-mapped sentences were recalled more than cross-mapped sentences in both conditions. Issues for future research are addressed.

Introduction

One of the central issues in analogical reasoning theory concerns the degree analogical similarity affects reminding (e.g., McDougal, Hammond, & Seifert, 1991). This issue is important because what one is reminded of in problem-solving affects further action. If one is reminded of an analogous problem in which a similar goal was solved, the plan that achieved that goal can be used to solve the current problem. However, if one is reminded of past situations on the basis of superficial resemblances, little of the knowledge associated with that reminding can be used to achieve the current goal. This question

remains unanswered. In the present experiments, we explore the retrieval conditions that are necessary to demonstrate analogical reminding.

Computer Models of Reminding

The extent to which reminding theories assert that memory retrieval is influenced by analogical similarity appears to be partially a function of the amount of domain expertise assumed about the reasoner. Some case-based reasoning (CBR) models implicitly represent reasoning done within a domain of expertise (see review in Reisbeck & Schank, 1989). As such, memory access in these systems is determined mostly by the plan or goal similarity between the current problem and cases in memory. Case retrieval will be only minimally based on surface similarity. In contrast, models of more general reminding such as ARCS (Thagard et al., 1990) and MAC/FAC (Gentner & Forbus, 1991) or assert that memory retrieval is influenced by surface (i.e., cue/target lexical overlap) as well as analogical similarity.

In ARCS theory, reminding is governed by three types of constraints: direct semantic similarity of concepts, isomorphism (consistent mapping of predicates and arguments), and pragmatic centrality (problem-solver's goals). Retrieval first proceeds with a search of all symbolic representations in memory to find targets that overlap semantically with propositions in the cue. Second, a connectionist "mapping network" is formed to represent competing potential mappings that are created between the cue and semantically-related episodes in memory. The connections of this mapping network form pressures on reminding. Excitatory and inhibitory connections embody the three constraint types. Finally, a connectionist process of this mapping network produces retrieval of the episode(s) in memory that best satisfy these constraints.

MAC/FAC theory is similar to ARCS in that retrieval is based on semantic and structural overlap. Computationally, retrieval is a two-step process. In the first stage (MAC), the episode in long-term memory that has the most surface commonalities

*Preparation of this article was supported by Contract MDA 903-89-K-0179 from the Army Research Institute and a Keck Foundation grant and NSF Grant DIR-9024251 to the UCLA Cognitive Science Research Program. We thank Mary Jo Ratterman and Dedre Gentner for assistance with Experiment 1.

with the probe story is retrieved. Stories are represented in MAC as vectors, where each element represents a word. Similarity is computed by taking the dot product between the probe and each story in memory. Any other stories whose dot product is within 10% of the best match are also retrieved. The second stage (FAC) computes how well each retrieved first stage story matches the cue, based on common relational structure and object descriptions. Stories are represented in FAC in predicate calculus form. The episode with the highest match, along with any story within 10% of the best match, is retrieved.

The predictions of ARCS and MAC/FAC seem relatively simple: Both semantic and relational similarity should influence reminding. However, actually demonstrating the affect of relational similarity empirically has proven difficult.

Empirical Studies of Reminding

Analogical reminding has been directly examined by Ratterman and Gentner (1987). Here, subjects initially read a number of stories. Two weeks later subjects read more stories and were asked to write down any stories that they were reminded of from the previous session. Ratterman and Gentner varied object attribute similarity and higher-order relational similarity in their materials. Crossing these two types of similarity resulted in four different cue/target similarity matches, (a) shared object attributes and higher-order relations (literal similarity), (b) shared object attributes only (mere appearance), (c) shared higher order relations only (true analogy), and (d) no shared object attributes or higher-order relations (false analogy). No cue story was ever matched to more than one target story. The only reliable differences found were advantages for the literal-similarity and mere-appearance conditions relative to the true-analogy and false-analogy conditions (recall proportions were 0.56, 0.53, 0.12, and 0.09, respectively). Ratterman and Gentner concluded that reminding is primarily, though not exclusively, influenced by object or "surface" similarity (see also Gentner & Forbus, 1991; Gentner & Landers, 1985; Seifert et al., 1986).

Why might the role of analogical similarity in reminding be so difficult to demonstrate? An answer can be found by examining computational models of analogical and case-based retrieval. Most such reminding models attempt to retrieve at least the single *best* or most similar target that can be accessed semantically for a given cue. The most similar target in memory is usually considered to be one that which is both analogous to the cue *and* shares significant semantic overlap with it. If there are no analogous targets that are semantically related to the cue, then a semantically-related non-analogous target will be

Table 1. Example of One Materials Set (Lexically-Associated Cue and Targets) from Wharton et al. (1991)

Consistent Target
Having just been fired from a high level job, he decided to go to his church for counseling. <i>The pastor calmed the businessman.</i>
Inconsistent Target
The church was having trouble approaching local corporations for contributions to the shelter. <i>The executive soothed the priest.</i>
Target-Cue Sentence
<i>The rabbi reassured the chairman.</i>

considered to be the best match, and therefore be retrieved. Accordingly, most contemporary models of analogical and case-based reminding would predict only small effects of cue/target analogical similarity in experiments in which each cue is semantically associated to a *single* target (a *singleton design*). This is because a single related target is likely to be retrieved regardless of whether it is analogous or not. This, in fact, is what previous studies of analogical reminding, all of which have used singleton retrieval designs, have generally found, .

In Wharton et al. (1991), we demonstrated an effect of analogical similarity on reminding by using a design in which each cue was semantically associated with *both* a consistently- and inconsistently-mapped target in memory (a *competition design*). An example of our materials and design is shown in Table 1. Targets were short texts built around a single sentence whose roles either mapped consistently with those of the cue (e.g., *rabbi/pastor, chairman/businessman*), or mapped inconsistently with those of the cue (e.g., *rabbi/executive, chairman/priest*).

A schematic of our design is shown in Figure 1. Subjects saw items in both competition (left boxes) and singleton conditions (right boxes). Wharton et al. found an overall effect of analogical similarity on reminding (in the form of consistent object-level mappings). However, the effect of analogical similarity was much larger in the competition condition than in the singleton condition.

Figure 1. Design of Wharton et al. (1991)

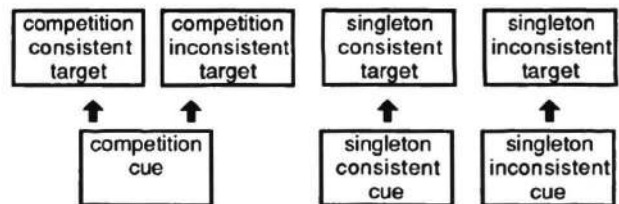


Table 2. Example of a Materials Set Used in Experiment 1 (after Seifert et al., 1986).

Set A, Theme 1: Ernie was really encouraged about his interview for a security guard at the new factory in town. He thought he was saved. Ernie went to the shopping mall, hunted for a dark blue security guard uniform, and bought several. The next day he received a phone call from the factory personnel office about the security guard position. Ernie was dismayed that he had wasted money. He didn't have a job.

Set A, Theme 2: Dan wasn't working and he was very concerned because he had very little left in his bank account. Several days later he had lunch with the chairman about becoming a broker. Dan thought he had made a good impression when he gave his resume to the investment partnership. Dan went to the department store and tried on some suits, and got a few. He felt that he was moving up again.

Set B, Theme 1: Ronnie thought she had it made because she thought she had done well in the audition for a musician. Ronnie went to the music showroom, played some electronic keyboards, and then purchased one. Later she got a message from the guitarist about her playing keyboards. She wasn't in a band. Ronnie was dejected that she had run up her credit card.

Set B, Theme 2: Pam was worried she that she had blown her savings. She was between jobs. Pam was really excited about her tryout as a dancer for a new musical. That evening she met the director about the dancer role. Pam got over to some stores, searched for, and bought some leotards. She believed her troubles were over.

Experiment 1

The view we took in Wharton et al. (1991)—that our results with varied object (or role)-level mappings generalized to mean that analogy influences reminding—could be disputed. Although our materials varied the consistency of object-level mappings, we never manipulated the consistency of higher-order relational mappings. As can be observed with the materials set in Table 1, *reassured* consistently maps to both *calmed* and *soothed*. That is, there is no difference in higher-order relational consistency between the consistent and inconsistent targets. Thus, our previous experiments did not constitute a complete test of the effect of analogical similarity on reminding.

We constructed new set of stories in which cue/target relational consistency has been varied. An

example of a story set is shown in Table 2. Here, the same underlying set of propositions are used to construct two different story lines, *counting your chickens before they're hatched* (theme 1), and *finding desperately-needed employment* (theme 2). No content words are used in more than one story and the characters and objects used in set A that map to those in set B are not closely related semantically (e.g., stock broker and dancer). Thus, the paired stories represent analogous rather than literally similar instantiations of the same common theme.

In the present study, subjects rated 12 stories of this type for imageability. Later, subjects read 10 more stories and were asked to write down any of the rated stories of which they were reminded. We used the same design as in Wharton et al. (1991). The event-related cue and target stories were either instantiations of the same theme (analogous) or of different themes (disanalogous). We predicted that analogous stories would be retrieved more than disanalogous stories, and that this difference would be greater in the competition than in the singleton condition.

Method

Materials. Materials consisted of 14 sets of 4 stories (see Table 2). Some stories were derived from Ratterman and Gentner (1987) and Seifert et al. (1986). Within each set of stories, we constructed one core set of propositions. Two different unique story plots were created for each story set by rearranging the sequence of propositions. To avoid having the surface order of propositions covary with analogical similarity, event sequences shared between thematically similar stories were changed as much as was possible without altering the underlying story plot. Each story within a set was written about a different set of actors such as roommates, countries, or siblings. No content words or proper names were used in more than one story across the entire set of materials.

In order to determine if, when reading our materials, people are sensitive to the factors we manipulated, 20 undergraduates attending the University of California, Los Angeles (UCLA), completed questionnaires designed to assess the perceived similarity of the cue and target stories. Subjects were asked to rate "...how similar are the scenes being described" on a 6-point Likert scale (range: 1, completely dissimilar to 6, completely similar). Analogous story pairs were rated more similar than disanalogous story pairs (4.70 vs. 3.40), $F(1, 19) = 34.86, p < .0001$; disanalogous story pairs were rated more similar than unrelated story pairs (2.00), $F(1, 19) = 67.37, p < .0001$.

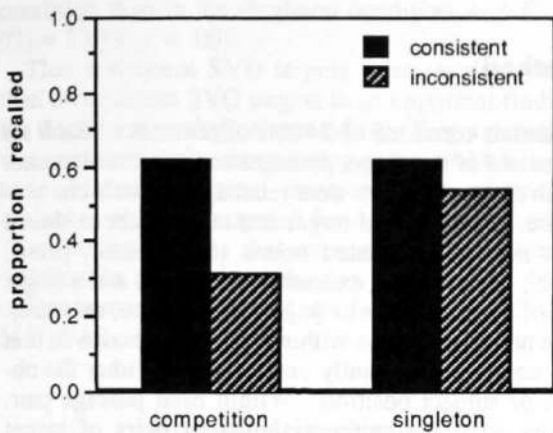


Figure 2. Proportion of reminders for each condition of Experiment 1.

Two supersets of materials were created by assigning the stories in each of story sets to one of two groups. The two stories from each story set assigned to each superset contained dissimilar themes. An equal number of cues and targets came from each superset. In order to have all stories appear in all conditions, we created 14 separate configurations of cue and target stories. Each configuration of materials was administered to three subjects.

Target and cue stories were made into separate booklets. A different random order of targets and cues was created for each subject. Each booklet of target stories contained 12 stories (2 stories each in the competition analogy, competition disanalogy, singleton analogy, singleton disanalogy conditions, and 4 stories not related to any cue). Each booklet of cue stories contained 10 stories (2 stories each in the competition, singleton analogy, singleton disanalogy conditions, and 4 stories not related to any target).

Procedure. Subjects were 42 UCLA undergraduates who participated either for pay or in order to meet requirements of one of several psychology survey courses.

During the encoding portion of the experiment, subjects rated the target stories on a 6-point Likert scale for imageability. Subjects were not informed that they would later try to recall these stories. Subjects were given a 5 min distractor task, after which they were asked to read the cue stories and write down anything that they were reminded of from the stories that they had previously rated.

Data Analysis. Reminding protocols were scored with respect to which story was accessed in response to a given cue. For each separate attempt at recalling a story (i.e., each attempt to report what the subject considered a single passage), credit for access was given to whichever passage had content words recalled; if content words from two passages were in-

cluded, credit was given to the passage that contributed the most content words. If subjects wrote down content words from two passages in *separate* retrieval attempts in response to a single cue, access credit was given for each. Because synonym substitutions could be confused with interchanges of hyponyms across paired passages, a criterion of literal recall had to be used in scoring content words.

In order to be able to generalize our findings beyond the specific materials we created, the conservative min F' formulation of analysis of variance was calculated for all tests of mean differences (Clark, 1973). One story from each of the 28 story themes was used in every condition. Thus, there were 28 observations per condition in the item ANOVAs.

Results and Discussion

The proportion of target story types of which subjects were reminded by cue stories is shown in Figure 2. Subjects recalled more analogous (consistent) stories than disanalogous (inconsistent) stories, min F' (1, 54) = 6.29, $p < .05$. The interaction between competition and analogical similarity was not reliable, min F' (1, 54) = 2.39, $p > .10$; however, both the subject and item ANOVAs were significant, $p < .05$ (F (1, 28) = 4.38, F (1, 27) = 5.27). As predicted, there was more access of analogous stories than disanalogous stories within the competition condition, min F' (1, 54) = 6.76, $p < .05$, but not within the singleton condition, min $F' < 1$.

These findings strongly support the claims of general reminding models such as ARCS and MAC/FAC that analogical similarity influences memory. As in Wharton et al. (1991), an effect of analogical similarity was only obtained with a competition design. This would seem to indicate that the effect of higher-order cue/target relational consistency on reminding is not very strong. Thus, our results also imply that direct semantic similarity of individual concepts (objects and predicates) dominates reminding for novice subjects.

Experiment 2

In Experiment 1, we demonstrated an effect of analogical similarity with story materials. Wharton et al. (1991) showed a similar effect with target passages consisting of several sentences that described a single scene (see Table 1). Thus, in the present Experiment 1 and Wharton et al. (1991), target materials consisted of multi-sentence scene or story descriptions. It is likely that the text contexts encouraged subjects to make inferences that augmented their text representations and therefore increased the chance that

analogous targets would be recalled more than disanalogous targets.

The passages in Table 1 can be used to illustrate how sentential context might make readers add additional knowledge to their text representations. For the consistent target in Table 1, it is natural to infer from the surround statement, *Having just been fired from a high level job...*, that the businessman went to the pastor about a personal problem, and that the priest helped the businessman to deal with this problem. For the inconsistent target, the surround statement, *The church was having trouble approaching local corporations for contributions...* might lead to the inference that the priest went to the executive about a financial problem, and that the executive helped the priest to deal with it. The cue sentence, *The rabbi reassured the chairman*, although not itself embedded in a text context, seems more likely to elicit inferences that parallel those generated when reading the consistent target, *The pastor calmed the businessman*, than those generated when reading the inconsistent target, *The executive soothed the priest...* Thus, additional inferences triggered by the surrounding text context should increase semantic and structural overlap between subjects' representations of the cue sentence and of the consistent target, relative to that of the cue with the inconsistent target.

It is possible that the inferences that readers create from multi-sentence scene descriptions are necessary for analogy and consistency effects in reminding. This view contrasts with that of ARCS and MAC/FAC, which predict that targets with consistent object mappings will be recalled more than targets with inconsistent object mappings, even in the absence of differential inferences. In order to test these contrasting predictions, single sentence targets *without* any surrounding text context (e.g., *The pastor calmed the businessman*) were presented to subjects during initial encoding.

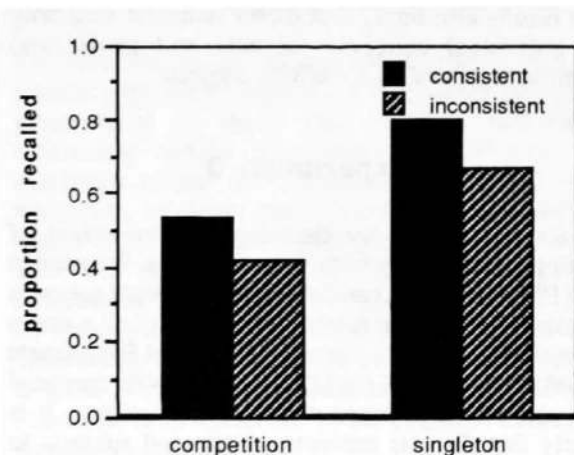


Figure 3. Proportion of reminders for each condition of Experiment 2.

Method

Materials consisted of 24 sets of sentences. Each set consisted of two target passages and one cue sentence. Both target sentences were related to the set's cue sentence. The matched target and cue sentences shared two sets of associated nouns (e.g., pastor, priest, rabbi; businessman, executive, chairman) and a single set of associated verbs (e.g., calm, soothe, reassure). The nouns and verbs within a set were chosen so that the nouns would jointly make sense in either the object or subject position. Within each passage pair, verbs were randomly assigned to pairs of target nouns, after which each target sentence was randomly assigned to one of the two passages. In order to avoid confounding cue/target consistency with surface order of the noun hyponyms, an equal number of active and passive cue sentences and target sentences were constructed in each condition. Random assignment was used to decide whether cue and target sentences would be active or passive and which target passage would be analogously cued.

The methodology for similarity ratings was virtually identical to that Experiment 1. Subjects were 96 UCLA undergraduates. Consistent pairs were rated more similar than inconsistent pairs (4.10 vs. 3.12), $\min F'(1, 56) = 19.76, p < .001$. Inconsistent pairs were rated more similar than unrelated pairs (2.07), $\min F'(1, 51) = 19.11, p < .001$.

There were 24 different configurations of cue and target sentences. Each configuration was administered to three subjects. The design of target and cue booklets was identical to that in Experiment 1.

Subjects were 72 UCLA undergraduates who participated either for pay or in order to meet requirements of an introductory psychology course. We used the same design and procedure as in Experiment 1 except that subjects made three separate ratings passes, respectively, for plausibility, meaningfulness, and imageability. Subjects were given 20 s to read and rate each sentence during each pass through their target sentence booklets.

All 48 target sentences appeared in all conditions. Consequently, we treated each target as an observation for the item ANOVAs (i.e., $n = 48$ observations). We used Experiment 1's scoring rules.

Results and Discussion

The proportion of target sentences retrieved is shown in Figure 3. Although the difference is relatively small, access was more likely with structurally consistent targets than with structurally inconsistent targets, $\min F'(1, 93) = 5.76, p < .05$. There was less access of each target type in the competitor

condition than in the singleton condition, $\min F' (1, 92) = 33.39, p < .001$.

That consistent SVO targets were recalled more than inconsistent SVO targets is an important finding for theories of analogical reminding. Target passages were not seen with a surrounding context which could have biased subjects' interpretations of target sentences and so accounted for our previous results. This finding implies that even minimal configural differences between episodes in memory will influence memory retrieval, as claimed by ARCS and MAC/FAC.

Future Directions

There are several general areas in which investigation would contribute much to theories of analogical reminding. One possible weakness with our studies concerns the short time delay that we have always used, 5 min. CBR systems generally model retrieval of episodes long after they have been experienced. Thus, the effect of time delay on reminding with story material is a critical area in need of investigation.

Perhaps the most pressing area for future research on case-based reminding is on the differences between expert and novice remindings. Such work is necessary to resolve the claims of some CBR models relative to those of general reminding systems such as ARCS and MAC/FAC.

Finally, future work needs to address the role of comprehension and inference processes in reminding. ARCS, which uses only taxonomic relations between individual concepts as retrieval paths, lacks any capacity to infer, for example, that the overall theme of a passage is "retaliation". As a result, the model is unable to use such implicit abstractions to guide retrieval by indexing episodes with similar abstract themes. Accordingly, ARCS might have difficulty modeling Experiment 1. We have addressed this problem with REMIND (Lange & Wharton, 1992), a structured connectionist model that integrates text inferencing and episodic reminding. Future experimental studies should address the questions of *why* structural and analogical similarity affect remindings — do syntactic structural isomorphisms play a role in retrieval in and of themselves, or can the entire effect of analogical similarity be explained by inferred similarities produced in the course of understanding the theme of a story?

References

- Clark, H. H. 1973. The Language-as-Fixed-Effect Fallacy: A Critique of Language Statistics in Psychological Research. *Journal of Verbal Learning and Verbal Behavior* 12: 335-359.
- Gentner, D., and Forbus, K. D., 1991. MAC/FAC: A Model of Similarity-Based Retrieval. In Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society, 504-509. Hillsdale, NJ: Lawrence Erlbaum.
- Gentner, D., and Landers, R., 1985. Analogical Reminding: A Good Match is Hard to Find. In Proceedings of the International Conference on Systems, Man and Cybernetics, 607-613. Tucson, AZ.
- Lange, T., and Wharton, C., 1992. REMIND: Integrating Language Understanding and Episodic Memory Retrieval in a Connectionist Network. In Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum.
- McDougal, T., Hammond, K., and Seifert, C., 1991. A Functional Perspective on Reminding. In Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society, 510-521. Hillsdale, NJ: Lawrence Erlbaum.
- Ratterman, M. J., and Gentner, D., 1987. Analogy and Similarity: Determinants of Accessibility and Inferential Soundness. In Proceedings of the Ninth Annual Meeting of the Cognitive Science Society, 22-34. Hillsdale, NJ: Lawrence Erlbaum.
- Reisbeck, C. K., and Schank, R. C., 1989. *Inside Case-Based Reasoning*. Hillsdale, NJ: Lawrence Erlbaum.
- Seifert, C. M., McKoon, G., Abelson, R. P., and Ratcliff, R. 1986. Memory Connections Between Thematically Similar Episodes. *Journal of Experimental Psychology: Human Learning and Memory* 12: 220-231.
- Thagard, P., Holyoak, K. J., Nelson, G. and Gochfeld, D., 1990. Analog Retrieval by Constraint Satisfaction. *Artificial Intelligence* 46: 259-310.
- Wharton, C. M., Holyoak, K. J., Downing, P. E., Lange, T. E., and Wickens, T. D., 1991. Retrieval Competition in Memory for Analogies. In Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society, 528-533. Hillsdale, NJ: Lawrence Erlbaum.