# The Effect of Similarity on Memory for Prior Problems

**Jeremiah M. Faries**
Department of Psychology and
The Institute for the Learning Sciences
Northwestern University
Evanston, Illinois,60201
jerry@swim-two-birds.psych.nwu.edu

**Karen R. Schlossberg**
School of Education and Social Policy
Northwestern University
Evanston, Illinois, 60201
karen@swim-two-birds.psych.nwu.edu

## Abstract

Students often rely on prior work or previously studied examples to help them solve their current problems. In this paper we investigate the relative contributions of easily accessed superficial similarity and deep, solution relevant, structural similarity to memory for prior problems. Some models of memory for analogy suggest that superficial similarity initially selects or constrains memory for prior examples and predicts that analogs that share both surface and structural similarities will be more likely noticed by novices. An experiment is reported in which subjects are observed as they learn how to program. We find that people remember the examples that are related in terms of structural features alone as frequently as those that are related in terms of both structural and superficial features but there is no advantage to having superficial similarities as well. Moreover, even though superficial features sometimes are associated with helpful similarities and sometimes associated with unhelpful similarities people still do not get misled by superficial similarity when that is the only basis for similarity. This finding suggests that models that require superficial similarity as a major selection procedure for analogical reminding may need to be modified for conditions in which people are learning a new skill.

People often learn from experience by either solving problems themselves or by observing the problem solving behavior of others in the form of worked examples or apprenticeship. To understand how novices in a domain learn from these kinds of experiences it is important to understand what kind of information they acquire from seeing an example or from performing a set of problem exercises and how they subsequently use their memories for those experiences to either recall solutions or search through external media for helpful material. Some experimental evidence suggests that people tend not to be very good at noticing important solution relevant similarities between study examples and problems to be solved (Gick & Holyoak, 1983; Ross, 1987) or between vignettes that are read at separate times (Gentner, Ratterman, & Forbus, 1993). This has led some researchers to conclude that novices are misled by superficial similarity and are therefore not very good an using analogy or case-based reasoning when learning a new domain (VanLehn, 1989). Instead, they focus on the kinds of rules novices acquire by studying examples or working through problem sets (VanLehn & Jones, 1993). However, in addition to filling in knowledge gaps, prior experience with a problem can provide a rich memory base for problems solved in the new domain which can facilitate subsequent problem solving if the prior experience is recalled (e.g., Hammond, 1986).

In earlier experiments in our lab we have shown that novice programmers are sensitive to structural similarities and are able to ignore superficial similarity under a variety of circumstances. This is in contrast to several experiments which show that novices are misled by superficial similarity (e.g., Ross, 1987; 1989). In particular, the novices we studied are much more likely to search a computer memory base for their prior problems that are structurally related to the problem they are working on than they are to search for a problem that is only related in terms of the superficial content (Faries & Reiser, 1988). In these experiments novice programmers spend between 10 and 17 hours learning how to program. They save all of their problems (or examples studied) and can later search for these by providing a distinctive keyword or phrase. Although the subjects are reading text and solving several problems, the last few problem sets represent the experimental "target" problems. For each of these problems one prior problem was seen that bore superficial similarity only, one that was superficially dissimilar and several others that were neither superficially or structurally similar. We recorded all their interactions in the environment including the particular problems they chose to consult for each target problem. In general, the majority of their searches were for the prior problems that were structurally related and they were rarely ever misled by the superficial similarities. We have also shown that novices are less likely to be misled if they solve the source problems than if they study well-annotated examples (Faries, 1991). This suggests that even novices have a memory for prior work that is accessible via structurally relevant memory routes.

Our earlier experiments support a conclusion that novices are not automatically misled by surface similarity and that they can, indeed, use structural retrieval routes even when surface similarity is not helpful and is potentially misleading. We deliberately contrasted the similarity that was primarily structural with that which was primarily superficial for each target problem because the subject had an equal chance of noticing either for each target problem. This enables us to use each item as a control because every source item appeared in both of the similarity conditions and each target item had a potential analog from each of two conditions. Our interpretation of the competitive

design is that people are at least as sensitive to structural similarities as to superficial similarities when remindings occur during a learning context. These remindings occur when particular learning and problem solving goals are in place (Seifert, 1988) and, perhaps more importantly, our subjects were encoding the source items as a result of developing a solution for the problem. We have begun to develop a model that requires that byproducts of the encoding process be stored as part of the event memory for example study and prior problem solving episodes (cf. Carbonell, 1986). This suggests that in cases where students participate in the development of the solution they will have more relevant (or structural) retrieval routes to other related problems than if they only superficially process the problem (lazily studying examples for instance). Some students are more likely to self-explain examples and thereby participate in the development of the solution (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). In this way students who lack a deep structural understanding of the domain are able to notice similarities that are structural because of similarities in processing. This provides a mechanism that enables novices to the bias to notice superficial similarity. In processing any event or word problem the superficial details must be processed so, by our reasoning, the superficial details will always be part of the representation and retrieval will always be possible by superficial routes. However, when students process examples or problems in more detail other more beneficial processing details are stored thus enabling retrieval based on those similarities.

One interesting aspect of this competitive design is that subjects may have learned that surface similarity would never be useful. In fact the predominance of structural remindings may arise because even though subjects may have been inclined to look up prior superficial isomorphs, they ignored their inclinations and sought, instead, the structural isomorphs. If this explanation were true it would still be very important to understand how the metaknowledge about predictability so affects their memory that they are able to notice structural similarity when ordinarily they cannot. It is not clear how such metaknowledge fits into existing models of analog retrieval (e.g., Gentner, et al., 1993; Thagard, Holyoak, Nelson, & Gochfeld, 1990).

In this paper we replicate our earlier work observing novice programmers who encode source problems during the development of a solution, but we designed source and target problems so that superficial similarity was helpful for retrieval of a structurally related problem only half the time. In this design subjects cannot exploit the knowledge that superficial similarity will never be helpful and thus ignore remindings that are based on the superficial similarity thus enabling an examination of the relative contributions of Superficial and Structural similarity. If subjects are able to notice structural similarity primarily because of the meta-awareness of the lack of utility of superficial information then we would expect that when they have no particular knowledge of the utility of superficial information people will retrieve many analogs related both superficially and structurally and will be often misled by analogs that are only superficially similar. In other words, people will not ignore their bias (e.g., Gentner et al., 1993; Ross, 1984) to pursue superficial similarity Models like the Analogical Constraint Mapping Engine (ACME) (Thagard, Holyoak, Nelson, & Gochfeld, 1990), where superficial similar is one main constraint in setting up comparison networks, or MAC/FAC (Gentner, et al., 1993) where the initial selection of candidate analogs is driven by superficial similarity, predict that superficial similarity will increase the chances of noticing a structural analog when it is also superficially similar and increase the likelihood of being misled when it is paired with structurally dissimilar problems. If, on the other hand, being reminded of structural similarity does not require an initial rejection of more obvious superficial similarity, but is instead based on the reasoning encoded as part of the source memory, then subjects will retrieve analogs that are structurally related (with or without superficial similarity) and will, as in previous experiments in our system, infrequently pursue misleading merely superficial analogs.

## Method

### Subjects

Subjects were 30 students from the Northwestern University community. All the students were novice programmers. The average math SAT score for the subjects who completed the full experiment was 674 and the range was from 540 to 800. Four subjects were excluded because of equipment failure or because they chose to quit prior to completion of both chapters. The average SAT for those who did not complete the experiment was 610 and the range was from 540 to 760.

### Materials

The Behavioral Analogical Tracing Electronic Book environment (BATBook) was used to observe novices learning how to program by keeping an extensive record of all activities within the system. The system is described in detail elsewhere (Faries & Reiser, 1989) so we will only briefly describe the key features. BATBook uses a window system to present a textbook, workspace, and examples or problems at particular points within each learning session. The system records a subject's use of text, access of previous problems or examples, and current problem solving activity. Subjects read text and are presented with problem sets every few pages to illustrate the points made in the text. After solving each problem the student must submit the problem to the system. BATBook gives minimal feedback if the solution is incorrect and requests that the subject attempt at least one fix. There are several button options permitting search back either through the previously read text, the

previously submitted problems, or the verbatim history of problem solving activity in the workspace. This allows us to observe the similarities subjects notice and pursue as well as obtain a detailed record of all problem solving steps. In addition to the behavioral protocol we asked half the subjects to think aloud while reading and problem solving. A cassette recorder and lapel microphone were used to collect these protocols. We have not analyzed the protocol data yet but have included this in the description because it was a major part of the procedure for half the subjects.

The experimental stimuli consisted of 6 source problems and 18 target problems. For each source problem we constructed one problem that shared the cover story with the original source problem but required a novel solution (Superficial Only), one problem had a novel cover story and required a similar solution (Structural Only), and one problem had both and similar cover story and required a similar solution (Both). Each subject saw the same 6 source problems. Later in the sequence each subject saw 2 randomly ordered examples of each of the three types of target problems described above. All experimental variables were varied within subjects and particular instances of the target problems were counterbalanced across all subjects.

To verify our intuitions about the types of similarities built into the construction of the stimuli we had independent raters rate the similarites between the problems. Each pair of target and source problems were rated for similarity (structural and superficial) by independent LISP programmers on a scale of 1 to 5, with 1 being the most similar and 5 being the least similar. The problems that shared structural and superficial similarity were rated the most similar in terms of story (average 1.8 out of 5) and highly similar in terms of task (average 2.3 out of 5), with superficial and structural similarity getting very close ratings. Problems with a novel story and the same task were rated low in superficial similarity (average 3.3 out of 5) and high in structural similarity (average 1.1 out of 5). Problems with a similar story but novel task were rated high in superficial similarity (average 2.2 out of 5) but low in structural similarity (average 4.1 out of 5).

Subjects were run separately and completed the experiment in one session lasting from three to five hours. Half of the subjects were trained in the techniques of think-aloud protocol generation (Ericsson & Simon, 1984) before initiating the session and were instructed to continue thinking aloud throughout the session including during problem solving, text comprehension phases as well as during search activity. Subjects worked at their own pace through two chapters of an introductory LISP text (Anderson, Corbett, & Reiser, 1987). The first chapter familiarized subjects with the LISP environment and the BATBook features as well as introduced them to very basic programming constructs. During Chapter 2 the subjects began to define functions according to the specifications layed out in word problems. It was during Chapter 2 that we collected the data we report in this paper. The six source problems were presented early in Chapter 2 and the six target problems were presented towards the end of Chapter 2. Problem sets were interspersed within expository text as is typical of textbook presentations.

## Results

The main finding we report is for the reminding behavior averaged over all participants in the experiments. We do not as yet have the protocol data transcribed and analyzed so all results are collapsed across protocol condition. We also identified subjects as good, moderate or poor learners based on their performance to enable an investigation of whether the subjects' performance correlated to the kind of search activity they engaged in. For this part of the analysis subjects were divided on the basis of the number of correct final solutions they produced; good learners had 11-12 correct, medium learners had 9-10 correct, and poor learners had 8 or less correct. Furthermore, good learners decreased their incorrect solution attempts between problem sets by 50%, whereas medium learners decreased by 25% and poor learners did not decrease at all. Each group of learners spent the same amount of time, on average, in the sessions. For each set of target problems subjects could have searched for problems that were superficially similar (.33 of the problems), problems that were structurally similar only (.33), or problems that were both structurally and superficially similar (.33).

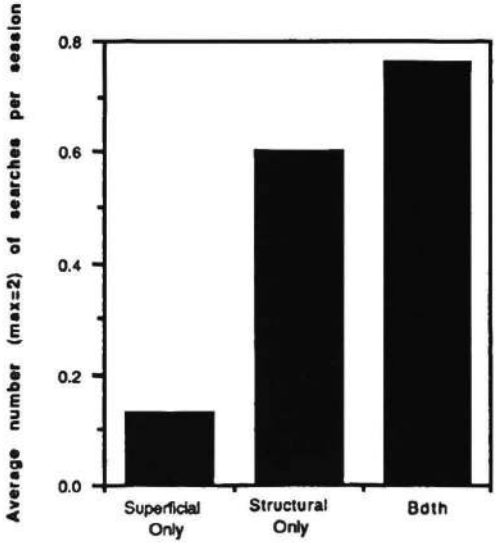


Figure 1: Average of searches for each type of relation between source and target.
(Total possible is 2 per subject per session)

### Remindings

As seen in Figure 1, averaged across learner type most of the searches occurred when the source target

problem was related in terms of the structural similarity (91% of all searches). A 3 X 3 Analysis of Variance using Ability (Hi, Med, Low) and Search Type (Superficial Only, Structural Only, Both) as variables shows a significant effect for Search Type, $F(2,54) = 9.08$, $p < .001$. Newmann-Keuls comparison of the three pairwise comparisons shows that both Structural only searches (.6 per subject) and the Both (.77 per subject) searches were higher ($p < .01$) than Superficial searches (.13 per subject). There was no difference between the Structural only and the Both search frequency.

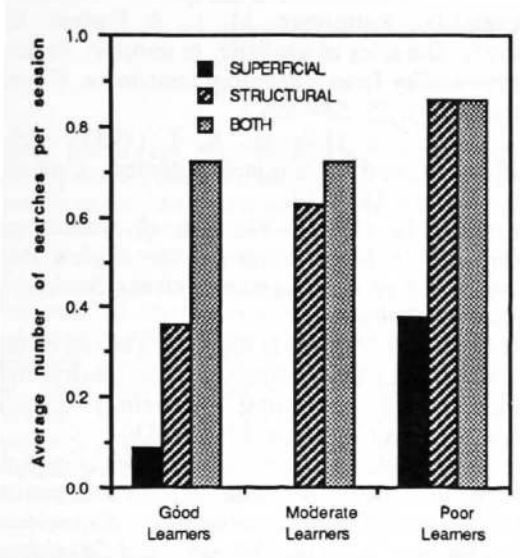


Figure 2: Number of searches per session by ability of the student.

## Good Versus Poor Subjects

Although the frequency of search activity within our post performance ability groupings was too low to really test for significance (e.g., the number of poor ability subjects was 8) we report them here because they are suggestive. In particular, they may soften the strength of our argument about the non-contribution of surface similarity to reminding patterns. Overall, Poor subjects conducted an average of 2.1 searches per Target session (out of 6 possible). Moderate subjects conducted 1.4 searches and good subjects conducted 1.1 searches per Target session. These differences are not significant ($p > .4$) but are in accord with the typical finding that subjects who struggle with a problem are more likely to be reminded of earlier examples and use those examples more frequently (e.g., Ross, 1989). While most of the searches conducted for Superficial Only source problems were by the poor subjects (.75) there were too few superficial searches altogether (total = 4) to make much of this difference. Good learners tend also to search slightly less often for items that are similar only in terms of structural properties than for those that are also superficially similar and, although this is not reliable, it suggests there may be circumstances under which surface features facilitate retrieval of structurally similar problems.

## Discussion

We have replicated our earlier results that novices are capable of noticing and using structural similarity. Very few problems were retrieved that were only superficially similar. People are not misled by superficial similarity when they are engaged in learning and when they participate in the development of the solutions to the examples. Moreover, we have provided support for our prior contention that subjects in our experiments who demonstrate predominantly structural search behavior do not simply learn that superficial similarity is not useful. In this experiment, it predicted structural similarity half the time so was of no help in forecasting whether the easily notice superficial similarities would correlate either way with structural similarity. The addition of superficial similarity to the structurally similar source and target problems made little or no difference to the likelihood that subjects would retrieve a useful structurally related prior problem.

It appears from these results that the effects of superficial and structural similarity are additive but that the contribution of superficial similarity is greatly overwhelmed by the contribution of structural similarity. In order to reconcile these results with a 2 stage model of retrieval such as MAC/FAC (Gentner, et al., 1993) one could suggest that during the second stage or the reminding process these spurious remindings were ruled out and that subjects only pursued remindings that were structurally sound. We certainly cannot rule such an explanation out based on data from this experiment but the data reported here occur exclusively during attempts to solve target problems for which the subject lacked knowledge (therefore opted to search). This absent knowledge may be the same knowledge needed to discriminate between candidate analogs.

As argued, these findings contradict some interpretations of other research that has shown analogical or structural remindings to be difficult to elicit. We have claimed that our experiment differs from other experiments in that the subjects encode the source problems within the context of trying to learn a new skill. In fact, we argue that our subjects were very likely to develop the solutions themselves and that the development of the solutions are part of the memory (Carbonell, 1986). The significance of developing solutions for memory for the problem is that the reasoning steps taken toward the solution are part of an event memory for the source problems. We argue that models used to explain reminding during problem solving should not operate on static descriptions of problems. The event as stored in memory is a trace of the procedures and reasoning used to understand the

event or solve the problem. Similarities between past and present processing and reasoning stages produce remindings. Things are considered to be similar if similar processes are used to understand them or, in the case of problem solving or example studying, if similar processes are used to solve components of the problem.

When students initially read an example or begin to solve a problem the processing of superficial features predominate. If this were as far as they went their representation of the problem would be heavily determined by the processes used to understand the superficial aspects of the problem. As students develop a solution or elaborate on the problem, however, the processes used to understand the superficial components become just a few of many processes used to understand and solve the problem. In cases where subjects solve the source problem (storing solution relevant procedures as part of the event memory) and solve the target problem (thus generate appropriate procedure related cues) we expect that the influence of superficial similarity is greatly diminished. This is essentially what we observe in the behavior of our subjects. In less ideal cases, where subjects do not solve the source problems or if they are lazy about solving the target (therefore do not generate appropriate cue conditions) superficial similarity will provide a stronger influence.

Our findings suggest that models of retrieval should modify the role of superficial similarity as a general strong force in reminding. In cases where subjects participate in the development of the solution, as our results show, it will have only a mild influence, even on the memories of novices who cannot likely articulate the structural similarities that form the bases for their remindings.

## Acknowledgments

This work is supported in part by the Office of Naval Research grant number N00014-91-J-1742. We thank Brian Reiser for comments on the design of these experiments. We would also like to thank Tim Wienrich for help in programming the data acquisition software and Joan Stahl for her help in collating data and transcribing protocol data.

## References

Anderson, J. R., Corbett, A. T., & Reiser, B. J. (1987). *Essential LISP* . Reading, MA: Addison-Wesley.

Carbonell, J. G. (1986). Derivational analogy: A theory of reconstructive problem solving and expertise acquisition. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning*. Los Altos, CA: Morgan Kaufmann.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data* . Cambridge, MA: MIT Press.

Faries, J. M. (1991) *Reasoning-based retrieval of analogies*. Doctoral Dissertation, Princeton University.

Faries, J. M., & Reiser, B. J. (1988). Access and use of previous solutions in a problem solving situation. In *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*, (pp. 433-439). Montreal:

Faries, J. M., & Reiser, B. J. (1989). *BATBook: An on-line book and problem solving environment for the study of skill acquisition* No. Cognitive Science Laboratory, Princeton University.

Gentner, D., Ratterman, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, 25, 524-575.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, **15**, 1-38.

Hammond, K. (1986). The use of remindings in planning. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, (pp. 442-451). Erlbaum.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 629-639.

Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 456-468.

Seifert, C. M. (1988). Goals in reminding. In *Proceedings of DARPA Workshop on Case Based Reasoning Workshop*, . Clearwater Beach, FL:

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

VanLehn, K. (1989). Problem solving and skill acquisition. In M. I. Posner (Eds.), *Foundations of Cognitive Science* Cambridge, MA: MIT Press.

VanLehn, K., & Jones, R. M. (1993). What mediates the self-explanation effect? Knowledge gaps, schemas or analogies? In *15th Annual Conference of the Cognitive Science Society*, (pp. 1034-1039). Boulder, Colorado: Hillsdale NJ, Lawrence Erlbaum Associates.