Converging on a New Role for Analogy in Problem Solving and Retrieval

Kenneth J. Kurtz (kkurtz@binghamton.edu)
Department of Psychology, P.O. Box 6000
Binghamton, NY 13902 USA

Jeffrey Loewenstein (jeffrey.loewenstein@columbia.edu)
Columbia Business School, 3022 Broadway
New York, NY 10023 USA

Abstract

A novel approach to generating retrieval and transfer of structured knowledge is presented. We investigate the effect of comparing two analogous unsolved problems at test as opposed to comparing two solved analogous stories during initial study. We found that both procedures facilitate transfer relative to a standard baseline group studying one solved story and then attempting to solve a new analogous problem. In two studies we demonstrate that: 1) comparing two unsolved problems at test promotes analogical problem solving at least as effectively as comparing two fully solved problems during study; and 2) comparing two unsolved problems is helpful even when no source story is made available for retrieval.

Introduction

There is a wealth of cognitive science research about how people learn from examples and use them to solve new problems (Reeves & Weisberg, 1994). We also know that people are unlikely to spontaneously compare examples that seem different on the surface even though such comparison can provide learning and transfer advantages (Gick & Holyoak, 1983; Kurtz, Miao & Gentner, 2001). Retrieving analogous matches is therefore both important and demonstrably difficult. Research on retrieval shows that people have an easier time accessing examples on the basis of surface features than structural match (Catrambone, 2002; Gentner, Rattermann & Forbus, 1993; Ross, 1987). It is not that structural matches, particularly partial matches, are impossible or even rare, just that surface matches tend to predominate among novices, whereas experts seem able to exhibit structure matches more reliably (Dunbar, 2003; Novick, 1988).

We know that comparing examples can lead people to focus on common systems of relations which can in turn facilitate knowledge transfer (e.g., Loewenstein, Thompson & Gentner, 1999). The conventional wisdom in the field is that upon encountering a test problem, people are able to retrieve the earlier analogous cases, or a schema abstracted from those cases, to generate potential solutions to the problem (Gick & Holyoak, 1983). The implication is that the similarity function used in memory retrieval can link a current case to a prior case if the prior case is represented well in long term memory. The specific nature of such a superior representation is a challenging question for the field, but we take as a starting point the idea that a good representation is one that accurately encodes pertinent systems of relational structure and does so with sufficient generality to support transfer. This generality can be considered in terms of domain generality of encoded relational content (Clement, Mawby & Giles, 1994), uniformity of representational elements (Forbus, Gentner, & Law, 1995), or filtering out of mismatching irrelevant case details (Hummel & Holyoak, 1997).

We are currently intrigued by a new role for analogy in memory retrieval (see also, Loewenstein, Gentner, & Thompson, 2004). Our question is whether the benefits of this kind of representational "improvement" to the analogical source might also be observed with respect to the target (probe). Is a structural reminding more likely with a target that is better encoded? Theories of memory retrieval rely on a similarity function between the probe and stored items. Such similarity functions are symmetric. Since the empirical data suggest that only one side (i.e., the source) needs to be well-encoded to encourage a match, then it is plausible that a relevant, but regularly encoded source might become more retrievable on the basis of applying a probe with a superior encoding. In addition to being a theoretical possibility, there is a phenomenon, admittedly rare, of recalling an example with the sense of having a new understanding about it as a result of something we have just learned. The current line of thinking could explain such occurrences. Furthermore, it suggests a mechanism by which reflection upon a newly learned principle or abstraction could be a prod to retrieve prior examples, reinterpret them, and integrate the new knowledge with the old. Drawing analogies might then not only be a source of changes in knowledge from this point forward, but could also be a means for reorganizing the knowledge we already have and retrieving further analogous matches.

To reiterate, one of the seminal findings in the analogy literature is that problem solvers are more successful in retrieving an available solution strategy when they have previously made use of comparison to improve the encoding of the source analogs (Gick & Holyoak, 1983). We adapt this highly influential paradigm to ask the following question: Can comparison of target problems be used to facilitate analogical retrieval? There is considerable reason for skepticism. First, the advantage of source comparison is thought to rely on storing a generalized version of the solution principle, but in the case of target comparison the solution is not part of the compared cases—only the two problem statements are available. Secondly, the traditional

account suggests that structural reminding depends on having a well-represented source in memory—it may well follow that structural reminding is largely a dead end without a well-encoded source. Third, it is easy to imagine that having two problems to solve rather than just one could divide attention and processing resources in a detrimental manner. Finally, there is an extensive tradition of failed attempts to improve analogical problem-solving performance. Even so, if comparison at test can improve the encoding of targets such that retrieving structural matches is facilitated, this would have significant theoretical and applied ramifications. In the following two studies, we explore comparison-improved representation at the point of actual problem solving in hopes of gaining new insights into learning, retrieval, and transfer.

Experiment 1

In the current studies we use classic materials to study a novel set of questions about analogical problem solving: Duncker's (1945) tumor problem and its associated materials generated by Holyoak and colleagues (Gick & Holyoak, 1980, *inter alia*). In Experiment 1, we use these materials to examine whether retrieval is a two-way street. That is, if comparing two examples at study facilitates transfer at test (as shown by Gick & Holyoak, 1983), then can comparing two examples at test facilitate retrieval from study? In addition to the comparison being on-line rather than during initial study, the other key difference is that compared target problems do not include the solution.

We include two conditions to replicate prior data: a baseline group receiving one solved story at study and one problem at test (which will presumably yield little transfer) and a group comparing two solved stories at study and then receiving one problem at test (which will presumably show transfer). The key question is what will result in a new condition with one solved story at study and a comparison of two unsolved problems at test. Will participants who compare two test problems show greater success than participants in the baseline condition? Furthermore, to address the question of whether success hinges on transfer via retrieval of the source story, we include a group asked to compare and solve two problems without having first seen a solved story at study. Participants in this group are the only ones to receive no exposure at all to the relevant solution strategy.

Method

Participants A total of 293 undergraduate students at Binghamton University participated in partial fulfillment of a course requirement. Participants were randomly assigned to one of four conditions: Baseline, Source Comparison, Target Comparison, or Just Targets.

Materials The target case in all conditions was the well-known Radiation problem developed by Duncker (1945) and further studied by Gick & Holyoak (1980, 1983). The source and comparison cases were analogs based on the convergence principle used by Gick & Holyoak (1983). The source case was "The General" set in a military context and the comparison case was "Red Adair" set in a firefighting context. The comparison case was given with solution included during the study phase in the Source Comparison condition. We used the same Red Adair problem for comparison at test without the last lines that give the convergence solution in the Target Comparison and Just Targets conditions.

Procedure All phases of the experiment were conducted using paper packets for the presentation of instructions and materials as well as for the collection of responses. Separate packets were created for study and test phases. Participants did not receive the test packet until they completed and handed in their study packet (if a study packet was required by their condition).

In the Baseline condition (1:1, meaning participants were given 1 source story and 1 target problem), participants were instructed at the beginning of Part I to read the story (General) carefully and to gain sufficient familiarity that they could retell the story in their own words. Toward the bottom of the page, participants were asked: "What critical insight allowed the problem in the story to be solved?" In Part II, participants were asked to read the problem (Radiation) and to "use the space at the bottom of the page to explain how the problem can be solved."

In the Source Comparison condition (2:1), participants were instructed in Part I to carefully read two stories (General and Red Adair). The two stories were shown on the same page with General appearing first. At the top of the second page were two tasks to encourage better encoding. As in the control condition, participants were asked to gain sufficient familiarity that they could retell the stories in their own words. In addition, participants were asked to "Consider the parallels between the two stories" and complete a task in which five elements of Column A (General) had to be matched with elements of Column B (Red Adair). Each element had exactly one appropriate match. The columns were prepared in a jumbled order so that no correctly corresponding elements were directly across from one another. In Part II, participants were asked to solve the Radiation problem. The exact same procedure was used as in the Baseline condition.

In the Target Comparison condition (1:2), participants were presented with the General story using the exact same procedure as in the Baseline condition. In Part II, these participants were given two problems to solve (Radiation and Red Adair) The first page of the packet gave the following instructions: "What approach would you take to solve both of the following problems? After reading the problems carefully, please complete the matching task and then explain your proposed solutions in the space provided. Here's an important hint: The same strategy can be used to solve both problems."

Below the instructions were the two problems: Radiation followed by Red Adair. On the second page was a matching task between the Radiation and Red Adair problems

constructed in the same manner as above. On the last page, participants were again given the hint that "The same type of solution can be used" and asked to: "Please write down how these two problems can be solved."

In the Just Targets condition (0:2), participants were given Part II of the Target Comparison condition only. That is, they were asked to solve the Radiation and Red Adair problems without any prior exposure to the General story and its convergence solution. There was one additional procedural difference. The same-solution hint was provided, but in this case it was given only at the point when participants were actually asked to produce their solutions (rather than mentioning the hint twice).

An important point to clarify here is that this hint is distinct in type from the well-known use of a hint in the Gick & Holyoak studies. In that prior work the hint was to use the initial story as a basis for solving the target problem. That hint removed the fundamental obstacle in analogical problem solving: achieving a spontaneous structural reminding. The manipulation was of critical theoretical importance since it revealed that the Radiation problem was widely solved once participants accessed the source analog. In our current work, the hint has nothing to do with retrieval; instead it enforces mutual consideration of the two target problems.

Scoring A rater blind to condition scored each response for success in solving the Radiation problem in terms of the convergence solution. Responses were scored as correct if they captured the key principle of a multiplicity of low-intensity rays acting in concert on the tumor. The rater marked any responses they considered questionable for discussion with another rater. The agreed-upon scoring was then recorded. In occasional cases in which more than one solution was proposed, participants were given credit for the correct answer if it was produced.

Results

As expected, we replicated prior data showing that people who compared two source stories showed a transfer advantage relative to a baseline group who only read one source story (as shown in Table 1, 38% vs. 13% generated convergence solutions), $\chi^2(1, N=146) = 12.12$, $p < .01$. The important new result is that the group comparing at test, rather than study, performed as well or better than all other groups. The Target Comparison group performed better than the Baseline group (51% vs. 13%), $\chi^2(1, N=142) = 24.06$, $p < .001$. There was also a trend toward better performance by the Target Comparison group than the Source Comparison group (51% vs. 38%), $\chi^2(1, N=142) = 2.62$, $p = .11$. Critically, the Target Comparison group also performed better than the Just Targets (25%) group, $\chi^2(1, N=147) = 10.58$, $p < .005$. This suggests that participants who compared target problems were drawing upon the story from study since this was the major difference between the Target Comparison and Just Targets conditions. Finally, the Just Targets participants were marginally more likely to derive the convergence solution than were Baseline participants, $\chi^2(1, N=145) = 3.62$, $p = .06$. This suggests that comparing two unsolved target problems facilitated reaching the correct solution as compared to the Baseline condition of single cases at study and test.

Table 1: Proportion of convergence solutions by condition

Condition	N	Proportion generating convergence solution
Baseline (1:1)	70	.13
Source Comparison (2:1)	76	.38
Target Comparison (1:2)	72	.51
Just Targets (0:2)	75	.25

Discussion

We were able to replicate the well-known finding that comparing two examples at study yielded transfer benefits at test relative to a control group reading just one story at study. The intriguing result is that comparing two problems at test resulted in higher performance than the control group. Perhaps even more surprising, the Target Comparison group performed slightly, but not significantly, better than the comparison at study group. Confronted by one hard problem, these results suggest that a reasonable course of action would be to seek another problem with the same underlying structure!

Prior research suggests that abstracting the convergence schema was critical for success on the Radiation problem. Yet it is highly unlikely that the Target Comparison group abstracted the convergence schema from one example (otherwise the control group should have done well too). Our interpretation is that comparing analogous problems can lead to better representations of one or both problems. Such an encoding is likely to serve as a more effective retrieval cue for analogical problem solving. Due to having better representations of the problems via comparison, participants recalled the initial source story and borrowed its convergence solution. That is, retrieving prior examples on the basis of structure might be feasible if the probe is sufficiently well encoded, just as the comparison at study condition suggests that retrieval is feasible if the stored item is sufficiently well encoded.

The lower level of performance by the group who compared test problems, but did not receive a story at study, provides support for the claim that retrieval was a factor. Further tests are needed however to determine whether the single versus repeated hint played any role in this finding. The marginal advantage obtained in the Just Targets (0:2) condition over the Baseline condition indicates potential, not only for problem comparison as a means to achieve analogical retrieval, but also as a means to generate problem insight right then and there via analogical encoding.

In sum, we found that performance on a difficult problem can be greatly facilitated by an on-line technique. It is not necessary to construct improved representations at the time of encoding because one can do the necessary work through

comparison at test. Furthermore, such comparison is between problems, not between solved stories. The power of this comparison is not based on highlighting the convergence principle, but arises from comparison of two problem scenarios both amenable to a convergence solution.

Experiment 2

A second study was designed to replicate our basic finding and to further test whether drawing comparisons was an important component of the Target Comparison group's strong performance in Experiment 1. In this study we contrast the Target Comparison condition with a condition also receiving one study story and two test problems, but not guided with a hint to seek one solution for both problems. This Separate Targets condition still includes a matching task and the task to write down how "these problems can be solved," but the specific suggestion to work toward a single solution strategy is removed. If the Target Comparison group outperforms the Separate Targets group, this would serve as an indication that the depth of comparison of the problems is critical, just as comparing study problems is critical (Catrambone & Holyoak, 1989; Loewenstein, et al., 1999; Kurtz, et al., 2001). Additionally, in Experiment 1, the Source Comparison group tended to perform less well than the Target Comparison group, so a Source Comparison condition was included to test for a reliable difference.

Method

Participants A total of 224 undergraduate students at Binghamton University participated in partial fulfillment of a course requirement. Participants were randomly assigned to one of three conditions: Source Comparison, Target Comparison, or Separate Targets.

Materials, Procedure and Scoring The same source and target cases, the same use of paper packets, and the same scoring procedures were used as in Experiment 1. The Source Comparison (2:1) and Target Comparison (1:2) conditions were conducted using the same experimental and scoring procedure as in Experiment 1. The Separate Targets (1:2 without hint) condition followed the Target Comparison condition exactly with the exception that the initial hint and hint repetition were excluded from the text of the instructions.

Results

The main focus of this study was the contrast between the Target Comparison and Separate Targets conditions. People who received two problems, but no hint to compare them generated the convergence solution infrequently (16%, see Table 2). As in Experiment 1, the Target Comparison group frequently generated convergence solutions (40%), and did so reliably more often than did participants in the Separate Targets condition, $\chi^2(1, N=147) = 10.77$, $p < .005$. Thus an explicit instruction to compare and generate a common solution was critical to the effectiveness of the Target Comparison manipulation.

There was little difference between the Source Comparison (35%) and Target Comparison (40%) groups in this study, $\chi^2 < 1$. The previous study suggested there might be a difference between the two conditions, and the ordering of the means was consistent, but the difference in this study was minimal.

Table 2: Proportion of convergence solutions by condition

Condition	N	Proportion generating convergence solution
Source Comparison (2:1)	77	.35
Target Comparison (1:2)	72	.40
Separate Targets (1:2) without hint	75	.16

Discussion

This study replicated the effectiveness of comparing two target problems. It also confirmed an important boundary condition, namely that comparing the target problems toward drawing out a common solution was important. Merely receiving two target problems with minimal encouragement to assess their parallels was not effective. Indeed, solving two target problems separately led to comparable performance as solving one target problem (i.e., the baseline condition) in Experiment 1.

General Discussion

With these two studies, we provide grounds for a new emphasis, if not a new interpretation, of analogical retrieval and transfer. The usual assumption is that comparing examples facilitates generation of a representation of the common schema that clarifies the key structure and is less cluttered by unrelated contextual details than the original examples. It is clear that without drawing a comparison, people are unlikely to represent the structure in such a way that it can be retrieved and used to solve a new problem—an effect we replicated in Experiment 1. The current results open up the possibility that the benefit of comparison at study may be due to: 1) improving the encoding of the examples rather than creating a new general knowledge structure; or 2) allowing people to form better encodings of subsequent cases using a more sophisticated or general representational vocabulary.

The current studies were aimed at addressing this issue by turning it around: what if people study just one example (so they are unlikely to form any particularly clear or uncluttered representation), but they compare examples at test and then profit from having read the earlier single case. The results of our two experiments are consistent with people being able to retrieve single stored cases in just this fashion. We showed a distinct transfer advantage for a group that was: (1) specifically encouraged to compare two unsolved test problems and (2) had previously studied a single case. One may not need to store cases in a

particularly good fashion if one can later construct a superior retrieval probe.

There are multiple implications if a comparison today can facilitate retrieving a case learned yesterday. First, with respect to models of the retrieval process, it suggests a constraint on the similarity process that matches stored items to probes: it may well have to be symmetric. Second, it suggests a mechanism by which people can reorganize their knowledge. One may not have to "learn it right the first time" if, after appreciating a new abstraction, one is able to retrieve and perhaps re-represent a prior matching example. This supplies a concrete mechanism for gradual conceptual change in both development and the acquisition of expertise. Third, this implies that educators, particularly those who teach adults, can look to integrate people's prior experiences in their formal acquisition of domain expertise.

A second point arising from these studies is that drawing comparisons can facilitate learning in a new way. Typically, people draw comparisons to understand a principle or solution in a more general way. In our studies, people used comparison to generate better formulations of the problem at hand, not a better understanding of provided solutions. There are at least three reasons as to why this should facilitate problem solving. First, comparing two problems with the knowledge that that they have a common solution type means that idiosyncratic information can be ignored. Second, potentially misleading example-specific solution types can be ruled out. Third, it may allow people to formulate a more abstract or general version of the problem at hand. As Polya (1945) suggested, despite it seeming counterintuitive, sometimes a more general problem is easier to solve than a more specific problem. There may be interesting and important applications of this use of comparison both in education and in discovery.

We find these studies an intriguing first step. We are pursuing several related issues that might influence our interpretation of these studies. The Just Targets condition was given a weaker hint to compare than the Target Comparison condition, and as the Separate Targets condition showed: hints are important. We are running a new study that examines equal encouragement to draw comparisons. We are also interested in whether the Target Comparison condition benefits from one problem being easier to solve than the other (in which case its solution would be tested on the second problem) or whether it is the development of a more general version of the problem that is driving participants' success.

In conclusion, drawing comparisons may facilitate learning and transfer in multiple ways. It may enhance recalling prior experiences as much as generating knowledge that is likely to be later transferred. It may enhance clarifying a problem formulation as much as deriving generalizations from solved problems.

Acknowledgments

Special thanks to Jessica Federman and Aliza Nelson. We also thank Janeen O'Connor, Olga Boukrina, and the rest of the Learning and Representation in Cognition (LaRC) Laboratory at Binghamton University.

References

Catrambone, R. (2002). The effects of surface and structural feature matches on the access of story analogs. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*(2), 318-334.

Clement, C.A., Mawby, R., & Giles, D.E. (1994). The effects of manifest relational similarity on analog retrieval. *Journal of Memory and Language, 33*, 396-420.

Dunbar, K. (2003). The analogical paradox: Why analogy is so easy in naturalistic settings yet so difficult in the psychological laboratory. In D. Gentner and S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought*. (pp. 313-334) Cambridge, MA: MIT Press.

Duncker, K. (1945). On problem-solving (L. S. Lees, Trans.). *Psychological Monographs, 58*(5), Whole No. 270.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*(2), 141-205.

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

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

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

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*(3), 427-466.

Kurtz, K. J., Miao, C., & Gentner, D. (2001). Learning by analogical bootstrapping. *Journal of the Learning Sciences, 10*(4), 417-446.

Loewenstein, J., Gentner, D., & Thompson, L. (2004). Analogical Encoding: Facilitating Knowledge Transfer and Integration. Working paper, Columbia University.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review, 6*(4), 586-597.

Novick, L. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*, 510-520.

Polya, G. (1945). *How to solve it*. Princeton, NJ: Princeton University Press.

Reeves, L. M., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin, 115* (3), 381-400.

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 & Cognition, 13*(4), 629-639.