

Not taking the easy road: When similarity hurts learning

Hee Seung Lee · Shawn Betts · John R. Anderson

Published online: 18 February 2015
© Psychonomic Society, Inc. 2015

Abstract Two experimental studies examined the effects of example format and example similarity on mathematical problem solving across different learning contexts. Participants were more successful inducing a correct problem-solving rule when they were provided with annotated examples rather than nonannotated examples. The effects of example similarity varied depending on learning context. In Experiment 1, by presenting an example and problem simultaneously, a direct comparison was possible between the cases. When the examples were similar, participants relied on superficial analogies that hurt learning. When an example was dissimilar from the given problem, participants appeared to study the example first to induce a solution procedure and then apply the rule to the problem, thus resulting in better learning and transfer. However, in Experiment 2 where the example and problem were presented in a sequential manner, the effect disappeared because the learning context did not support a direct comparison. We conclude that comparison is not inherently good for promoting learning and transfer, rather its effect depends on whether it supports relational mapping that is essential for schema acquisition.

Keywords Example format · Example similarity · Mathematical problem solving · Comparison · Mapping

Electronic supplementary material The online version of this article (doi:10.3758/s13421-015-0509-3) contains supplementary material, which is available to authorized users.

H. S. Lee (✉)
Department of Education, Yonsei University, 50 Yonsei-ro,
Seodaemun-gu, Seoul, Korea
e-mail: hslee00@yonsei.ac.kr

S. Betts · J. R. Anderson
Department of Psychology, Carnegie Mellon University,
Pittsburgh, PA, USA

Introduction

A large number of studies have shown that comparing multiple examples promotes learning by facilitating abstraction and discovery of deep relational structure (e.g., Ankowski, Vlach, & Sandhofer, 2013; Gentner, Loewenstein, & Hung, 2007; Gentner, Loewenstein, & Thompson, 2003; Gick & Holyoak, 1983; Loewenstein, Thompson, & Gentner, 1999; Namy & Gentner, 2002; Oakes & Ribar, 2005; Rittle-Johnson & Star, 2007). These studies generally suggest that simultaneous presentation of multiple instances is better than sequential presentation of each instance. For example, Loewenstein et al. (1999) showed that in learning negotiation strategies, students were more likely to choose better negotiation strategies when they drew an analogy between examples than when students were provided with the same examples separately. Likewise, in the domain of mathematics, Rittle-Johnson and Star (2007) demonstrated that students gained greater procedural knowledge and flexibility in the equation-solving task when they simultaneously compared alternative solution methods than when they studied the same solution methods in a sequential manner.

Sharing and comparing solution methods is now believed to be an effective learning strategy and a high-quality instructional method. Experienced teachers tend to emphasize the importance of comparing activity (Ball, 1993; Fraivillig, Murphy, & Fuson, 1999; Huffred-Ackles, Fuson, & Sherin, 2004; Lampert, 1990; Silver, Ghouseini, Gosen, Charalambous, & Strawhun, 2005). National Council of Teachers of Mathematics (NCTM, 1991, 2000) also advises teachers to encourage students to share and compare their own problem-solving methods with others.

When does similarity help or hurt learning?

Although there is strong empirical evidence *and* strong belief that comparison helps learning, it is not clear about how

similar or dissimilar examples should be when comparing them (Renkl, Stark, Gruber, & Mandl, 1998; Rittle-Johnson & Star, 2009). Is it better to provide similar examples, or is it better to provide dissimilar examples to promote learning and transfer? Many studies that have shown benefits of similar examples claim that similarity among examples facilitates structure mapping and helps the learner to identify deeper relational commonalities. When examples are very dissimilar it may make mapping harder; thus, it may be harder to discover common structural features between the cases (Gick & Holyoak, 1983; Richland, Holyoak, & Stigler, 2004; Ross, 1989a; Ross & Kennedy, 1990). For example, many developmental studies demonstrated that similarity helped learning by bootstrapping, especially early in the learning process (Gentner & Rattermann, 1991; Kotovsky & Gentner, 1996; Namy & Gentner, 2002). According to the progressive alignment hypothesis experience in concrete similarity prepares learners for abstract similarity. Alignment on the basis of concrete features leads to alignment of more general representations (Gentner & Markman, 1993; Goldstone, 1994; Goldstone & Medin, 1994; Markman & Gentner, 1993; Novick, 1988). Therefore, young children can use surface-level similarities as a starting point to further find more abstract, relational commonalities. For example, Gentner and Namy (1999) reported that young children use perceptual similarity to evoke the alignment process. Similarly, Holyoak, Junn, and Billman (1984) reported that a minor surface dissimilarity significantly decreased the percentage of analogy among children populations. Some researchers also suggested that dissimilar examples might overtax learners and prevent them from focusing on structural aspects (Renkl et al., 1998; Ross, 1989b).

However, similarity does not always lead to better learning. When similar examples share domain-specific features that are irrelevant to the problem schema, it may prevent acquisition of a more general and abstract schema (Gick & Holyoak, 1983; Merrill & Tennyson, 1978; Quilici & Mayer, 1996; Reed, 1989; Tennyson, 1973). Several studies have distinguished surface and structural similarity (Holyoak & Koh, 1987) and demonstrated that students sometimes use irrelevant surface problem features as heuristics to solve problems. A well-known example of disadvantage of similarity is that of cross-mapping. Cross-mapping is a comparison, where two analogous situations contain similar objects that play different relational roles (Gentner & Toupin, 1986). In Gentner and Toupin's (1986) study, children showed worse performance when similar characters played different roles rather than similar roles. Many other studies also have shown that people are often misled by surface similarity, and transfer accuracy decreases when surface similarity does not correlate with structural similarity (Reed, 1987; Ross, 1987). For example, Ben-Zeev and Star (2001) found that when there were spurious correlations between an irrelevant problem feature and the algorithm used for problem solution, they used this misleading information when solving new problems.

The studies reviewed above can be largely summarized by two conclusions: (1) Similarity helps learning when it supports alignment process and structural mapping between examples; (2) similarity harms learning when it guides learners' attention to irrelevant features. Thus, it may not be a simple matter of whether similarity helps or hurts. Rather, it seems that advantages or disadvantages of example similarity depend on whether a learner's attention is directed to relevant or irrelevant features for attaining the learning goal (e.g., problem solution). Accordingly, we predict that similarity will help learning only if it elicits a relevant mapping and promote abstraction of a schema in a learning context. This proposal is quite consistent with that of structure-mapping theory (Gentner, 1983; Gentner & Markman, 1997). Structure-mapping theory predicts that similarity among examples facilitates structure mapping and helps the learner to identify deep relational commonalities. Also, it emphasizes relational similarity over surface similarity; thus, similarity between cases should help learning when it is correlated with relational similarity. In the current study, we investigate these issues in a domain of mathematical problem solving, focusing on how effect of similarity interacts with learning context by systematically manipulating whether it supports superficial or relational mapping.

Two alternative mappings in mathematical problem solving

Mathematical problem solving offers a particularly clear example of a structure that characterizes many problem-solving domains. Students are faced with a problem (e.g., an equation) and must find a solution (e.g., a set of transformations solving the equation). An example presents both the problem and a solution, while a test item provides just the problem. In terms of the proportion framework problem for analogy (Sternberg, 1977), the student is being asked to solve a problem of the form " $A:B :: C:X$ " where A refers to the example problem, B refers to the answer of the example problem, C is a new problem, and X is the answer to the problem C . A problem can be solved by finding the mapping between A and C first, and then applying that mapping to B to find the answer X . Alternatively, it can be solved by finding the relation between A and B first and then applying that relation to C in order to find the answer X . While these two methods may be equivalent for some kinds of analogy (e.g., Raven's Progressive Matrices Test; Raven, 1938) they are not equivalent for learning mathematical problem solving. The goal is not to solve a specific problem but to learn general rules for solving any problems. From this perspective, the second way of using the example is to be preferred.

One factor that might influence which method a student attempts is similarity between an example and problem. For instance, if a student is learning to combine like terms in a problem, like $[6x - 3x + 7x = ?]$, they might be shown a very similar example, like $[5x - 3x + 7x = 9x]$. Because the example

and problem are identical except for the first component (from $5x$ to $6x$), the answer can be derived by the superficial rule “increasing the integer number of the example answer by 1” (from $9x$ to $10x$), resulting in the correct answer. On the other hand, if the example were different, like $[5x + 2x - 3x + 4x = 8x]$, it is hard to map components between the example and test item; thus, it would be better to figure out the relation $[5x + 2x - 3x + 4x:8x]$ and then apply this relation, “8 is the result of performing arithmetic operations among four numbers in order” to the test problem.

We expect that students’ mapping preference will differ depending on how easily they can see similarities between the example and problem, and this will in turn affect learning of problem-solving skills. Although there could be multiple factors that contribute to ease of recognizing similarities between the example and problem, in the current study we focus on two things: (1) how perceptually similar or dissimilar they are, and (2) whether they are presented simultaneously or sequentially.

Current study

The current study aimed to investigate how example similarity affects the mapping and solution strategy learners adopt in mathematical problem solving. We expect that when an example and problem look perceptually similar it will be relatively easy to draw a mapping between the example and problem and thus increase the chance of superficial strategy between the example and the problem. In contrast, when an example and problem look perceptually dissimilar, it could be hard to draw a direct mapping between them, and thus students will be more likely to choose the strategy of abstracting the relation in the example first and then applying it to the problem.

However, we also expect that such negative effect of similarity will depend on learning context. More specifically, it will depend on whether examples and problems are presented simultaneously or sequentially. When a similar pair of example and test problems are presented simultaneously, similarity between the example and problem will be likely to promote the wrong strategy of mapping the example to the problem. However, when the identical pair of example and test problems are presented separately in a sequential manner,

similarity will not be apparent, and thus students will not be likely to use a superficial mapping.

To test this idea, learning context was varied in two experiments. In one experiment, a problem and example were presented simultaneously (Experiment 1), and in the other experiment, they were presented sequentially (Experiment 2). Figure 1 shows the schematic representation of procedures used in each experiment. E and P represent an example and a problem, respectively. The E and P inside the one single rectangle indicate that they were presented simultaneously on the same page. In Experiment 1, participants were provided with a problem to solve and were able to request an example for each problem. The requested example appeared next to the problem on the same page, and it was either similar or dissimilar to the problem. In Experiment 2, a single example appeared before each problem on a separate page, and it was either similar or dissimilar to the problem. Because direct comparison is not easy in this sequential learning context, we expect that the effect of example similarity will be abolished.

Along with the manipulation of example similarity, the current study also examined the effect of example format. Participants were given either a fully worked-out example with final answers (annotated example condition) or an example with final answers only (nonannotated example condition). In the annotated example condition, intermediate procedure steps were shown. These annotations were constructed based on some of our past research (Anderson, Lee, & Fincham, 2014; Lee, Fincham, Betts, & Anderson, 2014), where they were proven to be effective for learning. In the nonannotated example condition, participants had to induce procedure steps for themselves based on the example problem with final answers filled in. Our annotation provided clues to facilitate a nonsuperficial analogy (i.e., finding the relationship between A & B and applying it to C in the $A:B :: C:X$ proportion framework). We thought participants might not succeed in finding this mapping without the annotation. To the extent that our similarity manipulation influences the tendency to seek this nonsuperficial analogy, we would expect a stronger effect of similarity in the annotated condition.

Besides learning, we will look at transfer to new types of problems to assess the breadth of what participants have learned. We expect the benefits of annotating the example will be particularly strong in transfer because annotation should

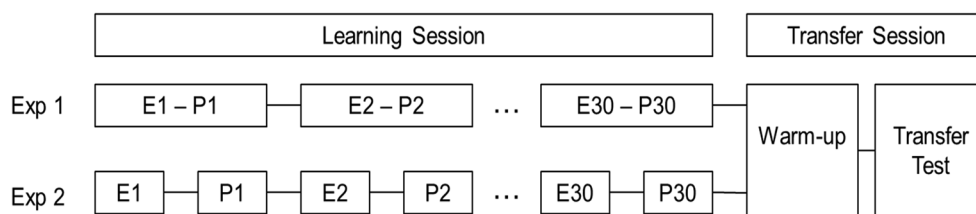


Fig. 1 Schematic representation of the procedure of the two experiments. E and P represent an example and a problem, respectively. The E and P within the same rectangle were presented on the same page

help prevent learning superficial rules that only work for the learning material. Thus, in both studies we tested example format and example similarity in a 2×2 between-subjects design, and this was tested across the two different learning contexts (simultaneous vs. sequential presentation).

Learning materials

We wanted a learning task that would be like algebra but which would be novel to all participants. Therefore, we used a computer-based instructional system like the one used by Lee et al. (2014), that was originally adapted from Brunstein, Betts, and Anderson (2009). Isomorphs of algebraic expressions are represented as data-flow diagrams with multiple boxes and arrows (see Fig. 2, which is the data-flow equivalent of $3 * 4 - 9 = x$). The task is to determine what values to fill into the empty portions of the boxes. The unknown value on the bottom box can be simply determined by “propagating” the number down from the top, performing the arithmetic operations- for the problem of Fig. 2, this involves placing 4 in the empty tile below the top box, then placing 12 in the tile below it (because $3 * 4 = 12$), and finally placing 3 in the bottom unknown box (because $12 - 9 = 3$). In our current and other studies, we observed that most students (both college-level and algebra-level students) found propagating numbers easy and intuitive and thus did not need much assistance when learning to solve this kind of problem.

However, when problems cannot be solved by this simple propagation strategy, students tend to have difficulty understanding the problem structure and inducing a correct

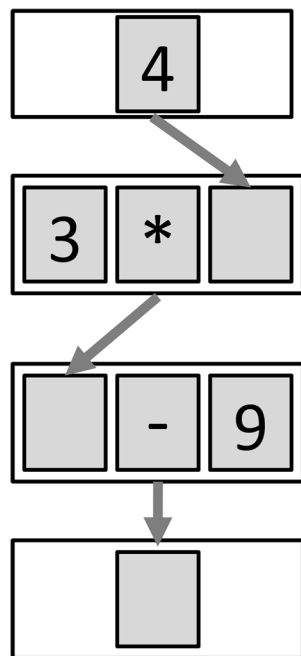


Fig. 2 An example of data-flow diagram. The task is to fill a number into the empty portions of the diagram

procedure required to solve the problem. The current study focused on these harder problems that could not be solved by this simple strategy (The entire experimental materials are available in the [supplementary online material](#)). Figure 3 shows an example problem used in this study. An unknown number flows from the top box into the boxes below, the arithmetic operations are performed, and the final result is another unknown number in the bottom box. The problem requires relating the rightmost path with the colored boxes to the rest of the graph on the left. These two paths converge on the bottom box, and the task is to fill in values into the two colored boxes along the rightmost path so that the same number flows into the bottom box from both paths, no matter what number starts out in the top box. To correctly solve the problem, students have to understand that the values for the two boxes on the right can be calculated by collecting like terms. For example, the left path of the diagram in this figure is equivalent to the algebraic expression $(6 + x) + 13 * (3 + x)$. Essentially, the right path of the diagram has a structure that corresponds to the algebraic equivalent of the result of collecting like terms. That is, $(6 + x) + 13 * (3 + x)$ is simplified into the expression $14x + 45$, which is the structure that needs to be created for the right path. Regardless of the value of the top unknown number, the two alternative paths will have the same operational effect and will always lead to the same resulting number. Therefore, this problem can be solved by putting the coefficient term 14 into the green box next to the “*” operator and the constant term 45 into the yellow box next to the “+” operator.

As described above, making a reference to algebra makes it easier to understand the structure and solution of this problem. When students were given verbal instruction on procedural steps without making a reference to algebra, they had difficulty mastering this problem solving skill (Lee, Anderson, Betts, & Anderson, 2011; Lee et al., 2014). In contrast, when the problem structure was explained with reference to algebra, participants were more successful at mastering the problem-solving skill even without explicit verbal instruction on how to

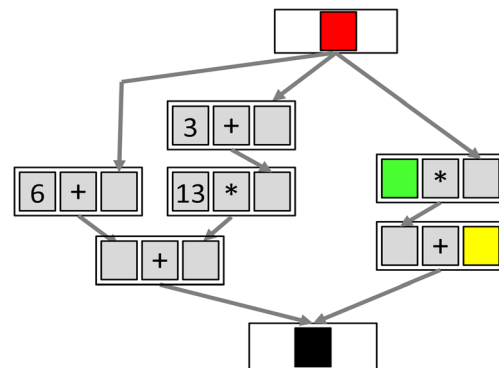


Fig. 3 An example problem used in the learning session of Experiments 1 and 2. The task was to find a number to fill in onto the green and yellow tiles

compute the answers (Lee et al., 2014). The current experiments will include annotated and nonannotated examples to manipulate whether this algebraic connection is made apparent or not in the provided examples.

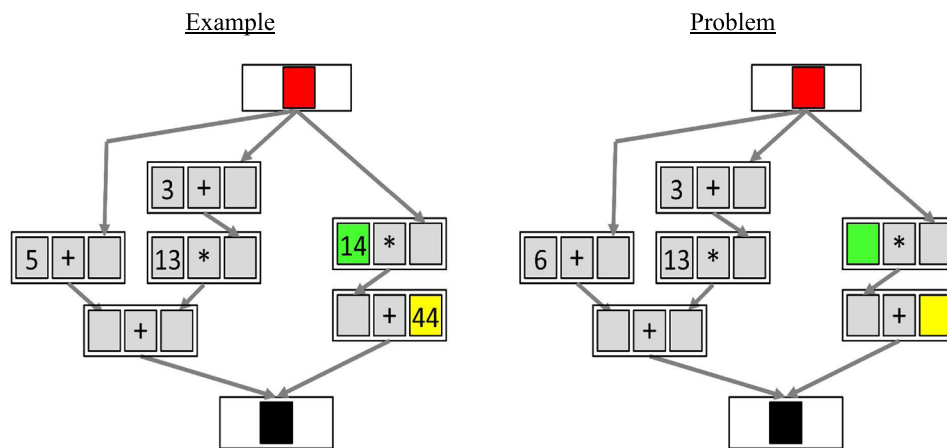
Experiment 1

Experiment 1 investigated the effect of example format and example similarity in a situation where participants can directly compare an example and a problem by having them both available at the same time. As noted in the introduction, there are two possible ways to utilize the example. The first is to draw a direct analogy from the example to problem. The other is to study the example thoroughly, generate procedural steps from the example, and then apply the generated rule to the problem. Only the second approach will produce a general competence for solving a problem that does not depend on always having an example to

map from. We expect that if the example is similar to the problem, comparing an example with a problem will tend to promote the first, inappropriate method of solution.

Figure 4 shows two possibilities of example-problem pairs depending on example similarity. In the similar condition (Fig. 4a), the provided example and problem are identical in diagram structure, operators, and all numbers except only one value, which results in a change in only one of the final answers. When an example is perceptually similar to the current problem, it is easy to align commonalities between the cases, and participants will be more likely to use the relation between example and current problem to find the solution. By comparing what is the same and what is different between the example and problem, they can generate the answer to the current problem by adjusting the answer of the example. For instance, in this example, participants could generate the answer by copying the common value and focusing on how the one difference changed the other value. Here, the only difference

(a) Similar example-problem pair



(b) Dissimilar example-problem pair

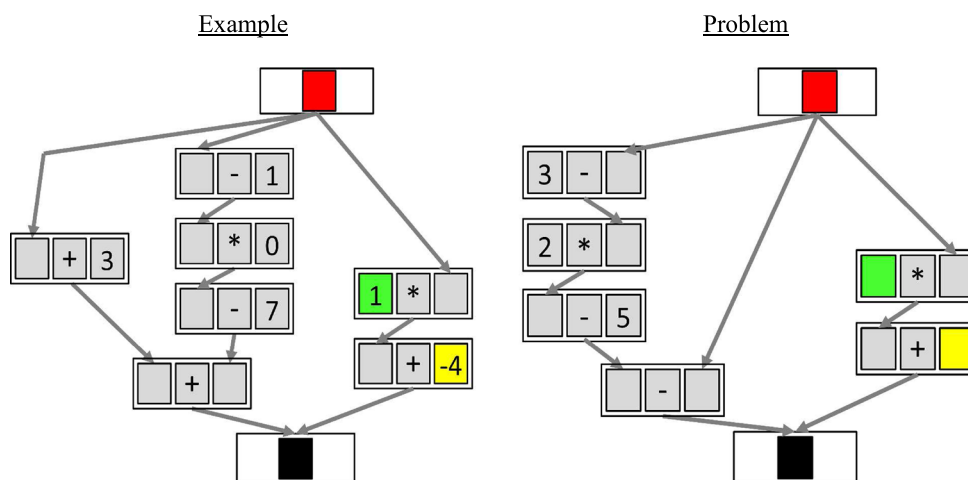


Fig. 4 Two possibilities of example-problem pairs depending on example similarity. **a** Similar example-problem pair. **b** Dissimilar example-problem pair

is on the left-side path between the 5 in the example and the 6 in the problem. This number affects the change in the constant term (from 44 to 45) in the current problem while keeping the coefficient term ($14x$) unchanged. In this example, problems can be solved using a superficial analogy between the example and problem without understanding the principles. Solely based on the changed values (from 5 to 6), participants might induce the relation “increase by 1” and then apply this relation to the number 44. Increasing the number 44 by 1 generates correct answer, 45, to the current problem. This scenario illustrates how participants can rely on superficial analogy without understanding the underlying rules. In such scenario, comparison between simultaneously presented cases does not help learning, rather it harms learning. Superficial analogy to example from problem can be a characteristic of poor learning. For example, Chi and her colleagues (Chi, Bassok, Lewis, Reimann, & Glaser, 1989) examined how good and poor students utilize examples while solving a problem and identified different patterns of referring to examples. Poor students tended to refer an example more often than good learners, and this was because they used the example in this superficial way to essentially “copy” the solution.

In contrast, in the dissimilar condition (see Fig. 4b), this kind of superficial analogy is not possible. The example has a different structure and different answers from the given problem. Because the provided example is very dissimilar to the problem to solve, participants will have a hard time relating the example to problem. In this case, it would be easier to first study the example thoroughly and come up with a solution procedure rather than making a direct mapping from example to problem.

Method

Participants

Sixty undergraduate and graduate students (31 male and 29 female, $M = 22$ years, $SD = 4.1$ years) at Carnegie Mellon University participated in the study. Participants were randomly assigned to one of the four experimental conditions (15/condition). Participants received \$10/hour plus a performance-based bonus.

Design, materials, and procedures

A 2×2 between-subjects design was employed. The study crossed example format and example similarity. The first contrast compares annotated examples with nonannotated examples. In the annotated example condition, the example indicated the algebraic operations that were required to perform the collection of like terms and final answers to the problem. Figure 5 shows an example used in the annotated example condition. Algebraic expressions were directly drawn on top of the data-flow diagram, and special colors were used for the coefficient and constant terms. The coefficient term was colored green on the left path of the

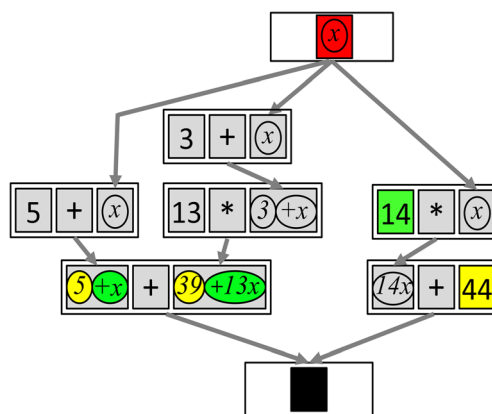


Fig. 5 An example of annotated examples used in the learning session of Experiments 1 and 2. Algebraic expressions are directly drawn on top of the diagrams with a special color-coding. Coefficient term is colored green and constant term is colored yellow

diagram, and this matching color was used on the right for the coefficient term. Likewise, the constant term was colored yellow on the left and the right. Below the example, a text message appeared, such as, “The correct answer to the green tile is 14 and the answer to the yellow tile is 44. The bubbles show the solutions of this problem.” The expression bubbles used in examples were never used in the problem, and thus participants solved identical problems in all experimental conditions. In the nonannotated example condition, the example simply provided final answers to the problem without algebraic expressions (like the examples shown in Fig. 4). The right-path diagram had green and yellow coloring for coefficient and constant term, as in the annotated example condition, but this color coding did not have a special meaning because algebraic expressions were not provided in this condition. The same text message appeared as in the annotated condition, minus the sentence mentioning bubbles. In neither case was any explicit instruction given as to how to calculate the answer, but the annotated example revealed the algebraic structure of the problem and solution steps in calculating the answer.

Example similarity was crossed with example format. As described earlier, Fig. 4 illustrates the two possibilities for the example-problem pair, depending on similarity. In the similar condition, example and problem were identical in diagram structure, operators, and all numbers except for one value or operator, which resulted in a change in only one of the final answers. In the dissimilar condition, examples always had a different structure and different answers from the given problem. The same set of examples was used in both similar and dissimilar conditions, but in different order. By simply changing the presentation order of examples while holding the order of problems constant, similarity was manipulated between the example and problem.

The study consisted of two sessions, a learning session and a transfer session. In the learning session, there was a total of 30 problems and 30 examples paired to each problem based on experimental conditions. All problems were identical across the conditions, and experimental manipulations occurred only in the

provided example (in terms of both example format and similarity). The problems were divided into three separate blocks (10 problems per block) to allow participants to have a break between blocks if needed. The three blocks were identical in terms of problem difficulty. Each problem had four or five boxes to combine in collecting like terms. There were three buttons below the problem, “Show example,” “Hint,” and “Next problem.” For the first problem, the example was presented from the beginning, but for the rest of problems the example was only presented on request or when a time limit was reached. Participants were not allowed to move on to the next problem until they solved the problem correctly or the maximum time limit was reached. If participants could not complete the step within 1 minute,¹ an example for the current problem would automatically appear. After the next time limit (1 min) was reached, the text hint was automatically shown. The text hint was simply about general interface use such as “Move the mouse to the green tile and enter a number,” and it was identical across all experimental conditions. Then each of the empty tiles was automatically filled by the system following the next time limit (1 min) consecutively. This time limit was intended to prevent excessive floundering. Also, immediate feedback was provided whenever participants entered an incorrect value. To prevent overuse of examples and hints, participants were instructed that use of those buttons would deduct 4 cents each from the bonus they would receive.

A transfer session immediately followed the learning session. The transfer task was designed to measure whether participants could apply what they learned to a new set of problems by adapting the learned procedure, separately from how well they could remember their learning. The transfer session was thus conducted on the same day as the learning session. The transfer task was identical across all experimental conditions, and the experimental manipulations occurred only during the learning session. There were neither examples nor hints provided in any experimental conditions. In the transfer session, there was a total of 40 problems, and the problems were divided into five blocks (eight problems per block). The first block consisted of eight warm-up problems, and the remaining four blocks consisted of 32 transfer problems. The warm-up problems were similar to the problems used in the learning session. The one change was that the warm-up problems showed a final value at the bottom, whereas the bottom box was empty during learning. These warm-up problems were included to help participants become familiarized to the new format of upcoming transfer problems.

The transfer problems presented students with problems that they could solve only if they understood the solutions they were producing in learning. Four different subtypes of transfer problems were constructed by changing the structure of the diagram.

To correctly solve the problem, participants had to perform collection of like terms and then adjust the values according to the changed feature of the problem. The subtypes involved sign change, factoring, operator change, and division, respectively. Figure 6 shows one example of transfer problems, and some other examples are included in the Appendix. Changed parts of the diagram were pink colored so that participants could easily notice the changed parts of the diagram. Participants were given 1 minute for each transfer problem. They were asked to enter two numbers for each problem and click a “Done” button. The response was followed by an immediate feedback page. The feedback page showed one of the three messages, “Correct,” “Incorrect,” or “Time’s up,” on the top of the screen. In cases of incorrect and timed-out responses, the correct answers were presented next to the participant’s own answers simultaneously so that they could compare them. The feedback was presented for 2 seconds for correct responses and 10 seconds for incorrect and timed-out responses. The entire experiment took between 1.5 and 2 hours.

Results and discussion

Figure 7 presents the proportion of perfectly solved problems (i.e., both numbers correctly provided with no errors, hints, or examples) in the two phases of the experiment. Two-by-two between-subjects analyses of variances (ANOVAs) were conducted to determine the effect of example format and example similarity on learning and transfer separately. The results from the warm-up problems did not add anything over the initial learning; thus, we focused on the performance of the learning and transfer problems. The left panels in Fig. 7 show the mean percentages of correctly solved problems out of the 30 problems in the learning session. There was a significant main effect of example format, $F(1, 56) = 13.47, p = .001, \eta_p^2 = .194$. Regardless of example similarity, participants who were given annotated examples with an algebraic expression ($M = 58.56, SD = 27.23$) correctly solved problems significantly more often than those who were given

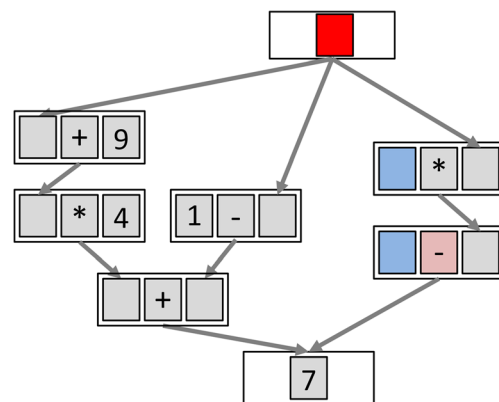


Fig. 6 An example of transfer problems used in Experiments 1 and 2. This example is one of the four subtypes “sign change” problem, $(x + 9) * 4 + (1 - x) = 3x + 37 = -(3x) + 37$, thus, correct answers are -3 and 37. Other examples of transfer problems are shown in the Appendix

¹ The 1-minute duration for each step was determined based on average time participants needed for solving a problem obtained from some of our earlier studies (Lee et al., 2011; Lee et al., 2014).

nonannotated examples ($M = 33.44$, $SD = 30.56$). This in itself is clear evidence that participants were using the examples to guide their learning, but of more interest is the highly significant effect of the example similarity, $F(1, 56) = 13.00$, $p = .001$, $\eta_p^2 = .188$. As predicted, participants who studied examples that were dissimilar to problems ($M = 58.33$, $SD = 31.36$) performed significantly better than those who studied examples similar to the problems ($M = 33.67$, $SD = 32.00$). There was no interaction effect of example format by similarity, $F < 1$.

In the transfer session, significant main effects were obtained of both example format, $F(1, 56) = 17.33$, $p < .001$, $\eta_p^2 = .236$, and similarity, $F(1, 56) = 5.70$, $p = .02$, $\eta_p^2 = .092$. More interestingly, there was a significant interaction between example format and similarity, $F(1, 56) = 6.61$, $p = .01$, $\eta_p^2 = .106$. The right panels of Fig. 7 show the mean percentages of correctly solved problems out of the 32 transfer problems. Among the group of participants who were given annotated examples in the learning session, those who studied dissimilar examples ($M = 66.46$, $SD = 13.39$) showed a better transfer performance than those who studied similar examples ($M = 32.71$, $SD = 31.34$), $t(28) = 3.84$, $p = .001$, $d = 1.40$. However, among the group of participants who were provided with non-annotated examples, there was no performance difference between the similar ($M = 21.88$, $SD = 30.48$) and dissimilar ($M = 20.63$, $SD = 26.25$) conditions, $t < 1$.

To summarize, the effect of example format was evident in both the learning and transfer tasks. Regardless of example similarity, annotated examples were more effective for both learning and transfer than nonannotated examples. When problem structure was clearly revealed by algebraic expressions as in annotated examples, participants were able to understand examples better and, in turn, to solve problems better. However, when only final answers were provided in an example, participants had a hard time generating correct solution procedures for themselves from the given example. This replicates our previous findings (Lee et al., 2011, 2014) in that it is critical to reveal any hidden structure in the problem so that learners pay attention to relevant features of the problem.

Regarding the effect of example similarity, dissimilar examples resulted in better learning than similar examples. It appeared that simultaneous presentation of an example similar to the problem promoted superficial mapping, preventing the generation of a more general understanding. This interpretation is consistent with the reports collected in the debriefing session. Many participants from the similar condition reported that they found that one of the answers were always the same as the example answers, and the other could be increased or decreased by the changed amount. To use this superficial strategy, they had to keep asking for examples throughout the learning session. As a matter of fact, examples were more often used in the similar condition (16.2 times) than in the dissimilar condition (8.3 times) while solving a total of 30 problems in the learning session, $F(1, 56) = 13.90$, $p < .001$, $\eta_p^2 = .199$. Because of the heavy reliance on examples during the learning session, they had difficulty in the transfer session where there were no examples from which to make a superficial analogy. In contrast, in the dissimilar condition, examples, and problems looked so different from each other that participants did not try to engage in a superficial mapping. Instead, they appeared to try to understand the relations in the example and then apply their understanding to the problem. Therefore, comparing similar example-problem pairs harmed learning by promoting a superficial mapping strategy, whereas comparing dissimilar example-problem pairs helped learning by promoting a relational mapping strategy.

One interesting finding from this study was the interaction effect between example format and example similarity found in the transfer test. Dissimilar examples were superior regardless of example format during the learning session. However, this effect disappeared in the nonannotated example condition while it remained in the annotated example condition during the transfer session. This seemed to be because participants in the nonannotated/dissimilar condition sometimes acquired solution procedures that only worked for the learning session problems. In fact, in the debriefing session, when participants were asked to describe the solution procedure they used

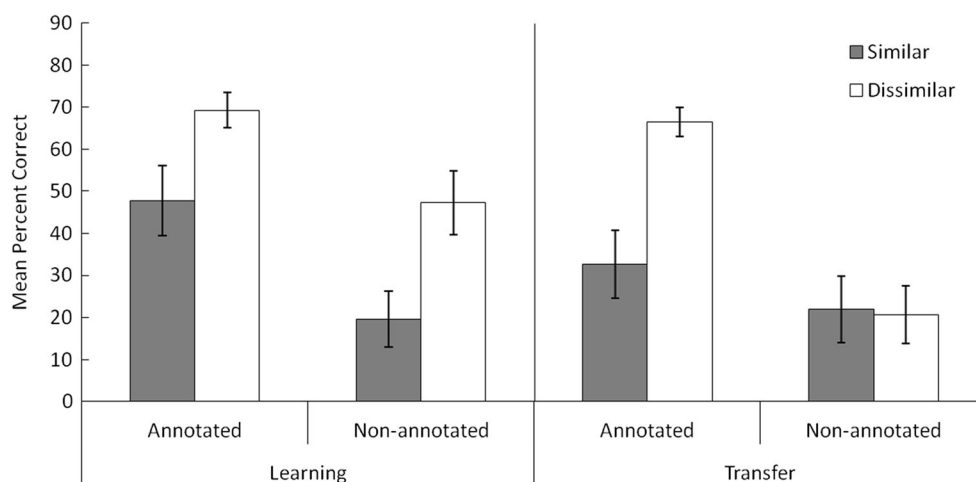


Fig. 7 Mean percentages of correctly solved problems without errors in each session of Experiment 1. Error bars represent 1 standard error of mean

during the learning session, many of them reported incorrect solution steps. Surprisingly, 5 of the 15 participants in this particular condition reported the same but incorrect procedures. Those procedures were plugging in 0 on the top box and propagating it down to get the value of the constant term, and then plugging in 1 on the top box and propagating it down to get the value of the coefficient term. By assuming the top unknown number is 1, they could come up with a hypothetical resulting number, and in turn constrain the value of the coefficient term. This algorithm works perfectly for the problems in the learning session, but becomes broken in the transfer session. Participants who saw annotated examples would realize that this strategy would not work on the transfer problems without some modification of the result to accommodate the different structure of the transfer problems.

Experiment 2

In Experiment 1, we attributed the disadvantage of example similarity to a superficial mapping strategy that was available only in the similar conditions. Similarity between example and problem allowed participants to directly map the solution from the example to the problem. Simultaneous presentation of an example and problem seemed to work as a spatial cue that evokes comparison (Richland, Zur, & Holyoak, 2007). In Experiment 2, we examined effect of example similarity after removing such a spatial cue. Instead of presenting an example and a problem on the same page, an example was alternated with a problem in a sequential presentation. Without simultaneous presentation, learners may have difficulty in mapping between corresponding elements of the example and problem because this would require holding the complex examples in working memory. Such sequential presentation of alternating examples and problems has frequently been used in past research on examples (e.g., Paas, 1992; Paas & Van Merriënboer, 1994; Trafton & Reiser, 1993). By testing the same design in a sequential presentation, we will be able to identify how example similarity differs based on learning context (simultaneous vs. sequential presentation). We expected that presenting the example and problem on a separate page would abolish the similarity effect of Experiment 1 by preventing participants from making a comparison, especially because the examples were complex and hard to remember in our current domain.

Method

Participants

Sixty undergraduate and graduate students (31 male and 29 female, $M = 21$ years, $SD = 2.3$ years) at Carnegie Mellon University participated in the study. Participants were randomly

assigned to one of the four experimental conditions (15/condition). Participants received \$10/hour plus a performance-based bonus.

Design, materials, and procedures

Design, materials, and procedures were identical to those used in Experiment 1, except for one change. In Experiment 2, instead of presenting an example and a problem on the same page, examples alternated with problems in a sequential manner. Participants were given up to 1 minute to study a provided example. By clicking the next button below the example, participants were able to move on to the subsequent problem. Depending on the experimental conditions, participants were provided with either annotated examples with algebraic expressions or nonannotated examples without algebraic expressions. Thus, participants saw all of the examples in this experiment but could choose whether to study them. Crossed with the example format, example similarity was manipulated in the same manner as in Experiment 1. The identical transfer problems were used as in Experiment 1.

Results and discussion

Figure 8 shows the mean percentages of correctly solved problems in the learning session and the transfer test. In the learning session, there was a significant main effect of example format, $F(1, 56) = 12.57, p = .001, \eta_p^2 = .183$. Regardless of similarity condition, participants who were provided with annotated examples ($M = 66.00, SD = 23.53$) solved significantly more problems correctly than those who were provided with nonannotated examples ($M = 40.22, SD = 31.78$). As in the first experiment, this clearly establishes that participants were using the examples to help their learning. However, as predicted, there was no overall effect of similarity in this experiment, nor was there an interaction between example format and similarity type, $F_s < 1$. In the transfer test, the overall effect of example format was only marginally significant, $F(1, 56) = 3.95, p = .052, \eta_p^2 = .066$, although the same directionality was again found. There was no overall effect of similarity, nor was there an interaction between example format and similarity type, $F_s < 1$.

To summarize, as in Experiment 1, participants showed better learning and transfer given annotated examples that clearly revealed the structure of the problem. However, unlike the findings from Experiment 1, there was no effect of example similarity. When cases were not presented simultaneously, participants could not compare them and thus were not affected by example similarity. The only way to learn from the example was to find a mapping from the left-hand side to the right-hand side within the example. Other studies also have found that sequential presentation of the cases led to different training effects from the simultaneous presentation of the identical cases (e.g., Loewenstein et al., 1999; Rittle-Johnson & Star, 2007).

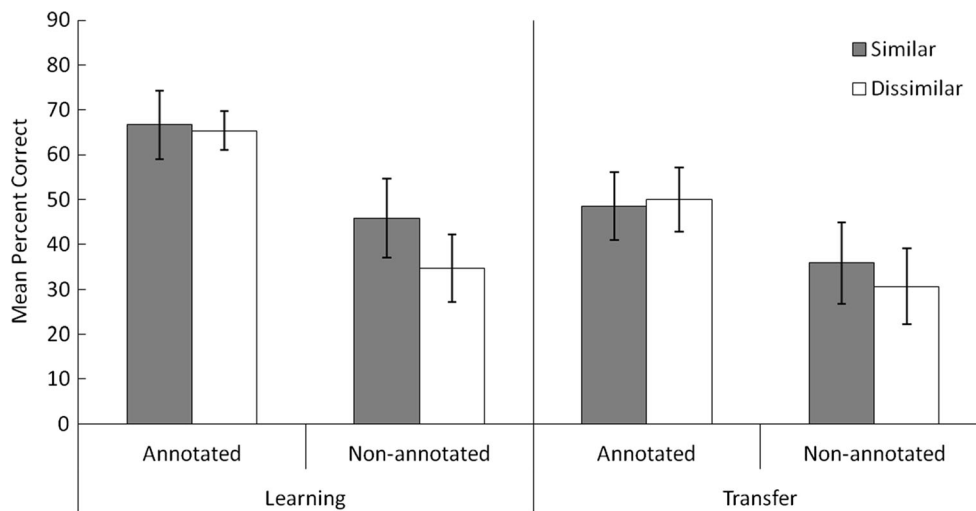


Fig. 8 Mean percentages of correctly solved problems without errors in each session of Experiment 2. Error bars represent 1 standard error of mean

General discussions

Summary

Two experiments have replicated our previous findings (Lee et al., 2011, 2014) regarding the effect of example format such that appreciation of problem structure is critical for learning and transfer. When participants were provided with an annotated example with algebraic expressions, they were better able to understand the problem solution procedures and apply them to solve new cases of problems. However, when participants were given an example with final answers filled without algebraic expressions, they had difficulties generating solution procedures for themselves. In Experiment 1, some participants from the nonannotated example condition even induced an incorrect solution step that worked for learning problems but did not work for any type of transfer problems.

Regarding the effect of example similarity, different patterns of results were obtained as predicted in the two experiments. In Experiment 1 with simultaneous presentation, example similarity had negative effects on learning and transfer. In the similar condition, participants could use superficial mappings that prevented learning of the desired problem solving rules. In Experiment 2 with sequential presentation, it was no longer possible² to use such a superficial strategy and the results showed no effect of similarity in any of the sessions.

This research demonstrates that when learners can directly compare examples and problems, similarity between example and problem can result in superficial mapping between them and poor learning. This finding is consistent with many previous studies that showed harmful effects of similarity (e.g., Ben-Zeev & Star, 2001; Gick & Holyoak, 1983; Merrill & Tennyson, 1978; Quilicci & Mayer, 1996; Reed, 1987, 1989;

Ross, 1987; Tennyson, 1973) in that similarity seemed to draw attention toward irrelevant superficial detail rather than underlying deep relations, thus inhibiting learning.

In order for comparison to be effective, example and problem had to be different so that participants could not rely on superficial mapping based on perceptual similarities between the cases. This finding has educationally important implication regarding the effect of comparison. Although comparing multiple cases is considered as a high-quality instructional method (NCTM, 1991, 2000), comparison does not always guarantee better learning. To promote learning, what matters is not simply whether students compare example and problem but how they compare them. Several studies suggest that effects of comparison can depend on various factors, including a learner's prior knowledge (Rittle-Johnson, Star, & Durkin, 2009), experience of pretraining (Braithwaite & Goldstone, 2014), which aspects of problem being compared (Rittle-Johnson & Star, 2009), and other task specifics, such as feature variation and category structure (Ankowski et al., 2013).

Aggregate data analysis

To examine how effect of example format and example similarity differs depending on learning context, we put learning and transfer data of Experiment 1 and Experiment 2 into a single design and performed $2 \times 2 \times 2$ between-subjects ANOVAs. Three between-subjects variables were example format (annotated vs. nonannotated), example similarity (similar vs. dissimilar), and learning context (simultaneous vs. sequential). Although manipulation of learning context occurred across the two experiments, two CMU populations were comparable in terms of mean age, $t(118) = 1.81, p = .073, d = 0.33$, and average SAT scores, $t(103) = 1.5, p = .137, d = 0.28$. The analysis of learning data showed that there was a significant effect of example format such that participants who studied with annotated examples showed better learning than those who studied with nonannotated examples,

² Or at least not easy, because the superficial strategy would require remembering a complex example.

$F(1, 112) = 25.99, p < .0001, \eta_p^2 = .188$. However, there was no overall effect of similarity, $F(1, 112) = 3.41, p = .067, \eta_p^2 = .030$, nor was there an effect of learning context, $F(1, 112) = 2.03, p = .157, \eta_p^2 = .018$. The only significant interaction effect was similarity by learning context, $F(1, 112) = 9.57, p = .002, \eta_p^2 = .079$, suggesting that effect of example similarity differed depending on whether an example and a problem were presented simultaneously (Experiment 1) or sequentially (Experiment 2). In Experiment 1, similarity harmed learning by allowing participants to use a superficial mapping, but in Experiment 2 such effect disappeared because participants were not really able to do a comparison.

Regarding the transfer performance, similar patterns of results were obtained as in learning session. Again, there was a significant effect of example format, $F(1, 112) = 17.67, p < .0001, \eta_p^2 = .136$, but no reliable effects of either similarity, $F(1, 112) = 0.02, p = .088, \eta_p^2 < .001$, or learning context, $F(1, 112) = 1.22, p = .272, \eta_p^2 = .011$. The interaction between similarity and learning context was only marginally significant, $F(1, 112) = 2.95, p = .088, \eta_p^2 = .026$. There were not any other two- or three-way interaction effects.

A comparison of the transfer results between the two experiments is suggestive: Participants performed non-significantly better in Experiment 2 in all conditions but the dissimilar annotated condition, where they were significantly better in Experiment 1, $t(28) = 2.09, p = .046, d = 0.76$. In three of the other conditions, sequential presentation appeared to be better, if anything, because it prevented superficial mapping. However, in the dissimilar annotated condition, participants were guided to the correct understanding of the example. Perhaps the better performance of dissimilar annotated participants in Experiment 1 (than in Experiment 2) reflects the fact that they did not have to remember their understanding of the example but could immediately see if it would apply to the problem due to simultaneous presentation. Performance in the dissimilar annotated condition in Experiment 1 was more than 15 % better than the transfer performance in any other conditions. We think this presents a good condition for learning from examples because the annotation highlights the critical features of the examples, the simultaneous presentation allows for comparisons that do not depend on memory, and the dissimilarity prevents any superficial mapping from example to test problem.

In addition to resulting in the best transfer performance, participants in the dissimilar condition of Experiment 1 showed the shortest learning time. Across experiments, participants were slower in the nonannotated conditions, an average of 41 vs. 26 minutes in the learning phase, $F(1, 112) = 23.81, p < .0001, \eta_p^2 = .175$. Although there was no significant difference between the two experiments in the dissimilar conditions, participants had somewhat

shorter learning times in Experiment 1 (22 mins) than Experiment 2 (23 mins). Taken together, all of our results are consistent with the conclusion that the optimal condition is one in which examples are simultaneously presented, annotated to facilitate their interpretation, and different from the problem. This combination appears to encourage and facilitate extracting the critical relations in the example.

One may imagine that given the lack of similarity between the example and problem (as in dissimilar conditions of Experiment 1), students just would not bother comparing them and simply treat each example as unrelated to the particular problem it is presented with. However, the best performance obtained from the annotated dissimilar condition in Experiment 1 strongly suggests that these participants were making the alignment between the example and problem.

Limitations and future directions

Our study should be distinguished from previous studies in several aspects. First, in the current study, we used sequential or simultaneous presentation to encourage or discourage participants to compare between cases and in turn use the mapping between compared cases for learning. However, participants were never explicitly asked to compare cases in any of the experiments. In contrast, several previous studies emphasized the importance of direct instruction of comparison to benefit from comparing multiple cases. Participants were explicitly told to think about similarities between cases (e.g., Gick & Holyoak, 1983; Gentner et al., 2003; Loewenstein et al., 1999), describe similarities and differences (Rittle-Johnson & Star, 2009), or answer a question that directly refers two cases at the same time (Rittle-Johnson & Star, 2007). Different from these previous studies, our study, however, suggests different presentation structure itself may encourage or discourage spontaneous comparison.

Second, many previous studies distinguished surface and structural similarity (e.g., Holyoak & Koh, 1987; Reed, 1989; Ross, 1989a, 1997; Ross & Kennedy, 1990; VanderStoep & Seifert, 1993) and current study did not separate such different types of similarity. We used overall perceptual similarity to encourage participants to do comparison. In future studies, it needs to investigate how different types of similarity can affect comparison activity and in turn mapping strategy involved in problem solving.

Conclusion

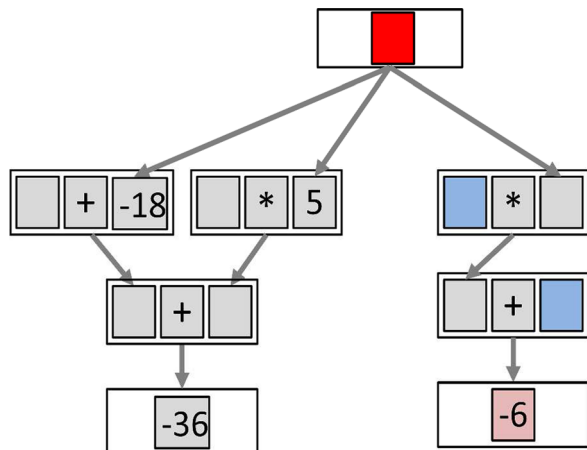
A large literature documented benefits of simultaneous presentation of multiple examples and comparison on learning. However, comparison does not always lead to better learning.

Rather, its effect appeared to depend on whether comparison actually promotes relevant mapping that is required to achieve a desired learning goal. Also, such comparison seemed to be affected by both similarity of examples and how they are presented. In order to make comparison effective, instructors need to consider what cognitive process is elicited by the learning context imposed on students.

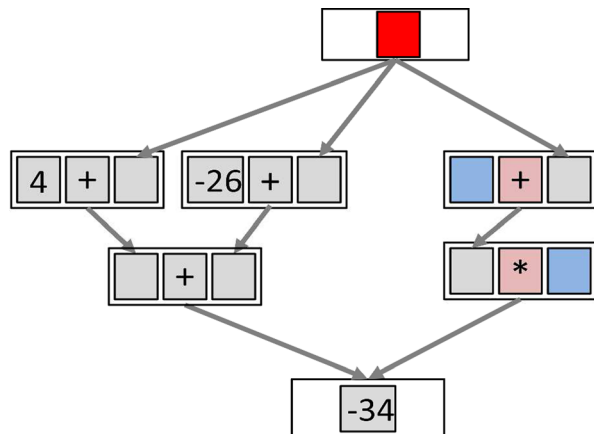
Appendix

Examples of transfer problems. There were four different types of transfer problems, and they included sign change, factoring, operator change, and division. The first type (sign change) is shown in Fig. 6, and the other three types are shown below.

(a) Factoring problem: $(x + -18) + (x * 5) = -36 \rightarrow 6x - 18 = -36 \rightarrow 1x - 3 = -6$, thus, correct answers are 1 and -3

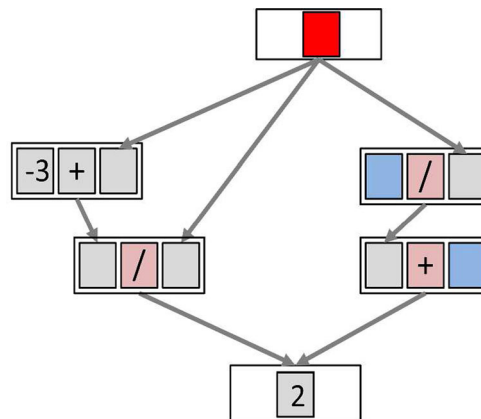


(b) Operator change problem: $(4 + x) + (-26 + x) = 2x - 22 = (x - 11) * 2$, thus, correct answers are -11 and 2



Acknowledgments The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A100109 to Carnegie Mellon University and by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014S1A5A8018026). We thank Abe Anderson and Aryn Pyke for valuable comments on the paper. Correspondence may be directed to Hee Seung Lee, Department of Education, Yonsei University, Seoul, Korea (hslee00@yonsei.ac.kr).

(c) Division problem: $(-3 + x) / x = -3/x + 1$, thus, correct answers are -3 and 1.



References

Anderson, J. R., Lee, H. S., & Fincham, J. (2014). Discovering the structure of mathematical problem solving. *NeuroImage*, *97*, 163–177.

Ankowski, A. A., Vlach, H. A., & Sandhofer, C. M. (2013). Comparison vs. contrast: Task specifics affect category acquisition. *Infant and Child Development*, *22*, 1–23.

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal*, *93*, 373–397.

Ben-Zeev, T., & Star, J. (2001). Spurious correlations in mathematical thinking. *Cognition and Instruction*, *19*, 253–275.

Braithwaite, D. W., & Goldstone, R. L. (2014). Benefits of variation increase with preparation. *Proceedings of the Thirtieth-Sixth Annual Conference of the Cognitive Science Society* (pp. 1940–1945). Quebec City, Canada: Cognitive Science Society.

Brunstein, A., Betts, S., & Anderson, J. R. (2009). Practice enables successful learning under minimal guidance. *Journal of Educational Psychology*, *101*, 790–802.

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.

Fraivillig, J. L., Murphy, L. A., & Fuson, K. (1999). Advancing children’s mathematical thinking in everyday mathematics classrooms. *Journal for Research in Mathematics Education*, *30*, 148–170.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170.

- Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development, 8*, 285–307.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*, 393–405.
- Gentner, D., & Markman, A. B. (1993). Analogy—Watershed or Waterloo? Structural alignment and the development of connectionist models of cognition. In S. J. Hanson, J. D. Cowan, & C. L. Giles (Eds.), *Advances in neural information processing systems, 5* (pp. 855–862). San Mateo, CA: Kaufmann.
- Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*(1), 45–56.
- Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*, 487–513.
- Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on thought and language: Interrelations in development* (pp. 225–277). London: Cambridge University Press.
- Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277–300.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38.
- Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 3–28.
- Goldstone, R. L., & Medin, D. L. (1994). Similarity, interactive activation and mapping. In K. J. Holyoak & J. A. Barnden (Eds.), *Advances in connectionist and neural computation theory: Vol. 2. Analogical connections* (pp. 321–362). Norwood, NJ: Ablex.
- Holyoak, K. J., Junn, E. N., & Billman, D. (1984). Development of analogical problem solving skill. *Child Development, 55*, 2042–2055.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15*, 332–340.
- Huffed-Ackles, K., Fuson, K., & Sherin, M. G. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education, 35*, 81–116.
- Kotovskiy, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*, 2797–2822.
- Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*, 29–63.
- Lee, H. S., Anderson, A., Betts, S., & Anderson, J. R. (2011). When does provision of instruction promote learning? In L. Carlson, C. Hoelscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3518–3523). Austin, TX: Cognitive Science Society.
- Lee, H. S., Fincham, J., Betts, S., & Anderson, J. R. (2014). An fMRI investigation of instructional guidance in mathematical problem solving. *Trends in Neuroscience and Education, 3*, 50–62.
- Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review, 6*, 586–597.
- Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 23*, 431–467.
- Merrill, M. D., & Tennyson, R. D. (1978). Concept classification and classification errors as a function of relationships between examples and nonexamples. *Improving Human Performance Quarterly, 7*, 351–364.
- Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General, 131*, 5–15.
- National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 510–520.
- Oakes, L. M., & Ribar, R. J. (2005). A comparison of infants' categorization in paired and successive presentation familiarization tasks. *Infancy, 7*, 85–98.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*, 429–434.
- Paas, F., & Van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*, 122–133.
- Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology, 88*, 144–161.
- Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence, individual form*. London: Lewis.
- Reed, S. K. (1987). A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(1), 124.
- Reed, S. K. (1989). Constraints on the abstraction of solutions. *Journal of Educational Psychology, 81*, 532–540.
- Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23*, 90–108.
- Richland, L. E., Holyoak, K. J., & Stigler, J. W. (2004). Analogy use in eighth-grade mathematics classrooms. *Cognition and Instruction, 22*, 37–60.
- Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science, 316*, 1128–1129.
- Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*, 561–574.
- Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology, 101*, 529–544.
- Rittle-Johnson, B., Star, J. R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: Influence on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology, 101*, 836–852.
- 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*(4), 629.
- Ross, B. H. (1989a). 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.
- Ross, B. H. (1989b). Reminders in learning and instruction. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 438–469). Cambridge: Cambridge University Press.
- Ross, B. H. (1997). The use of categories affects classification. *Journal of Memory and Language, 37*, 240–267.
- Ross, B. H., & Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 42–55.
- Silver, E. A., Ghousseini, H., Gosen, D., Charalambous, C., & Strawhun, B. (2005). Moving from rhetoric to praxis: Issues faced by teachers in having students consider multiple solutions for problems in the mathematics classroom. *The Journal of Mathematical Behavior, 24*, 287–301.

- Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.
- Tennyson, R. D. (1973). Effect of negative instances in concept acquisition using a verbal learning task. *Journal of Educational Psychology*, *64*, 247–260.
- Trafton, J. G., & Reiser, B. J. (1993). The contributions of studying examples and solving problems to skill acquisition. In M. Polson (Ed.), *Proceedings of the Fifteenth annual Conference of the Cognitive Science Society* (pp. 1017–1022). Hillsdale, NJ: Erlbaum.
- VanderStoep, S. W., & Seifert, C. M. (1993). Learning “how” versus learning “when”: Improving transfer of problem-solving principles. *The Journal of the Learning Sciences*, *3*, 93–111.