

# Learning From Mistakes: Inducing Analogous Solution Failures to a Source Problem Produces Later Successes in Analogical Transfer

Mary L. Gick and Susan J. McGarry  
Carleton University, Ottawa, Ontario, Canada

In 4 experiments we investigated the use of analogical problem solving to facilitate the solution to a problem that usually results in persistent solution failures without hints. The results of Experiments 1-3 indicate that spontaneous transfer was facilitated by a manipulation of the surface form of the source problem that tends to induce initial solution failures analogous to those produced to the target problem. Surface similarity of content words and diagrams had no effect on transfer in Experiments 1-3. In Experiment 4, facilitation of spontaneous transfer was not obtained when source solution failures were prevented. The importance of failed solutions in problem representation and the relationship between problem representation and surface similarity are discussed.

The use of analogy was recommended by the mathematician Polya (1957) as a useful heuristic for solving a novel problem. In the past decade, research in cognitive psychology has proposed processes involved in analogical problem solving and examined conditions under which analogy will be used. A critical first process that has been proposed in most theories of analogical transfer is *noticing* (e.g., Gentner, 1989; Holyoak & Thagard, 1989; Novick, 1988a, 1988b; Ross, 1987, 1989); the access or retrieval of the source problem and its solution during solution of a subsequent target problem. *Spontaneous transfer*, that occurring in the absence of an informing hint to use the source problem while solving the target problem, has been used as a measure of noticing (e.g., Gick & Holyoak, 1980, 1983).

Once the source problem has been noticed, the corresponding elements between source and target problems can be mapped together, and the solution of the source problem can be transformed and applied to fit the target problem. Learning or generalization may take place as a result of transferring the source solution to the target problem (e.g., Gick & Holyoak, 1983; Novick & Holyoak, 1991; Ross & Kennedy, 1990).

Theories of analogy have emphasized that analogical processes are affected by the similarity of different kinds of features (and their relations) between source and target problems.

Problem features thought to be important include surface (not related to the solution) versus structural (solution-related) features (Gentner, 1989; Holyoak, 1985; Ross, 1989); pragmatic or goal-related features (Holyoak, 1985; Holyoak & Thagard, 1989, 1990); and systematicity of features, determined by whether they figure in higher order relations (Gentner, 1983, 1989; Reed, 1987).

The importance of similar problem features is supported by recent research. For example, surface features (e.g., lasers and rays) influence noticing processes in transfer between Duncker's (1945) radiation problem and its isomorphs (Holyoak & Koh, 1987; the problem is to destroy an inoperable stomach tumor with rays without harming surrounding healthy tissue). Surface features (e.g., both problems are about pizza or both problems have people arranged in rows) also facilitate transfer between probability problems (Ross, 1984) and word problems (Novick, 1988a; Reed, 1987). Structural features (e.g., "glass filament might break" as compared to "healthy tissue surrounding tumor might be harmed") facilitate both noticing and application processes between isomorphs of the radiation problem (Holyoak & Koh, 1987). Recent research has focused on further distinguishing different types of similarity and determining their relative importance for the different processes involved in analogical transfer (Ross, 1989).

---

The research reported in this article was supported by Grant A1212 to Mary L. Gick from the Natural Sciences and Engineering Research Council of Canada. Experiment 1 was part of an honor's thesis and Experiment 3 part of a master's thesis, conducted by Janet Kanwisher and Susan J. McGarry, respectively, under Mary L. Gick's supervision.

We thank Michael deJong and Nausheen Waheed for data scoring and Bruce Cowin for data scoring, testing subjects, and other technical help. We are grateful to Keith Holyoak, Jo-Anne LeFevre, Bob Lockhart, and Brian Ross for helpful discussions of this research. We thank Rick Grant, Keith Holyoak, Jo-Anne LeFevre, Mark McDaniel, and two anonymous reviewers for valuable comments on a draft of this article.

Correspondence concerning this article should be addressed to Mary L. Gick, Department of Psychology, Carleton University, Ottawa, Ontario, Canada K1S 5B6.

## Analogical Transfer and Solution Failures

Although theories of analogy have emphasized the importance of the similarity of features of source and target *problems*, they have not addressed the similarity of features of source and target (failed) *solution attempts* as a potential means of facilitating transfer. Perhaps this is partly because the target problems used in some studies of analogical problem solving do not have only one solution procedure (e.g., Bassok & Holyoak, 1989; Novick, 1988a), and thus there may be no failure to solve the target problem even if transfer does not occur. For example, if subjects fail to transfer the efficient lowest common multiple procedure from a source problem, the target word problem can be solved using other methods (e.g., examining all multiples of the divisors in the problem;

Novick, 1988a). Similarly, research on Duncker's (1945) radiation problem has focused on transfer of the convergence solution (i.e., dividing the large ray into multiple small rays that converge simultaneously on the stomach tumor from different directions; e.g., Catrambone & Holyoak, 1989; Gick & Holyoak, 1980, 1983; Holyoak & Koh, 1987; Spencer & Weisberg, 1986), which Duncker considered to be the best solution among others (e.g., sending the rays down the esophagus or using a lead shield; Duncker, 1945; Gick & Holyoak, 1980).

In the present study, we examined the importance of source and target failed solutions in analogical transfer when the target problem led to frequent and persistent failure because of the use of inappropriate problem representations. (We use the term *problem representation* to refer to the problem solver's understanding of the problem situation and operators associated with it [Simon, 1978; Simon & Hayes, 1976].) We compared transfer from analogous problems initially solved (i.e., before correct feedback was given) using the correct representation with analogous problems initially attempted using inappropriate representations that failed to lead to the correct solution but were analogous to those that failed to solve the target problem. The hypothesis was that spontaneous transfer would be more effective from analogous problems initially solved incorrectly, because the representations of source and target problems would match more than when the source problem was solved correctly. Increased similarity in representation between source and target problems was expected to facilitate retrieval of the source problem during solving of the target problem (Novick, 1988a).

Intuitively, the hypothesis that target solution failures are used in searching for related problems is plausible, because individuals often use analogies when they are failing on a target problem (e.g., Novick, 1988b), and hence that failure may be used to guide retrieval of related problems. For example, in the same way that subjects can be reminded by similar problem features (e.g., both problems are about pizza; Ross, 1984), perhaps they can be reminded by similar features of failed solutions (e.g., this solution procedure seems to be failing the same way as that other solution procedure).

### *Previous Evidence for Use of Failures*

Is there any evidence that failures are used in processing? Recent computational models of problem solving and memory include failure-based processing as an important feature that facilitates reminding (e.g., Schank, 1982), analogical problem solving (Carbonell, 1986), and learning (Laird, Newell, & Rosenbloom, 1987; VanLehn, 1988, 1989). Psychological research on discovery learning indicates that use of an ineffective strategy by subjects for solving water-jar problems during training facilitated production of the strategy when it was effective for solving later transfer problems (McDaniel & Schlager, 1990). In studies of nonanalogical transfer in simple verbal riddles, transfer was found to be higher when acquisition statements were presented in a format in which a puzzling concept typically induced by the riddle (e.g., a man marries several women) was given before the concept that correctly solved it (e.g., the man is a clergyman; Adams, Kasserman,

Yearwood, Perfetto, Bransford, & Franks, 1988; Lockhart, Lamon, & Gick, 1988) was given. The puzzlement can be seen as a kind of failure in comprehension. Unlike in the water-jar example, the initial failure is inappropriate for the solution of the later riddle, but it is useful to access the correct solution that has been paired with it.

### *Failed Solutions and Problem Representation*

The hypothesis that failed solution attempts are used in problem solving gains some theoretical support from Duncker (1945), who suggested that problem solvers "learn from mistakes" in that they use aspects of perceived failures as "an additional demand" (p. 13) into the original form of the problem. Duncker gave the example of solving the radiation problem. After the subject generates a solution of sending rays down the esophagus to the tumor, the solution might be rejected if the subject realizes that the esophagus is not completely free of healthy tissue. The next reformulation of the problem might involve looking for a path to the stomach that is completely free of healthy tissue, or perhaps a completely different approach (e.g., desensitizing healthy tissue) once it is realized that there are no tissue-free paths to the stomach.

In Duncker's (1945) words, "The solution properties found at first . . . always serve as productive reformulations of the original problem" (p. 8). In contemporary information-processing terms, the subjects incorporate the failed solution into the representation of the problem, as an additional problem constraint.

### *Failed Solutions and the Mutilated Checkerboard Problem*

Because matching source and target problem representations facilitate retrieval of the source problem (Novick, 1988a, 1988b), and failed solutions of a problem are incorporated into its representation (Duncker, 1945), we hypothesized that when target solutions persistently failed, analogous failed solutions produced to the source problem may facilitate spontaneous transfer. To test this idea, we needed a target problem that typically produced persistent solution failures. The mutilated checkerboard problem (referred to hereinafter as the checkerboard problem; see Figure 1) was chosen. Persistent failures occur on this problem because there is only one correct solution method that requires a change in representation (Kaplan & Simon, 1990), and this method is rarely used without hints (Kaplan & Simon, 1990).

For example, one representation that subjects often use that fails to solve the checkerboard problem is "covering." A covering representation includes the numbers of dominoes and squares and their geometrical arrangement as problem elements. The associated operator is covering the board with dominoes, which fails to solve the problem (Kaplan & Simon, 1990). In order to solve the checkerboard problem, a "parity" representation must be used. This representation includes dominoes that cover two squares of alternating color as problem elements and a corresponding "parity-check" operator that counts the number of squares of different colors (Kaplan & Simon, 1990). The correct parity solution is that no cov-

You are given a checkerboard and 32 dominoes. Each domino covers exactly two adjacent squares on the board. Thus, the 32 dominoes can cover all 64 squares of the checkerboard. Now suppose two squares are cut off at diagonally opposite corners of the board. If possible, show how you would place 31 dominoes on the board so that all of the 62 remaining squares are covered. If you think it is impossible, give a proof of why.

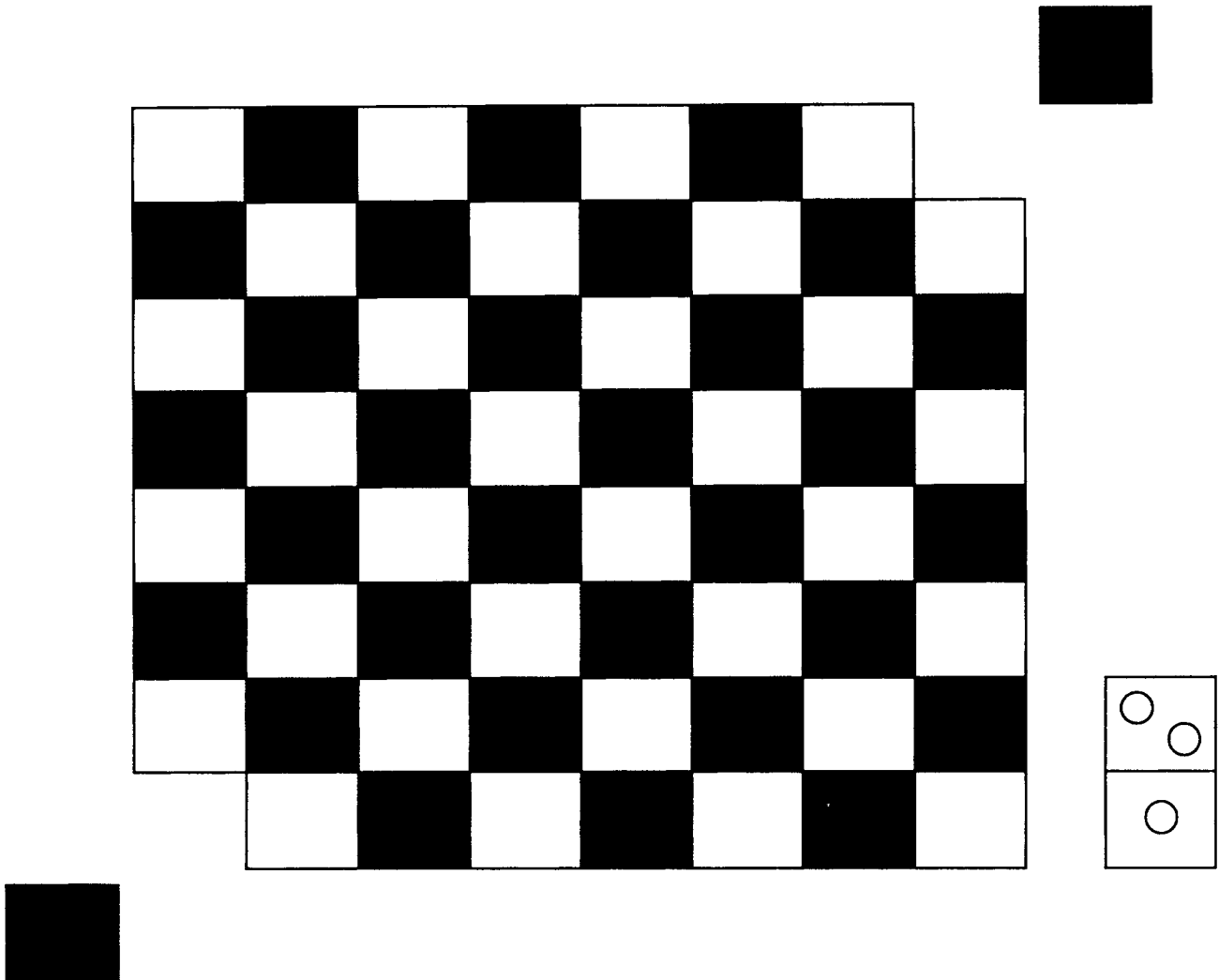


Figure 1. The standard mutilated checkerboard problem. (From *How to Solve Problems* [p. 29] by W. Wickelgren, 1974, New York: Freeman. Copyright 1974 by Freeman. Adapted by permission.)

ering is possible because two squares of the same color have been removed, leaving two squares of the same color that cannot be covered by a domino (Gardner, 1978; Kaplan & Simon, 1990; Wickelgren, 1974).

Kaplan and Simon (1990) found that one heuristic used by subjects in solving the checkerboard problem was *noticing invariants*, or noticing properties of the problem situation that did not change over repeated solution attempts. One invariant

of the covering solution attempts is that two squares are always left over when a covering attempt fails. If subjects noticed invariants while solving the same (i.e., checkerboard) problem, perhaps they would notice similar invariants across source and target problems. The noticing-invariants heuristic would operate across source and target problems only if the solution attempts were similar, however. Therefore, if subjects initially failed to solve both source and target problems anal-

ogously, subjects would be more likely to notice invariants of the failed solution attempts (facilitating spontaneous transfer) than if the source problem was initially solved correctly.

### Overview of the Experiments

The present research strategy was to use a source problem that objectively (Gick & Holyoak, 1987) had the same correct parity solution procedure as that required to solve the target checkerboard problem. However, the probability that subjects would reach the correct source solution procedure was manipulated by varying the surface form of parity features in the source problem. In particular, sometimes parity features were presented explicitly in a verbal statement of the problem, whereas at other times they had to be inferred from a verbal statement and a diagram presented with the problem. Because subjects' problem representations are influenced heavily by verbal statements of the problem (Hayes & Simon, 1974; Simon, 1978; Simon & Hayes, 1976), verbal exclusion of parity features may lead to a nonparity problem representation (e.g., covering) and corresponding solution failures. We predicted that the correct parity solution would be used less often to solve the source problem, and nonparity solution failures would be more frequent, when parity features were excluded explicitly from the verbal statement of the source problem than when they were included.

In Experiment 1 we examined transfer to the target checkerboard problem from a source problem that readily led to the correct parity solution. Our prediction was that spontaneous transfer performance would be poor because there were few nonparity solution failures to the source problem. In Experiments 2 and 3 we manipulated source solution failures by varying the explicit verbal inclusion of the parity feature, which was predicted to affect source solution failures. The prediction for target problem performance was that increasing analogous source solution failures would facilitate spontaneous transfer. Experiment 4 was a direct test of whether exposure to the analogous failed source solution was necessary for the transfer facilitation obtained in Experiments 2 and 3.

### Experiment 1

To ensure that subjects would easily obtain parity solutions to the source problem in Experiment 1, we used the problem elements *man* and *woman* in a dance situation that readily suggested pairing. However, because the concepts "man" and "woman" are quite different from the colors black and white, transfer failure could be due to dissimilarity between source and target problems (Holyoak & Koh, 1987; Novick, 1988a; Ross, 1987, 1989). Hence, in Experiment 1, surface similarity of target problem content also was varied in order to rule out problem dissimilarity as an explanation for any transfer failure.

### Method

*Subjects.* Subjects were 124 students in introductory psychology classes who received either course credit or \$5 for their participation.

*Design and materials.* Subjects were randomly assigned to one of four conditions: high-similarity analogy, low-similarity analogy, high-similarity control, or low-similarity control, with 31 subjects in each group. In the two analogy conditions, subjects received the "partner problem" (adapted from the "Russian matchmaker" analogy to the checkerboard problem; Hayes, 1978) as a source problem before solving the target problem. The partner problem involves an initial situation in which 20 pairs of men and women are at a dance, and two women leave. The problem is to determine whether the remaining people can form 19 man-woman couples (see Appendix A for the full text of the problem).

The answer to the partner problem is no, because after 2 women leave, there are 18 women and 20 men, who cannot form 19 heterosexual couples. The partner problem is analogous to the target checkerboard problem because both are parity problems, which have the following features in common. Pairs of elements exist (couples or dominoes), where a pairing consists of two elements of different kinds (man and woman or black and white squares). Two elements of one kind are removed (2 women or two black squares), leaving two more of the other kind of element (2 men or two white squares), which cannot be paired together. Solutions to both source and target problems involve applying the parity-checking procedure to check for equality of the number of the different elements (men and women or black and white squares).

The partner and checkerboard problems are dissimilar because they have different problem elements (men and women vs. white and black squares). Surface similarity between target problem content and source problem content was varied using the following procedure. Subjects in the low-similarity analogy condition received the standard black-white checkerboard at transfer (Gardner, 1978; Wickelgren, 1974; see Figure 1). In the high-similarity analogy condition, subjects solved a man-woman version of the target checkerboard problem at transfer. The two versions of the checkerboard problem are identical except that the word *board* was substituted for the word *checkerboard* in the verbal statement of the man-woman version, and the words *woman* and *man* were substituted for the colors black and white, respectively, in the squares of the diagram that accompanied the man-woman version (see Figure 2). Changing the square content in the checkerboard problem from white and black to man and woman in the present study did not affect the solution, because a domino still covers squares of different kinds, which can be represented by different kinds of labeling systems (Kaplan & Simon, 1990; Wickelgren, 1974). However, the source and target problems were higher in surface similarity in the high-similarity condition, because the words *man* and *woman* were common to both problems.

The two low- and high-similarity control conditions were used in order to establish baseline rates of solving the two versions of the target checkerboard problem without any source problem.

*Procedure.* All subjects were tested in small groups of up to 4 people. Subjects were told that they would be solving problems and that they should attempt to make their answers to the problems as explicit as possible. Subjects were given paper-and-pencil versions of all problems.

In the analogy conditions, subjects were allowed 10 min to solve the partner problem. In order to ensure that all subjects acquired the solution to the partner problem before solving the target problem, they were then presented with the solution to the partner problem and asked to rate it on a scale from 1 (*very poor*) to 5 (*very good*).

Subjects then proceeded immediately to the transfer task. They were given 10 min to solve either the standard or man-woman version of the checkerboard problem. During this first attempt at solving the checkerboard problem, no mention was made of its relationship to the previous partner problem. After their first attempt, subjects were given 5 min to recall the correct solution to the partner problem. This was done to determine whether failure to transfer was due to



problem. The first attempt lasted 10 min and the second lasted 7 min to equate total time on the target problem with the analogy conditions. During their second attempt, control subjects were asked to solve the problem again.

All subjects received a questionnaire at the end of the experiment that asked them about previous experience with the checkerboard problem. None of the subjects indicated prior knowledge of the problem.

## Results

In all four experiments, data for the source and target problems were scored independently by two coders, who were unaware of the experimental hypotheses. In each of Experiments 1 and 2, there were 11% disagreements overall between the two coders. In Experiments 3 and 4, there were 10% and 13% rates of disagreement, respectively, between the two coders. A third coder resolved the disagreements between the first two.

If a subject used more than one type of solution, the final attempt was included in analyses. In Experiment 1, scoring of the target problem was done without knowledge of analogy versus control condition, and scoring of the source problem was done without knowledge of low- versus high-similarity condition. Maximum likelihood chi-squares ( $G^2$ ) were calculated for the frequency data (Bishop, Fienberg, & Holland, 1975).

*Source problem performance.* Subjects' solutions to the partner problem were scored as follows. An answer was scored as a correct parity solution if it said that 19 couples could not be formed and it contained two of the following three criteria: (a) A couple consists of a man and a woman, (b) two women have left, and (c) therefore two men will be left over who cannot form a couple. This scoring is analogous to that used to score the checkerboard problem (discussed below in more detail).

As predicted, most subjects solved the partner problem, and there were no differences in the rates of solution between the two analogy conditions (77% in both the low- and high-surface-similarity conditions). Most (10 of 15) failed solutions to the partner problem consisted of violations of problem constraints (e.g., let the caller dance or let a woman dance with more than one man). There were no differences between the two analogy conditions in their ratings of the solution to the partner problem provided by the experimenter (3.71 vs. 3.87 in the low- vs. high-similarity condition),  $t(60) = .52$ ,  $p < .6$ . (Two subjects did not provide ratings.)

*Target problem performance: Correct solutions.* An answer to the checkerboard problem was scored as a correct parity solution if it stated that the solution was impossible and contained at least two of the following three features (solution features for the man-woman version are in parentheses): (a) A domino covers one black square and one white square (one man and one woman square), (b) two black (women) squares have been removed, and (c) therefore two white (men) squares are left over, which cannot be covered by a domino. This criterion of two features is similar to that required by Kaplan and Simon (1990) for a solution to qualify for a "rough proof." A solution that contained only one parity

feature was not considered correct, partly because subjects may mention a parity feature in passing while they are working in a nonparity problem representation (e.g., we can cover all the squares except for those white ones up there; Kaplan & Simon, 1990).

As in previous work in which the informing hint was administered within subjects (e.g., Catrambone & Holyoak, 1989; Gick & Holyoak, 1983; Holyoak & Koh, 1987), we used two measures of transfer. Spontaneous transfer used correct solutions that occurred before the informing hint, and total transfer used correct solutions that occurred either before or after the informing hint.

The overall prediction was that spontaneous transfer to the standard black-white checkerboard problem (i.e., the low-similarity analogy condition) would be poor because most subjects gave the correct parity solution to the source partner problem, instead of nonparity failed solutions that are typically generated to the target checkerboard problem. The results indicate support for the hypothesis and show that there was no spontaneous transfer. Subjects in the low-similarity analogy condition did not produce more solutions before a hint than did subjects in the low-similarity control condition (3.2% vs. 6.5%, respectively),  $G^2(1, N = 62) < 1$ . There were more solutions in the low-similarity analogy condition in total (25.8% vs. 6.5%),  $G^2(1, N = 62) = 4.55$ ,  $p < .03$ .

If the poor spontaneous transfer performance was due to lack of surface similarity between source and target problems in the low-similarity analogy condition, then spontaneous transfer performance in the high-similarity analogy condition should have been better than in the low-similarity condition. However, there were no differences between the low- and high-surface-similarity analogy conditions in either spontaneous transfer (3.2% vs. 0.0%),  $G^2(1, N = 62) = 1.40$ ,  $p < .24$ , or total transfer (25.8% vs. 35.5%),  $G^2(1, N = 62) < 1$ . Indeed, the high-similarity analogy condition had a transfer pattern similar to that of the low-similarity analogy condition. There was no spontaneous transfer (0.0% vs. 3.2% in the high-similarity analogy vs. control condition),  $G^2(1, N = 62) = 1.40$ ,  $p < .24$ , but total transfer was significant (35.5 vs. 6.5%),  $G^2(1, N = 62) = 8.52$ ,  $p < .004$ .

*Target problem: Incorrect solutions.* A critical manipulation in Experiments 2 and 3 was to induce failure into the source problem, analogous to the failure that occurred in the target checkerboard problem because of nonparity representations of the target problem. Hence, incorrect solutions to the target problem were examined and categorized into three groups: nonparity solution failures, one-feature solution failures, and other solution failures. Nonparity solution failures did not give any parity features, and instead showed clear evidence of a problem representation other than parity as a final answer. These answers usually stated that the covering was impossible for some reason other than parity (e.g., the shape of the board was wrong; see Appendix B for further details of the scoring procedure and percentages of incorrect solutions of different types). As indicated in Appendix B, in both the baseline and (especially) analogy conditions, there were many nonparity solution failures (51.6% and 77.4% collapsed across the two control and two analogy conditions before a hint, respectively). In the analogy conditions, expe-

rience with solving the prior source problem for which the correct answer was “no, it’s impossible” may have made subjects more confident about giving this same answer to the checkerboard problem. However, because the reasoning was usually wrong, this led to a higher percentage of nonparity solution failures in the analogy than in the control conditions. The main point is that many subjects generated nonparity solutions to the checkerboard problem, as did subjects in the Kaplan and Simon (1990) study.

### Discussion

The results of Experiment 1 can be summarized as follows. Neither subjects in the low-similarity analogy condition nor those in the high-similarity analogy condition showed any spontaneous transfer from the source partner problem to the target checkerboard problem. Most of the subjects succeeded in using the parity solution procedure to solve the partner problem, whereas most of them failed to use the parity procedure in attempting to solve the checkerboard problem before a hint. After a hint was given, there was significant total transfer in the analogy conditions, indicating that some subjects could apply the source problem to solve the target problem with a hint, even if they did not do so spontaneously. Increasing the surface similarity of content in the checkerboard problem diagram to words in the source problem (i.e., *man* and *woman*) did not affect transfer.

### Experiment 2

There were two purposes to Experiment 2. The main purpose was to test the hypothesis that increasing nonparity solution failures initially generated to the source problem (before parity solution feedback) would facilitate transfer of the parity solution to the target checkerboard problem. In order to implement the manipulation designed to vary source solution failures, a diagram similar to that used in the target checkerboard problem was added to the source problem. In the diagram-only version of the source problem, the parity information was present only in the diagram and not in the verbal statement of the problem, similar to the format of the target checkerboard problem. The verbal versions of the source problem stated the parity feature explicitly in the verbal statement. Our prediction for source problem performance was that subjects would generate the correct parity solution less often, and would produce more nonparity solution failures, while attempting the diagram-only version of the source problem than when they attempted the verbal versions of the problem (Hayes & Simon, 1974; Simon & Hayes, 1976).

An additional purpose of Experiment 2 was to test whether the surface similarity manipulation of the common words *man* and *woman* used in Experiment 1 was simply too weak to be effective. In Experiment 2 we tested whether the use of a diagram in the source problem, similar to that used in the target checkerboard problem, would facilitate transfer. Similar diagrams in source and target problems have been shown to facilitate transfer by providing a retrieval cue that effectively acted as an informing hint for the source problem while

subjects were solving the target problem (Beveridge & Parkins, 1987; Gick, 1985). In one verbal condition of Experiment 2 (verbal-plus-diagram condition), a diagram was included with the source problem.

The predictions for target problem performance in Experiment 2 were as follows. If the poor spontaneous transfer obtained in Experiment 1 was due to a weak surface similarity manipulation (i.e., lack of diagrams), then transfer should be higher in the verbal-plus-diagram than the verbal-no-diagram condition. If increasing nonparity solution failures to the source problem facilitated transfer, then subjects in the diagram-only condition should have superior transfer than subjects in the verbal conditions.

### Method

*Subjects.* Subjects were 81 undergraduate students, who received either \$5 or course credit for their participation.

*Design and materials.* In order to carry out the manipulations of Experiment 2, the partner problem was modified to be the verbal version of the dinner party problem (see Figure 3).

The corresponding features between the dinner party and checkerboard problems were as follows: The table was analogous to the domino, and the man and woman seated at a table were analogous to the man and woman squares that were covered by a domino.

Subjects were randomly assigned to one of three conditions: diagram only, verbal-plus-diagram, and verbal-no diagram, with 27 subjects per group. In addition, the 31 subjects in the high-similarity control condition of Experiment 1 were used for a measure of the baseline solution rate to the man-woman version of the checkerboard problem. The man-woman version of the problem was used to maximize surface similarity.

The three versions of the dinner party problem that were used for the three conditions of Experiment 2 were identical except for the following features. In both verbal versions, the problem was as written in Figure 3, explicitly stating that each table seats only two people, a man and a woman, and that two women leave. In the verbal-plus-diagram condition, a diagram similar to that presented in Figure 2 (modified to fit the numbers in the dinner party problem; see Figure 3) was also presented with the verbal statement. In the diagram-only condition, the phrase “a man and a woman” was omitted from the verbal statement of the problem, and the phrase “two women leave” was changed to “two people leave.” Otherwise, the problem was identical to that used in the verbal-plus-diagram condition.

It is important to note that the manipulation was one of surface form of presentation of the same information. In the verbal versions of the source problem, the parity information (i.e., that a table seats one man and one woman and that two women leave) was present explicitly in the verbal statement. In the diagram-only source problem, the parity feature that a table seats only one man and one woman could be inferred from the diagram and the statement that a table seats only two people in a horizontal or vertical direction. The parity feature that two women leave could be inferred from the arrows in the diagram and the verbal statement that “two people leave, as indicated by the arrows in the diagram below.” The presentation of the diagram-only dinner party problem was similar to the usual format of the checkerboard problem, in that the parity information could be inferred from the diagram and the verbal statements that a domino covers adjacent squares and that diagonally opposite squares have been removed. Hence, the parity information in the source dinner party problem was more explicit and perhaps more salient in the verbal conditions, but it could be obtained in the diagram condition. A condition in which parity information was not included in

Thirty-six people, 18 men and 18 women, are at a dinner party, as illustrated schematically below. The 36 people are seated at 18 tables. Each table only seats two people, a man and a woman, who are sitting immediately next to each other, either in a horizontal or vertical direction. If two women leave, as illustrated by the arrows in the diagram below, can 17 tables be arranged to seat the other 34 people? The people are not allowed to move. Explain your reasoning.

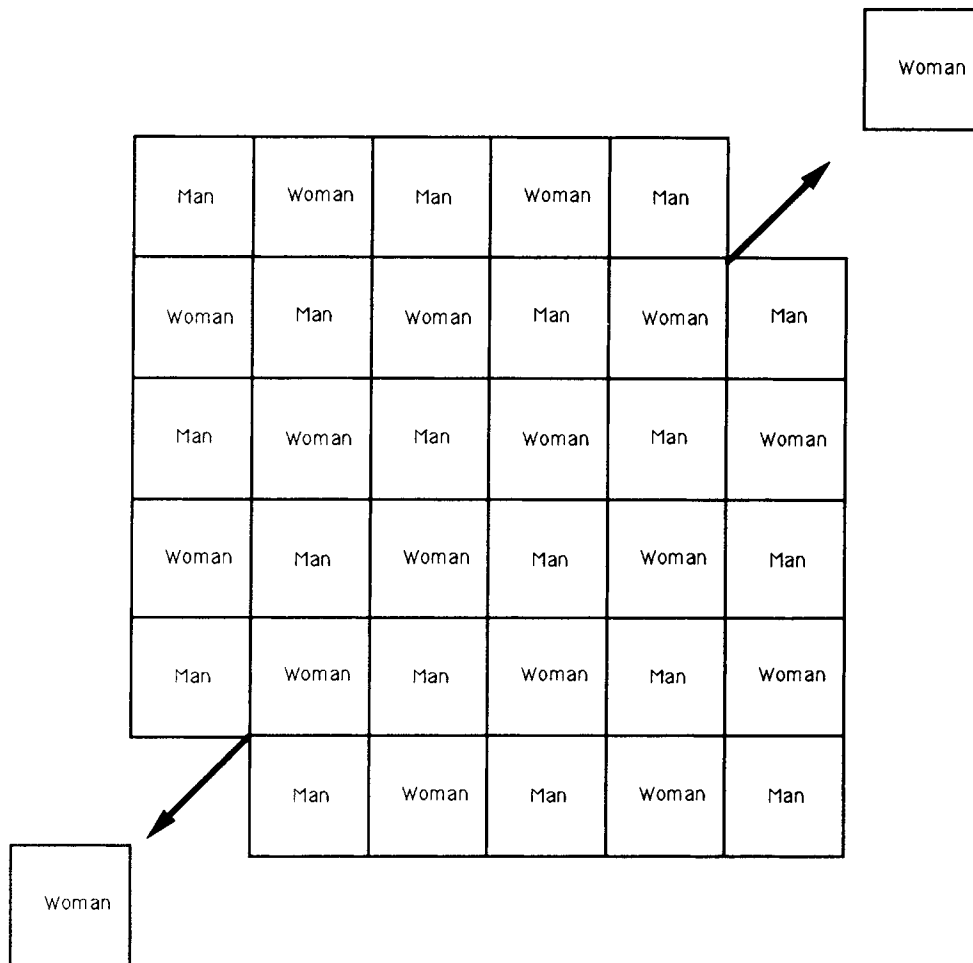


Figure 3. The verbal-plus-diagram version of the dinner party problem.

the verbal statement and no diagram was included was not used, because then parity information would be completely absent.

The solution to the source dinner party problem that was administered to subjects in all conditions contained the following information:

It is important to notice that, since a table seats two people who are immediately next to each other, any given table must always have a man and a woman seated at it. In removing two women, there are two fewer women than men who are left. So, if a table must always seat a man and a woman, and there are fewer

women than men, then the remaining 17 tables cannot seat the 34 people.

The diagram was presented with the solution in the diagram-only and verbal-plus-diagram conditions.

*Procedure.* Subjects followed the same procedure as in the experimental conditions of Experiment 1, with two minor modifications. When subjects received the source problem solution, they were asked to rewrite the solution in their own words, to provide a further measure of learning. (Rewrite performance was high, and no differences were obtained among conditions in Experiments 2–4; hence,

this measure is not discussed further.) In addition, subjects were warned 2 min before each problem-solving episode was over, to elicit some answer from subjects and reduce blank answers.

### Results and Discussion

Data from 2 subjects in the verbal-no-diagram condition were discarded because of familiarity with the checkerboard problem. Coders were unaware of condition in scoring the data from the checkerboard problem. In scoring the source problem data, coders could not be completely unaware of condition, because the verbal-no-diagram version of the problem had no diagram; however, coders were unaware of the verbal-plus-diagram and diagram-only conditions.

*Source problem performance.* In order to check the effectiveness of the manipulation of the verbal parity feature, subjects' final answers to the dinner party problem were scored for the presence of correct parity solutions and incorrect solutions. Scoring of the parity solution was analogous to that used for the partner problem in Experiment 1. Credit was given for a correct solution if it contained two of the following solution elements: (a) A table seats one man and one woman, and (b) two women have left, (c) leaving two men who cannot be seated at a table.

There were no differences in the rate of production of the correct parity solution to the two verbal versions of the dinner party problem (51.9% vs. 40% in the diagram vs. no-diagram condition),  $G^2(1, N = 52) < 1$ , and so they were combined for a comparison with the diagram-only condition. As predicted, the diagram-only version of the dinner party problem resulted in fewer correct parity solutions than the verbal conditions (7.4% vs. 46.2%),  $G^2(1, N = 79) = 14.06, p < .0002$ .

A crucial prediction was that subjects would produce more nonparity solution failures to the diagram-only version of the dinner party problem than to the verbal versions. Incorrect solutions were scored and classified into three categories: nonparity, one-feature, and other (analogous to the procedure used for the checkerboard problem of Experiment 1; see Appendix B). As expected, more subjects in the diagram-only condition produced nonparity solutions than did subjects in the combined verbal conditions (51.9% vs. 5.8%),  $G^2(1, N = 53) = 13.14, p < .0003$ .

*Target problem performance.* The first set of analyses tested the effects of including similar diagrams in source and target problems. Similar diagrams had no effect on transfer, because there were no differences between subjects in the two verbal conditions in either spontaneous transfer (7.4% vs. 8.0% in the verbal-plus-diagram vs. verbal-no-diagram conditions),  $G^2(1, N = 52) < 1$ , or total transfer (25.9% vs. 24.0% in the verbal-plus-diagram vs. verbal-no-diagram condition),  $G^2(1, N = 52) < 1$ . Lack of an effect of diagram was not due to the fact that it was ignored, because a total of 67% of subjects marked up the dinner party diagram somehow in the verbal-plus-diagram condition.

Because there were no differences in transfer performance between the two verbal conditions, they were combined and compared with the high-similarity control condition of Ex-

periment 1. Before receiving a hint, subjects in the verbal conditions did not perform significantly better than the control rate (7.7% vs. 3.2%),  $G^2(1, N = 83) < 1$ . Total transfer was significant in the verbal conditions (25.0% vs. 6.5% in the combined verbal conditions vs. control condition),  $G^2(1, N = 83) = 5.12, p < .02$ . This transfer pattern of no spontaneous transfer and significant total transfer replicates that found in the analogy conditions of Experiment 1 and shows no support for the hypothesis that the results of Experiment 1 were due to lack of surface similarity.

The second set of analyses addressed the effects of manipulating the parity feature in the source problem. If increasing nonparity solution failures to the source dinner party problem facilitated noticing, than spontaneous transfer performance should have been better in the diagram-only condition than in the verbal conditions. The results provide support for this hypothesis. Spontaneous transfer was significantly higher in the diagram-only condition than in the combined verbal conditions (33.3% vs. 7.7%),  $G^2(1, N = 79) = 8.07, p < .005$ . The facilitatory effect also was significant when considering total transfer (51.9% vs. 25.0%),  $G^2(1, N = 79) = 5.59, p < .02$ . When examining the percentages of subjects who generated the correct parity solution only on their second attempt (i.e., excluding the before-hint percentage), we found little difference between the diagram-only and verbal conditions (18.5% vs. 17.3%). Thus, total transfer obtained in the verbal conditions never caught up to that occurring in the diagram-only condition, and the advantage obtained in the latter group in spontaneous transfer remained in total transfer.

Because many subjects failed to solve the dinner party problem in the diagram-only condition, one hypothesis for its superior spontaneous transfer performance is that transfer was due to increased failure to solve the source problem per se, as opposed to increased failure due to the manipulation of form of presentation (verbal version vs. diagram). However, none of the 28 subjects who failed to solve either of the verbal versions of the dinner party problem solved the checkerboard problem before a hint. After a hint, only 3 (10.7%) of these subjects solved the checkerboard problem. Thus, there is no evidence that failure to solve the dinner party problem in the verbal conditions produced better transfer. Experiment 4 further addressed the reasons for the obtained transfer effects in the diagram-only condition of Experiments 2 and 3.

The results of Experiment 2 can be summarized as follows. The results obtained in the verbal conditions, which used the source dinner party problem, basically replicate the poor spontaneous transfer performance found in the analogy conditions of Experiment 1, which used the source partner problem. In Experiment 2, similar diagrams between source and target problems did not produce any spontaneous transfer to the checkerboard problem. Spontaneous transfer of the parity solution was facilitated by a manipulation that involved deletion of the parity feature from the verbal statement of the source problem, so that the feature had to be inferred from the diagram. This manipulation resulted in fewer correct parity solutions to the source problem and more nonparity failures. There was no evidence that failure to solve the source problem in the verbal conditions predicted transfer.

### Experiment 3

It is important to determine whether the success of the manipulation of source solution failures in Experiment 2 depended on the highly similar diagram used in both the diagram-only version of the source dinner party problem and the target checkerboard problem. If manipulating source solution failures is effective in the absence of high surface similarity of diagrams, then the results of Experiment 2 should be replicated by using the standard version of the checkerboard problem that has black and white squares, which are dissimilar to the man and woman squares used in the diagram of the source dinner party problem.

#### Method

*Subjects.* Subjects were 62 undergraduate students who received either course credit or \$5 for their participation.

*Design.* The design was identical to that of Experiment 2, with one difference. In Experiment 3, the verbal-no-diagram condition was dropped, because the results of Experiment 2 indicated that the diagram had no effect in the verbal conditions. Subjects were randomly assigned to conditions, with 31 subjects in each of the diagram-only and verbal-plus-diagram conditions. In addition, the 31 subjects in the low-similarity control condition of Experiment 1 were used for a measure of baseline solution rate to the black and white checkerboard problem.

*Materials and procedure.* The materials and procedure for Experiment 3 were identical to those of Experiment 2, except that all subjects solved the black-white checkerboard problem, instead of its man-woman version.

#### Results and Discussion

*Source problem performance.* Subjects' solution attempts to the dinner party and checkerboard problems were scored as in Experiment 2. Slightly more subjects in the verbal-plus-diagram condition than in the diagram-only condition generated the correct parity solution (19.4% vs. 12.9%), but this result was not significant,  $G^2(1, N = 62) < 1$ . Perhaps the reason for this difference between the results of Experiments 2 and 3 is that in Experiment 3 more subjects in the verbal-plus-diagram condition produced solutions that had only one parity feature in them (32.3% in Experiment 3 vs. only 7.4% in the verbal-plus-diagram condition of Experiment 2). We have no explanation for this difference in performance.

The critical result concerns the nonparity solution failures, because they were hypothesized to predict transfer to the checkerboard problem. There were more nonparity solution failures in the diagram-only condition than in the verbal-plus-diagram condition (41.9% vs. 9.7%),  $G^2(1, N = 52) = 8.45$ ,  $p < .004$  (see Appendix B), replicating the results of Experiment 2.

*Target problem performance.* The transfer results replicated the pattern found in Experiment 2. Subjects in the diagram-only condition, who produced more nonparity solution failures to the source dinner party problem than subjects in the verbal-plus-diagram condition, also produced more correct parity solutions to the target checkerboard problem before a hint (32.3 vs. 9.7%),  $G^2(1, N = 62) = 4.98$ ,  $p < .03$ . An additional 16.0% of subjects in each condition gen-

erated the correct solution on their second attempt only, so that subjects in the verbal-plus-diagram condition never caught up to the diagram-only condition, and the advantage obtained in spontaneous transfer remained in total transfer (48.4% vs. 25.8%),  $G^2(1, N = 62) = 3.43$ ,  $p < .06$ .

In addition, there was no spontaneous transfer from the verbal-plus-diagram version of the source problem to the target problem (9.7% vs. 6.5% in the verbal-plus-diagram vs. baseline condition),  $G^2(1, N = 62) < 1$ . Significant total transfer was obtained in the verbal-plus-diagram condition (25.8% vs. 6.5% in the baseline condition),  $G^2(1, N = 62) = 4.55$ ,  $p < .03$ .

Comparing the percentages in Experiments 2 and 3 revealed very little difference in spontaneous and total transfer in both the diagram-only and verbal-plus-diagram conditions, indicating that changing the surface similarity between source and target problems by using black and white colors instead of the words *man* and *woman* made no difference to transfer performance. This result further supports the hypothesis that surface dissimilarity is not responsible for the spontaneous transfer failure obtained in the analogy conditions of Experiment 1 and further suggests that surface similarity of diagrams has no effect on transfer from the diagram-only conditions.

### Experiment 4

In Experiments 2 and 3, transfer of the correct parity solution to the target checkerboard problem was facilitated in the diagram-only conditions, when the parity features of the source dinner party problem were not presented explicitly in the verbal statement. This resulted in the production of many nonparity solution failures, before feedback on the correct solution to the source problem was given. The failure hypothesis states that the nonparity solution failures produced to the diagram-only version of the source dinner party problem were important in facilitating transfer and predicted that if they were prevented, transfer to the checkerboard problem would be poorer than when source solution failures occurred.

However, in Experiments 2 and 3, source solution failure was confounded with problem type. That is, subjects who received the diagram-only problem received a different version of the source dinner party problem than did subjects in the verbal-plus-diagram conditions, and they also tended to produce more nonparity solution failures. An alternative to our failure hypothesis is that the version of the source problem that was given, and not the solution failures produced to it, was responsible for transfer. For example, perhaps the diagram-only version of the dinner party problem was more similar to the checkerboard problem than the verbal-plus-diagram version of the dinner party problem. This alternative hypothesis predicted that if solution failures were prevented in the diagram-only version of the dinner party problem, transfer performance would still resemble that obtained when failures occur (e.g., through matching source and target problem features).<sup>1</sup>

<sup>1</sup> We thank Kevin Miller for comments that motivated Experiment 4.

## Method

*Subjects.* Subjects were 87 undergraduate students who were given either course credit or \$7 for their participation in the experiment.

*Design and materials.* The diagram-only version of the source dinner party problem was used in all conditions of Experiment 4, because transfer facilitation was obtained only in that condition. Subjects were randomly assigned to one of three conditions: failure-solve, correct-given, and failure-given. In the failure-solve condition, subjects solved the problem as in Experiments 2 and 3. In the correct-given condition, subjects were given the problem and its correct solution to copy before regular solution feedback was given.

However, if transfer is higher in the failure-solve condition than in the correct-given condition, it could be partly because subjects solve the source problem, rather than read it (Adams et al., 1988; Needham & Begg, 1991). Hence, to separate the effects of solving the problem per se from those of producing failed solutions, we added a failure-given condition in which subjects did not solve the source problem, but instead copied it along with two nonparity failed solution attempts, before correct solution feedback was given.

Figure 4 illustrates the two failed covering solution attempts given to subjects. These attempts are typical of some non-parity solution failures produced by subjects as their final answer to the diagram-only dinner party problem in Experiments 2 and 3. Subjects were told that the two diagrams represent two attempts to match the 17 remaining tables and 34 people.

The correct solution feedback was the same as in Experiments 2 and 3. All subjects solved the standard black-white version of the checkerboard problem.

*Procedure.* The procedure for subjects in the failure-solve condition was identical to that used in Experiments 2 and 3. The procedure for the subjects in the correct-given condition was to first read the problem statement (presented with the diagram) and correct solution and then copy all of this information before the regular correct feedback was given. Subjects in the failure-given condition performed a similar procedure, except that they read and then copied the problem and diagram with the two (failed) solution attempts. In order to limit the opportunity of subjects in the correct-given and failure-given conditions to attempt their own problem solutions, these subjects were required to keep copying the information before the regular correct solution feedback was given. A total of 10 min was given for the acquisition task in each of the three conditions. The target problem was administered as in Experiments 1-3.

## Results and Discussion

Data from 3 subjects in the failure-given and 1 subject in the failure-solve conditions were eliminated because of familiarity with the checkerboard problem. Scoring of the target but not the source problem was done without knowledge of condition.

*Source problem performance.* As in Experiments 2 and 3, most subjects in the failure-solve condition of Experiment 4 failed to solve the dinner party problem. Only 17.9% of subjects produced the correct parity solution, and 35.7% produced clear nonparity solution failures. In the correct-given condition, 86.2% of subjects reproduced the problem and its correct parity solution. In the failure-given condition, 88% of subjects copied the problem and one or both of the two failed solution attempts.

*Target problem performance.* There were more correct parity solutions produced to the target checkerboard problem

in the failure-solve condition than in the correct-given condition before a hint (42.9% vs. 3.4%),  $G^2(1, N = 57) = 14.27$ ,  $p < .0002$ . Comparison of correct solutions generated only on the second attempt yielded similar percentages in the failure-solve and correct-given conditions (17.7% and 20.7%, respectively), with the result that differences obtained in spontaneous transfer remained in total transfer (60.7% vs. 24.1% in the failure-solve vs. correct-given condition),  $G^2(1, N = 57) = 8.02$ ,  $p < .005$ .

The transfer facilitation in the failure-solve condition was not due to a general problem-solving effect, because there were no significant differences in transfer performance between the failure-solve and failure-given conditions before a hint (42.9% and 34.6%, respectively),  $G^2(1, N = 54) < 1$ , or in total transfer (60.7% and 42.3%, respectively),  $G^2(1, N = 54) = 1.846$ ,  $p < .18$ . Moreover, there was significantly better spontaneous transfer obtained in the failure-given condition than in the correct-given condition,  $G^2(1, N = 55) = 9.91$ ,  $p < .002$ . However, this difference did not remain in total transfer,  $G^2(1, N = 55) = 2.07$ ,  $p < .15$ , perhaps because very few subjects (only 7.7%) in the failure-given condition generated the correct solution on their second attempt. This extremely low percentage of correct solutions given only with the informing hint suggests that when subjects in the failure-given condition tried to apply the source problem on their second attempt, there may have been interference from first attempts to solve the target problem (Gick & Holyoak, 1983; Perfetto, Bransford, & Franks, 1983). Interference effects in the present study are discussed further in the General Discussion section.

Comparison of the correct-given condition with the baseline rate revealed that there was no spontaneous transfer in the correct-given condition (3.4% vs. 6.5%). Subjects in the correct-given condition had significant total transfer (24.1% vs. 6.5% in the baseline condition),  $G^2(1, N = 60) = 3.84$ ,  $p < .05$ . The transfer performance of subjects in the correct-given condition was similar to that obtained in the verbal conditions of Experiments 2 and 3 and the analogy conditions of Experiment 1.

Because the results of Experiment 4 indicate that spontaneous transfer was present in the failure-given and failure-solve conditions but absent in the correct-given condition, one question is whether exposure to nonparity solution failures was the only thing that was important for spontaneous transfer. For example, would nonparity solution failures to the source problem facilitate transfer if they were given to subjects with the verbal-plus-diagram version of the source dinner party problem that is ineffective for transfer when subjects solve it themselves? The answer is probably not, because when subjects generated these failures themselves in Experiments 2 and 3, they were not effective. That is, many subjects in the verbal-plus-diagram conditions of Experiments 2 and 3 (68% and 52%, respectively, including those who generated the correct parity solution or incorrect solutions as a final answer) drew covering patterns on the diagram, similar to those given to subjects in the failure-given condition of Experiment 4. Yet we know that there was no spontaneous transfer in these verbal conditions of Experiments 2 and 3. In other words, because the results of Experiments 2 and 3

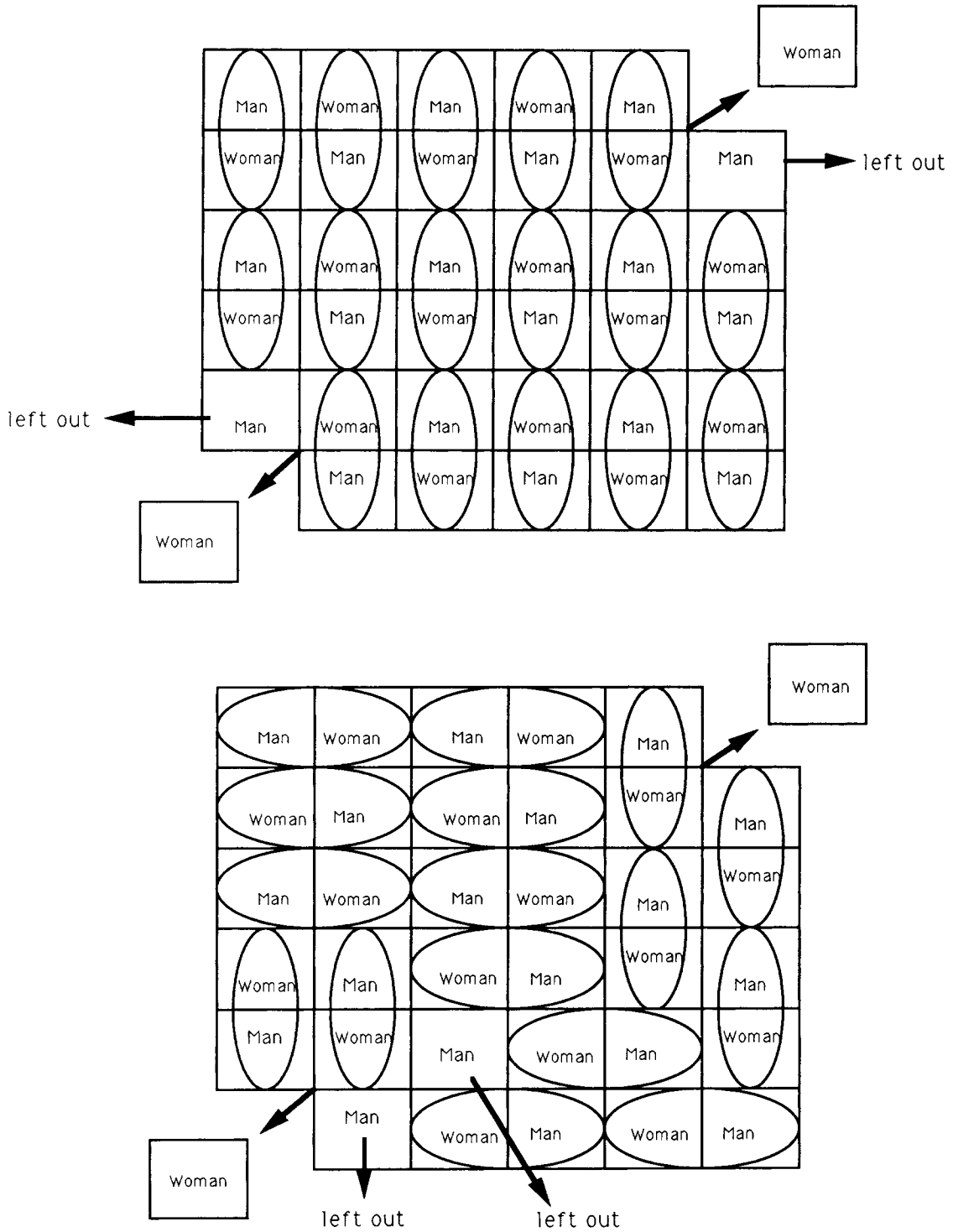


Figure 4. The two solution attempts presented with the problem in the failure-given condition of Experiment 4.

suggest that *generating* failed covering solutions to the verbal-plus-diagram version of the source problem was ineffective in transfer, and the results of Experiment 4 indicate that there were no differences in spontaneous transfer between providing

and generating failed solutions to the diagram-only version of the source problem, it is unlikely that *providing* failed solutions to the verbal-plus-diagram version of the source problem would be helpful.

Perhaps covering patterns are represented and used differently in the verbal-plus-diagram version of the dinner party problem than in both the diagram-only version of the dinner party problem and the checkerboard problem. For example, subjects might use covering patterns to check their final parity or other solutions. In general, parity features are more likely to be represented when covering attempts are included with the verbal-plus-diagram dinner party problem, because subjects generally are more likely to adopt parity representations of this problem. In other words, it may be the combination of the diagram-only version of the dinner party problem and failures that are induced by it that are effective for transfer. If these failures were included with another version of the dinner party problem (i.e., the verbal version), they might be represented differently, and the ensuing dissimilarity to the representation of failed solution attempts that occur in the subsequent checkerboard problem would not facilitate transfer.

In sum, the results of Experiment 4 indicate that there was no spontaneous transfer obtained when subjects were given no opportunity to process failed solutions to the source problem, but instead read and then copied the problem with its correct solution. Spontaneous transfer was obtained only when subjects processed (i.e., generated or were given) nonparity solution failures to the source problem. Total transfer performance in the correct-given condition never caught up to that obtained in the failure-solve condition.

### General Discussion

In general, the present experiments support our hypothesis that source solution failures that are analogous to target solution failures facilitate spontaneous transfer. In this section, we review our major findings, discuss their relation to other work, and suggest their implications for theories of analogy.

#### *Surface Similarity, Problem Representation, and Transfer*

Surface similarity has been shown to be a powerful variable that affects spontaneous (Holyoak & Koh, 1987; Novick, 1988a; Ross, 1984, 1987) and sometimes informed (Ross, 1989) transfer. Yet, in Experiments 1–3, increasing surface similarity of words or diagrams between source and target problems did not affect transfer. At first glance, the present results seem to contradict those found in the literature. However, our results can be explained by examining the kinds of surface similarity manipulations used in the present experiments, those used in previous research, and the relationship of surface similarity to problem representation.

In previous research, the surface similarity of problem features that were part of subjects' representations was manipulated (e.g., ray vs. laser; Holyoak & Koh, 1987). In contrast, we manipulated surface features of the problem that were not always part of subjects' problem representations. First, consider the content of the squares (i.e., black-white or man-woman). If subjects are using a covering representation in solving the target checkerboard problem, the numbers and

arrangements of the squares and dominoes are represented, but the fact that a domino covers squares of different contents is not (Kaplan & Simon, 1990). Hence, changing square content from black-white to man-woman did not have an effect in Experiment 1 or in comparisons of the diagram-only conditions of Experiments 2 and 3.

Similarly, there was no effect of adding a similar diagram to the source problem in the verbal conditions of Experiment 2 because subjects probably used the diagrams differently in the source and target problems. For example, in the nonparity representation of the target problem, the diagram may have been used for attempted covering solutions, which initially seem possible because the numbers of squares and dominoes match, and the content of the squares is ignored. In contrast, in the verbal version of the source dinner party problem, the content of the squares is not ignored, because it is explicitly given in the verbal statement. Subjects may initially think the solution is impossible because the numbers of men and women do not match, and the diagram is used only to check the parity solution (see Experiment 4 for further discussion of a related point).

In contrast to the effects of similar words and diagrams, the surface similarity manipulation of form of presentation of source problem information (verbal vs. diagram; Experiments 2 and 3) had an effect because it altered the representation of the source problem to match the representation of the target checkerboard problem. That is, parity problem elements are not included, and nonparity operators (e.g., covering) are used. Thus, surface similarity may be an effective variable only if the problem elements that are manipulated are likely to be incorporated into subjects' representations of the problems.

The present findings are consistent with recent work on transfer in algebra word problems. The particular content in which a problem is embedded (e.g., acceleration vs. finance) has less of an effect on transfer than do other surface features that affect how variables are represented in the two problems (e.g., expressing money variables in the algebra problem as rates vs. amounts; Bassok, 1990; Bassok & Holyoak, 1989). These and the present results support Ross's (1989) recent point that different types of surface similarity exist and have different effects on transfer performance.

#### *Source Solution Failures and Spontaneous Transfer*

The results of Experiments 1–4 show that noticing or retrieval of the source problem was facilitated by source solution failures, because spontaneous transfer was not obtained unless subjects had been exposed to failed solutions to the source problem that were analogous to those typically generated to the target problem. Noticing of the source problem may be facilitated through a noticing-invariants mechanism (Kaplan & Simon, 1990) that operates across similar failed solutions. For example, if subjects are attempting covering solutions that they think should work in the checkerboard target problem, but all solutions seem to fail, they may notice a similarity to failures that occurred in the source problem (e.g., two squares are always left over). Moreover, solution failures to both source and target problems may

become incorporated into their problem representations (e.g., as additional constraints; Duncker, 1945), which increases their similarity, thereby increasing transfer (e.g., Novick, 1988a). In addition, noticing might be accomplished through analogous impasses reached in solving source and target problems. Subjects may get stuck and not know what to do next in solving the target problem (VanLehn, 1988, 1989) and could be reminded of similarly being stuck in solving the source problem (Gick, 1990).

In contrast, if the initial source solutions are correct or are failures different from initial target attempts (e.g., "other" incorrect solutions that seat two men together), noticing can still occur through similar problem features (e.g., words and diagrams), but if they are used differently in solving the source and target problems (e.g., in different problem representations), then it is less likely that they will be noticed.

### *Source Solution Failures and Total Transfer*

In Experiments 2–4, the differences obtained in spontaneous transfer between subjects who did not generate non-parity solution failures to the source problem and those who did generate failures remained in total transfer; subjects who did not generate failures never caught up to the other subjects, even after the informing hint had been given. However, this result does not necessarily imply that processes such as mapping and application that occur after noticing are also facilitated by the failure manipulation. The within-subject procedure of administering the informing hint in the present experiments makes it difficult to determine the exact cause of the total transfer effects. For example, subjects' initial attempts to solve the target problem before the hint may have produced a set effect and limited the usefulness of the hint, making it difficult for the groups with low rates of spontaneous transfer (i.e., groups who did not process solution failures) to catch up in total transfer (Gick & Holyoak, 1983).

In addition, we must note that total transfer was far from perfect in the failure conditions. Approximately 40–50% of subjects who processed source solution failures in Experiments 2–4 still did not produce the correct target solution, even after the informing hint, suggesting possible interference in these conditions as well. A more general set effect may also exist in all conditions, in that subjects may give up trying to solve the difficult target problem, especially if they do not easily notice the relevance of the source problem. In the absence of more specific "mapping hints" (Novick & Holyoak, 1991), such as to consider the color of squares in the board (Kaplan & Simon, 1990), some subjects in all conditions may be unable to change to a parity representation of the checkerboard problem. Further research is needed to determine the source of difficulty of subjects in both the failure and non-failure conditions in applying the source problem to the difficult target checkerboard problem.

### *Relation of Present Studies to Other Work*

In the present experiments, correct source solutions resulted in no spontaneous transfer and low total transfer. In contrast,

other research on transfer in insight problems (e.g., Duncker's, 1945, ray problem; Gick & Holyoak, 1980, 1983) shows low spontaneous and high total transfer when the source problem is presented with the desired convergence solution in the form of a story. However, it must be noted that in transfer in the ray problem, the source story does in fact present an initial failed attempt in the first paragraph, before the convergence solution is given.

In addition, and perhaps more important, when subjects fail to solve the ray problem, they are not necessarily using an inappropriate problem representation. For example, some subjects realize they have to decrease the intensity of the rays on the way to the tumor, but they do not implement convergence to do this. Instead they may lower and then raise ray intensity at the tumor, for example (Duncker, 1945). Hence, subjects do not necessarily have to switch problem representations when solving the ray problem, although they may have to change their representation of the rays to allow for summative value and different directions.

In work on the nine-dots problem (drawing four lines through a  $3 \times 3$  array of dots without lifting the pen off the paper), Lung and Dominowski (1985) found that when subjects first practiced on dot problems that required extension of lines outside dot boundaries, performance improved on the nine-dots problem, provided the subjects eventually arrived at the correct solution to the practice problems. Subjects were not given correct feedback to the practice dot problems. Lung and Dominowski do not give information about the failed solution attempts, but presumably some initial failures on the practice problems were because subjects did not extend the lines. Hence, these results with the nine-dots problem may be consistent with the present findings that if subjects initially solved practice problems using inappropriate problem representations, and then arrived at the correct solution, transfer to other problems occurred. Indeed, Lung and Dominowski concluded that subjects learn the correct representation and strategy in the practice problems.

### *Implications for Theories of Analogy*

Analogical problem solving is used as a strategy when attempts to solve the problem directly fail (Novick, 1988b). The present experiments suggest that aspects of the attempted solutions that have failed aid the noticing process. Theoretical frameworks of analogical transfer have stressed aspects of the source and target problem features and relations (e.g., Gentner, 1989; Holyoak & Thagard, 1989, 1990; Novick, 1988b; Ross, 1989) but not initial failed solution attempts. However, current theories of analogy can accommodate the present findings if one accepts the present suggestion that solution failures become part of subjects' representations of problems (Duncker, 1945).

### Conclusion

The general conclusion that introducing failures into the source solution process will always improve spontaneous transfer cannot be drawn from the present work. In fact,

evidence from the present experiments suggests that failure to solve the acquisition problem per se was not predictive of transfer. Whether failures processed in acquisition are successful in producing transfer depends on whether analogous failures are induced spontaneously by subjects in solving the target problem. In other words, it is the similarity of the processing induced by the acquisition and transfer tasks that is important (Adams et al., 1988; Lockhart et al., 1988; Morris, Bransford, & Franks, 1977; Needham & Begg, 1991).

Further research is needed to establish the generality of the present findings with other problems that require a change in representation, such as physics problems (e.g., McCloskey, 1983), and other problems different from the checkerboard problem (e.g., the nine-dots problem or the matchsticks problem; Mayer, 1983) that are also known as "insight" problems, in which a sudden subjective feeling of insight or understanding into the correct solution often accompanies the change of representation (Duncker, 1945; Kaplan & Simon, 1990). In addition, in the present experiments, all subjects were novices, in that they had no experience with the checkerboard problem. The checkerboard problem, however, is one of a class of "tiling problems" (Gardner, 1978), found in combinatorial geometry. Recent work on expertise in transfer (Novick, 1988a) predicts that subjects with more background in this or another branch of mathematics might be less affected by the manipulations of the verbal statement of the parity feature used in the present research, because the subjects might naturally represent the parity features from the diagrams of the dinner party and checkerboard problems (Chi, Feltovich, & Glaser, 1981; Novick, 1988a).

The present results have implications for education. Pointing out and then correcting errors of representation in training problems may be more beneficial for students who will later be working on their own than is solving problems that readily lead to the correct representation of the transfer problem.

## References

- Adams, L., Kasserian, J., Yearwood, A., Perfetto, G., Bransford, J., & Franks, J. (1988). Memory access: The effects of fact-oriented versus problem-oriented acquisition. *Memory & Cognition, 16*, 167-175.
- Bassok, M. (1990). Transfer of domain-specific problem-solving procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 522-533.
- Bassok, M., & Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 153-166.
- Beveridge, M., & Parkins, E. (1987). Visual representation in analogical problem solving. *Memory & Cognition, 15*, 230-237.
- Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.
- Carbonell, J. (1986). Derivational analogy: A theory of reconstructive problem solving and expertise acquisition. In R. Michalski, J. Carbonell, & T. Mitchell (Eds.), *Machine learning: An AI approach* (Vol. 2, pp. 371-392). Los Altos, CA: Morgan Kaufman.
- Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1147-1156.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.
- Duncker, D. (1945). On problem solving. *Psychological Monographs, 58*(5, Whole No. 270).
- Gardner, M. (1978). *Aha! Insight*. New York: Freeman.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). New York: Cambridge University Press.
- Gick, M. L. (1985). The effect of a diagram retrieval cue on spontaneous analogical transfer. *Canadian Journal of Psychology, 39*, 460-466.
- Gick, M. L. (1990). Transfer in insight problems: The effects of different types of similarity. In K. J. Gilhooly, M. T. G. Keane, R. H. Logie, & G. Erds (Eds.), *Lines of thinking* (Vol. 1, pp. 251-265). Chichester, England: Wiley.
- 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.
- Gick, M. L., & Holyoak, K. J. (1987). The cognitive basis of knowledge transfer. In S. Cormier & J. Hagman (Eds.), *Transfer of learning* (pp. 9-46). San Diego, CA: Academic Press.
- Hayes, R. (1978). *Cognitive psychology: Thinking and creating*. Monterey, CA: Brooks-Cole.
- Hayes, J. R., & Simon, H. A. (1974). Understanding written problem instructions. In L. Gregg (Ed.), *Knowledge and cognition* (pp. 167-200). Potomac, MD: Erlbaum.
- Holyoak, K. J. (1985). The pragmatics of analogical transfer. *The Psychology of Learning and Motivation, 19*, 59-87.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15*, 332-340.
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295-355.
- Holyoak, K. J., and Thagard, P. (1990). A constraint-satisfaction approach to analogue retrieval and mapping. In K. Gilhooly, M. Keane, & G. Erds (Eds.), *Lines of thought* (pp. 205-220). Chichester, England: Wiley.
- Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*, 374-419.
- Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. *Artificial Intelligence, 33*, 1-64.
- Lockhart, R. S., Lamon, M., & Gick, M. (1988). Conceptual transfer in simple insight problems. *Memory & Cognition, 16*, 36-44.
- Lung, C. T., & Dominowski, R. (1985). Effects of strategy instructions and practice on nine-dot problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 804-811.
- Mayer, R. E. (1983). *Thinking, problem solving, cognition*. San Francisco: Freeman.
- McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 299-323). Hillsdale, NJ: Erlbaum.
- McDaniel, M. A., & Schlager, M. S. (1990). Discovery learning and transfer of problem-solving skills. *Cognition and Instruction, 7*, 129-159.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519-533.
- Needham, D., & Begg, I. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory oriented training promotes memory for training. *Memory & Cognition, 19*, 543-557.
- Novick, L. R. (1988a). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning,*

*Memory, and Cognition*, 14, 510–520.

Novick, L. R. (1988b). Analogical transfer: Processes and individual differences. In D. H. Helman (Ed.), *Analogical reasoning* (pp. 125–145). Norwell, MA: Kluwer.

Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 398–415.

Perfetto, G. A., Bransford, J. D., & Franks, J. J. (1983). Constraints on access in a problem solving context. *Memory & Cognition*, 11, 24–31.

Polya, G. (1957). *How to solve it*. Princeton, NJ: Princeton University Press.

Reed, S. K. (1987). A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 124–139.

Ross, B. H. (1984). Reminders and their effects in learning a cognitive skill. *Cognitive Psychology*, 16, 371–416.

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

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

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.

Schank, R. C. (1982). *Dynamic memory*. Cambridge, England: Cambridge University Press.

Simon, H. A. (1978). Information-processing theory of human problem solving. In W. Estes (Ed.), *Handbook of learning and cognitive processes* (pp. 271–295). Hillsdale, NJ: Erlbaum.

Simon, H. A., & Hayes, J. R. (1976). The understanding process: Problem isomorphs. *Cognitive Psychology*, 8, 165–190.

Spencer, R. M., & Weisberg, R. W. (1986). Context-dependent effects on analogical transfer. *Memory & Cognition*, 14, 442–449.

VanLehn, K. (1988). Toward a theory of impasse-driven learning. In H. Mandl & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems*. New York: Springer-Verlag.

VanLehn, K. (1989). Rule acquisition events in the discovery of problem solving strategies. In *Proceedings of the 11th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Wickelgren, W. (1974). *How to solve problems*. New York: Freeman.

### Appendix A

#### The Partner Problem

One Saturday night, at a local country dance, 40 people, 20 men and 20 women, showed up to dance. The dance was a “contra dance,” in which men and women face each other in lines. From 8 to 10 p.m., there were 20 heterosexual couples (consisting of one man and one woman each; i.e., two men or two women cannot dance together) dancing on the floor. At 10 p.m., however, two women left, leaving

38 people to dance. Could the dance caller make arrangements so that the remaining people could all dance together at the same time in 19 heterosexual couples? The dance caller must remain a caller only, and cannot take a partner. Answer yes or no, and explain the reasons for your answer.

### Appendix B

#### Incorrect Solutions

##### Scoring of Incorrect Target Solutions of Experiment 1

Examples of other nonparity solution failures were answers that stated that the solution was impossible because every way has been tried, there are always squares left over, or the shape of the board is

wrong. Nonparity solution failures often showed attempts to cover the board with dominoes.

Other solution failures included blank answers, answers that were not clear (e.g., they vacillated between possible and impossible), and answers that stated only that the solution was possible by violating

Table B1  
*Percentages of Subjects Who Produced Incorrect Solutions of Different Types to Target Problem Before Hint in Experiment 1*

Condition	One-feature	Nonparity	Other	N
Analogy				
Low similarity	3.2 (1)	83.9 (26)	9.7 (3)	31
High similarity	12.9 (4)	71.0 (22)	16.1 (5)	31
Control				
Low similarity	0.0 (0)	58.1 (18)	35.5 (11)	31
High similarity	6.5 (2)	45.2 (14)	45.2 (14)	31

Note. Frequencies are in parentheses.

Table B2  
*Percentages of Subjects Who Produced Incorrect Solutions to Source Problem in Experiment 2*

Condition	One-feature	Nonparity	Other	<i>N</i>
Diagram only	11.1 (3)	51.9 (14)	29.6 (8)	27
Verbal plus diagram	7.4 (2)	7.4 (2)	33.3 (9)	27
Verbal-no diagram	16.0 (4)	4.0 (1)	40.0 (10)	25

*Note.* Frequencies are in parentheses.

Table B3  
*Percentages of Subjects Who Produced Incorrect Solutions to Source Problem in Experiment 3*

Condition	One-feature	Nonparity	Other	<i>N</i>
Diagram only	22.6 (7)	41.9 (13)	22.6 (7)	31
Verbal plus diagram	32.3 (10)	9.7 (3)	38.7 (12)	31

*Note.* Frequencies are in parentheses.

problem constraints (e.g., a domino could cover diagonal instead of adjacent squares or a domino could be broken) or the representation used was not clear.

The scoring procedure used was conservative, in that it may underestimate the use of nonparity representations. For example, a subject who did not give any features of the correct parity solution and stated "Sure it can be done; break the domino in two" was probably using a nonparity representation. (See Table B1.)

#### *Scoring of Incorrect Solutions for the Source Dinner Party Problem in Experiments 2 and 3*

Nonparity solutions contained no parity features and usually stated that the solution was impossible for some other reason that was

clearly indicative of other problem representations, such as covering attempts at matching people to tables. One-feature solutions contained only one parity feature. Other incorrect solutions included those that stated that the solution was possible (e.g., by violating problem constraints and having more than two people at a table) and solution sheets that were blank or had no clear answer. (See Tables B2 and B3.)

Received May 17, 1991  
 Revision received October 2, 1991  
 Accepted October 27, 1991 ■