Journal of Experimental Psychology: General
1987, Vol. 116, No. 1, 50–67

Cognitive Constraints on Ordering Operations: The Case of Geometric Analogies

Laura R. Novick and Barbara Tversky
Stanford University

SUMMARY

Many tasks (e.g., solving algebraic equations and running errands) require the execution of several component processes in an unconstrained order. The research reported here uses the geometric analogy task as a paradigm case for studying the ordering of component processes in this type of task. In solving geometric analogies by applying mental transformations such as *rotate, change size,* and *add a part,* the order of performing the transformations is unconstrained and does not in principle affect solution accuracy. Nevertheless, solvers may bring cognitive constraints with them to the analogy task that influence the ordering of the transformations. First, we demonstrate that solvers have a preferred order for performing mental transformations during analogy solution. We then investigate three classes of explanations for the preferred order, one based on general information processing considerations, another based on task-specific considerations, and a third based on individual differences in analogy ability. In the first and third experiments, college students solved geometric analogies requiring two or three transformations and indicated the order in which they performed the transformations. There was close agreement on nearly the same order for both types of analogies. In the second experiment, subjects were directed to perform pairs of transformations in the preferred or unpreferred order. Both speed and accuracy were greater for the preferred orders, thus validating subjects' reported orders.

Ability differences were observed for only the more difficult three-transformation problems: High- and middle-ability subjects agreed on an overall performance order, but the highs were more consistent in their use of this order. Low-ability subjects did not consistently order the transformations for these difficult problems. The general information processing factor examined was working-memory load. A number of task factors have been shown to affect working-memory load during the solution of inductive reasoning problems. Of these, we chose to examine process difficulty. Because analogies are solved in working memory, performing more difficult transformations earlier may reduce working-memory load and facilitate problem solution. However, the observed performance order was not correlated with transformation difficulty. The first task-specific factor considered was that some transformations may be identified earlier, possibly because of perceptual salience, and that the performance order follows the identification order. However, the order of transformation identification also did not account for the order of transformation application in either experiment.

Solving geometric analogies is an imagery task in which the solution is constructed mentally by performing the inferred transformations on the given figure. Performance on geometric analogy tests is moderately to highly correlated with a variety of spatial visualization tests that involve image generation, transformation, and retention. The second task-specific factor considered was that the order of transformation application would resemble the order of using the corresponding information in an analogous physical task, the planning and execution of a drawing. Thus, in the fourth experiment, subjects were asked to indicate the order in which they would need specific types of information, each corresponding to a transformation, in order to draw a simple picture. This drawing order paralleled the order of performing mental transformations during analogy solution. Perceptual processing during object identification is also shown to proceed in a similar order. These results suggest that in solving geometric analogies, subjects tap into procedures and constraints common to the domains of imagery, drawing, and object identification.

In sum, we have shown that a logically unconstrained ordering of component processes does not necessarily imply that the observed ordering will be random. Rather, cognitive constraints imported from more familiar domains were shown to impose substantial agreement on the ordering of mental transformations during geometric analogy solution. We suspect that this type of finding is likely to obtain for other tasks that have theoretically unconstrained orders of component operations.

During the Rubik's Cube craze of the early 1980s, three books that described solutions simultaneously captured the first, second, and fourth places on the *New York Times* paperback bestseller list (Tierney, 1986). In the execution of the 52 steps on the shortest solution path, the cube appears disorganized until the very end. Two longer solution paths (as many as 120 moves in one case) yield a series of intermediate products that are recognizably closer to the final solution. Why did solvers generally prefer the longer methods? The longer solutions are chunked into meaningful units, so the extra steps are compensated for by a reduction in working-memory load. In contrast to the richness of Rubik's Cube, many tasks have essentially only a single solution path with a prescribed sequence of moves (discounting backtracking; e.g., Missionaries and Cannibals, puzzles in which two or more intertwined pieces of wire must be disentangled). For these tasks, solvers must perform the requisite steps in the prescribed order if they are to succeed.

More interesting are an intermediate set of tasks, from solving algebraic equations to running errands, that require the performance of a single set of component operations, but for which the order of executing the operations is optional. In these tasks, as in Rubik's Cube, solvers have a choice of solution paths. Thus, here also there may be cognitive reasons, such as reducing working-memory load, for selecting one sequence over another. Solving geometric analogies is a familiar and convenient example of such a task (see the top row of Figure 2 for a sample analogy). Several simple operations or transformations on geometric figures, such as *move* and *add half,* are inferred from the example terms (A and B) and then performed on the test term (C) to generate the solution (D). The order of performing the transformations is optional; that is, it does not matter whether one moves the figure first and then adds the other half to it or adds the other half first and then moves the figure, because the correct solution can be obtained in either case. Unlike the case of algebraic equations, there has been no formal instruction in geometric analogy solution techniques, instruction that biases the order of performing operations. And unlike running errands, there are no extraneous motivational factors (shall I go to the bakery first to get encouragement or last as a reward?) to bias the order.

Geometric analogies have been studied by several researchers interested in intelligence and analogical reasoning (Bethell-Fox, Lohman, & Snow, 1984; Mulholland, Pellegrino, & Glaser, 1980; Spearman, 1923; Sternberg, 1977; Whitely & Schneider, 1981). Using solution time as a dependent measure, Sternberg (1977) has isolated several solution stages, including (see also Spearman, 1923, for a discussion of the first, second, and fourth of these components) (a) *encoding* the terms of the analogy, (b) *inferring* the rules relating the two example terms (A and B), (c) *mapping* the rules relating the A and C terms, and (d) *applying* the rules identified at the inference stage: that is, performing the transformations on the C term to yield the answer, the D term. These global components of analogy solution have a prescribed order (e.g., a particular transformation must be inferred before it can be applied/performed), but the order of performing the individual transformations on the test term to produce an answer is unconstrained. Although problem solvers conceivably could perform multiple transformations in parallel, the work of Sternberg (1977) and Mulholland et al. (1980) suggests that transformations are performed serially during the application stage of solution. The optionality of transformation ordering then becomes important. In particular, one performance order may be more cognitively compelling than another. The order in which multiple transformations are performed on a single geometric figure may also have consequences for accuracy and solution time.

The research presented in this article will demonstrate that problem solvers have a preferred order for performing transformations during geometric analogy solution, despite the fact that the order is entirely optional. Three classes of explanations to account for this consistency will be considered. The first is based on general constraints of information processing, the second on task-specific considerations, and the third on individual differences. These classes of explanations are not mutually exclusive; thus, the task at hand is one of seeking evidence for each, rather than of deciding among them. For each class, the general claims will be described first, followed by the specific predictions tested experimentally.

General Information Processing Constraints: Working Memory

The solutions to many problems require the mental assembly and transformation of several different elements or processes. The greater the number of elements and/or processes, the greater the load on working memory. In general, there is a trade-

Portions of this research were presented at the annual meeting of the Psychonomic Society in San Diego, November 1983 and at the annual meeting of the American Psychological Association in Toronto, August 1984.

This research was supported by National Science Foundation Grant BHS-8002012 and National Institute of Mental Health Grant NH-34248 to Stanford University and by a National Science Foundation Pre-Doctoral Fellowship and a Stanford University Graduate Fellowship to Laura R. Novick.

Laura R. Novick is now at the University of California at Los Angeles.

Correspondence concerning this article should be addressed to Barbara Tversky, Department of Psychology, Building 420, Jordan Hall, Stanford University, Stanford, California 94305.

off between the amount of information that can be stored in working memory and the amount of processing that can be done (e.g., Case, 1978; Daneman & Carpenter, 1980; Newell & Simon, 1972). Working-memory load has been implicated in the performance of inductive reasoning tasks in general, and of analogies in particular (see Bethell-Fox et al., 1984, Mulholland et al., 1980, Sternberg, 1977, and Sternberg & Rifkin, 1979, for geometric and related analogies; Holzman, Pellegrino, & Glaser, 1982, 1983, for number analogies and number series; and Kotovsky & Simon, 1973, for letter series problems). These studies have shown that increasing the number and/or difficulty of the requisite processes increases demands on working memory and decreases solution success.

Working memory has proved to be an important theoretical construct in other domains as well, for example, reading (e.g., Daneman & Carpenter, 1980) and mental arithmetic (e.g., Hitch, 1978a, 1978b). Hitch's work on mental arithmetic is particularly relevant here because he has shown that intermediate results not immediately produced as responses are an added burden to working memory. Furthermore, he has suggested "that one feature of 'good' cognitive strategies is that they will minimize any deleterious effects of short-term forgetting on performance. It seems likely that this generalization will apply most clearly in tasks that are serial in nature, with a degree of option about the sequence of stages" (Hitch, 1978a, p. 337). The application of transformations in the solution of geometric analogies is a good example of such a situation.

One simple way geometric analogy solvers might reduce working-memory load is by performing transformations in decreasing order of difficulty. When two or more transformations are applied in series to a single geometric figure, the first transformation is applied to the externally presented (and internally represented) C term, but subsequent transformations are applied to the mentally represented products of the earlier transformations. More difficult transformations may take more time and/or more effort and therefore might benefit more from the externally represented figure than easier transformations. Alternatively, easier transformations might be more resistant to disruption than harder transformations when performed on a figure that is represented solely internally. Thus, for the general hypothesis that problem solvers order operations to reduce working-memory load, we will examine whether transformations are performed in order of decreasing difficulty. This is obviously not the only way of instantiating the working-memory hypothesis. However, it seemed that it would be a sensible and fairly easy strategy for subjects to adopt.

Task-Specific, Content-Dependent Considerations: Perceptual Processing

Persistent failures to find transfer between problems with similar solutions (e.g., Gick & Holyoak, 1983) and the lack of transparency of problem isomorphs (e.g., Hayes & Simon, 1977; Simon & Hayes, 1976) lend support to the idea that many important problem-solving strategies are content bound and task specific. In geometric analogy solution, perceptual factors may be important in determining the order in which transformations are initially inferred; some transformations may be more salient than others, and these may be detected earlier. One simple task-specific strategy would be to apply the transformations in the order in which they were identified. This strategy has the advantage of saving the step of deciding the transformation order separately for each problem.

A more complex task-specific constraint is suggested by the observation that constructing a geometric analogy solution typically entails constructing a mental image (Bethell-Fox et al., 1984). Many of the transformations used in geometric analogies have been studied separately in imagery tasks, in particular, mental rotation (Cooper & Shepard, 1973; Shepard & Cooper, 1982), mental size scaling (Besner & Coltheart, 1976; Bundeson & Larsen, 1975), and mental scanning (Kosslyn, Ball, & Reiser, 1978). Although theories of imagery posit that several transformations may be performed in sequence during image construction (Kosslyn, 1980; Kosslyn, Brunn, Cave, & Wallach, 1984; Shepard, 1984), imagery studies typically have examined transformations only singly, and the theories are moot with regard to the ordering of these transformations. Piaget and Inhelder (1956, 1971) have observed that construction of a mental image may bear resemblance to construction of a pictorial image. Thus, the order of performing transformations to mentally construct a geometric analogy solution may mirror the order in which different types of information about an object are needed to efficiently plan and execute a drawing of that object.

Individual Differences Considerations

Pellegrino and Glaser (1980) have suggested that individual differences may be expressed within the various components of analogy solution. With respect to transformation application, there are at least two possibilities. First, high- and low-ability subjects may adopt different performance orders that are determined by different constraints. For example, the order used by high-ability subjects might be determined by general constraints, such as working-memory load, whereas the order used by low-ability subjects might be determined by task-specific constraints, such as perceptual factors (see also Hunt, 1974). Consistent with this notion, Schiano, Cooper, and Glaser (1984) found that high-ability high school students tended to sort geometric analogies on the basis of the transformational relations among the figures, but that low-ability subjects generally sorted on the basis of the surface similarities of the figures. A second possibility is that high- and low-ability subjects might adopt essentially the same transformation order but differ in the consistency with which they follow that order. It is not unusual to find consistency of strategy use covarying with ability (see, e.g., Campione, Brown, & Ferrara, 1982; Sternberg, 1977).

Experimental Program

Four experiments will be reported, directed at both establishing and understanding the order of performing mental transformations in the solution of geometric analogies. In the first experiment, order was studied in two-transformation analogies. Throughout the research, it is assumed that transformations are performed in sequence or cascades, not in parallel. Although this assumption has empirical support (Mulholland et al., 1980; Sternberg, 1977), it does restrict the conclusions to problems for which the operations are not begun simultaneously. Determination of the order of transformation application relies, for the most part, on subjects' reports. The validity of these reports was checked in the second experiment, in which subjects were directed to solve analogies by performing the transformations in the preferred or unpreferred order, and speed and accuracy of solution were recorded. In the third experiment, the robustness of the performance order was tested using three-transformation problems. Finally, in the fourth experiment, order was examined in the construction of pictures.

Experiment 1

The first experiment was designed to determine whether there is a preferred order for performing transformations in the solution of geometric analogies, as well as to collect evidence bearing on the three classes of explanations. A consistent and transitive order of transformation application would be a strong finding because the task in no way constrains the order; problem solvers are free to perform the transformations in any order they choose because the order of application does not affect the identity of the final outcome. The transformation difficulty instantiation of the working-memory hypothesis was assessed by comparing the performance order to the order of the transformations in terms of difficulty. The hypothesis that transformations are applied in the order in which they are initially identified was tested by recording identification order. Finally, ability measures were collected to examine performance order as a function of ability.

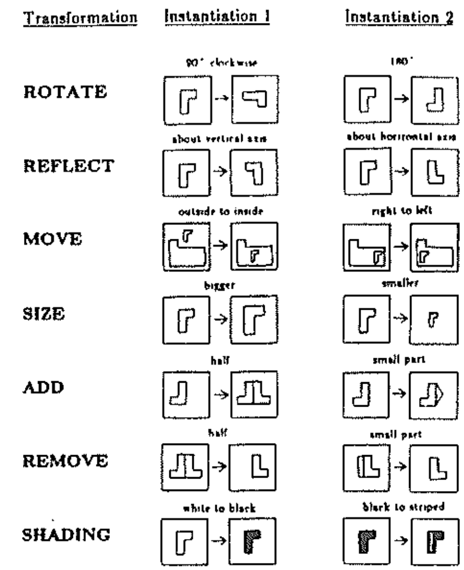


Figure 1 Examples of the two instantiations of each of the seven basic transformations.

Method

Subjects

The subjects were 98 Stanford University undergraduates who participated in partial fulfillment of course requirements. The data from 2 of these subjects were excluded from the analyses because those subjects failed to follow the instructions for the experimental task.

Stimuli

All of the analogies required two transformations to be performed on a single geometric figure. There were 21 problem types, which were chosen by constructing all possible pairs of the following seven basic transformations: *rotate, reflect, move, size, add, remove,* and *shading*. These transformations represent a variety of analog and discrete transformations that objects commonly undergo and that have previously been studied as component processes in cognitive tasks (e.g., Besner & Coltheart, 1976; Bundeson & Larsen, 1975; Cooper & Shepard, 1973; Shepard, 1975), and that are commonly found on psychometric analogy tests. Two different instantiations of each of these basic transformations were used in the analogies, as illustrated in Figure 1. Four analogies were constructed for each problem type by factorially combining the two instantiations of each transformation. For example, the four *rotate/size* problems were *rotate 90°/bigger, rotate 90°/smaller, rotate 180°/bigger,* and *rotate 180°/smaller*. Thus, there were 84 problems in all.

To construct analogies for which subjects would identify the intended transformations, there had to be two exceptions to the factorial combination rule. First, because a 180° rotation combined with a reflection looks identical to a single reflection about the axis perpendicular to the original reflection, the two rotate transformations used for the *rotate/reflect* problems were rotate 90° clockwise and rotate 90° counterclockwise. Second, because of difficulty in constructing an *add-half/remove-half* problem, an additional *add-part/remove-part* problem was used instead. To ensure generality across figures, many different geometric figures were used to construct the analogies.

Design

There were two conditions in this experiment: a *solution* condition ($N = 48$; 23 female and 25 male subjects) and an *identification* condition ($N = 48$; 20 female and 28 male subjects). In the solution condition, subjects saw the first three (A, B, and C) terms of an analogy and solved the problem by mentally constructing the fourth (D) term. The correct D term was presented on the other side of the page as one of five multiple-choice alternatives. Thus, subjects had to use the imaginal constructive matching solution strategy described by Bethell-Fox et al. (1984). The position of the correct answer was randomly chosen for each problem type, with the constraint that each position contained the correct answer approximately equally often. The position of the correct answer was held constant for all four analogies representing a given problem type. The four distractors for each analogy were constructed according to the following rules: (a) Transformation X alone applied to the third figure, (b) Transformation Y alone applied to the third figure, (c) Transformation X performed correctly and Transformation Y performed incorrectly, and (d) Transformation X performed incorrectly and Transformation Y performed correctly. A sample *move/add-half* problem appears in Figure 2. Subjects in the identification condition also saw the

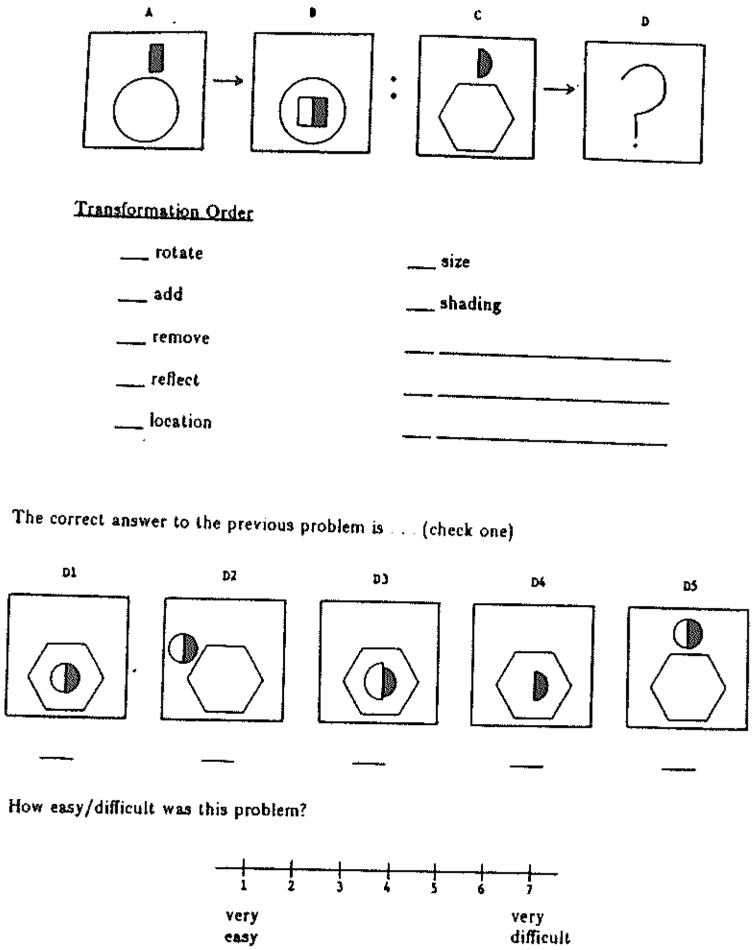


Figure 2. Example of a *move/add* problem from the solution condition, including the five alternative answers.

·st three terms of the analogies (see the top part of Figure 2) but simply ·d to identify the transformations that were used to change the A figure to the B figure.

Subjects completed 1 problem from each problem type for a total of problems. The problems were randomly chosen from among the 4 ·oblems for each problem type such that each problem was completed 12 subjects (hence the 48 subjects in each condition). Six different ·ders for presenting the problems were used, with 8 subjects receiving each order. The problem orders were random given the constraint that none of the seven basic transformations could appear in two consecutive problems.

Procedure

Subjects in both conditions began the experiment with a 24-item multiple-choice geometric analogy test, on which they were given 7 min to work. This test served two purposes. First, it gave subjects some practice solving geometric analogies before they had to solve the experimental problems, and, second, it served as a measure of analogical reasoning ability. The problems on this test were similar to those found on published psychometric tests. (A copy of the test may be obtained by writing to Laura R. Novick.)

After completing the analogy test, subjects were given detailed oral instructions that described the types of transformations they would encounter and the steps they were to follow in solving the problems. The instructions for the two conditions were identical until the experimenter began describing the specific task. The solution condition subjects' task was as follows: First, they were to identify the transformations used to change the A figure into the B figure. Second, they were to mentally apply those same transformations to the C figure to derive the correct answer. Third, they were to number the transformations in the order in which they had mentally performed them during the application stage. The seven basic transformations used in the experiment were printed below each analogy (see Figure 2). In addition, there were three blank lines at the end of the transformation list in case subjects used other transformations (a few subjects on a few problems actually used novel transformations; other subjects wrote descriptions of transformations that were functionally identical to the verbal labels we provided). Although subjects were not told that they had to perform the transformations sequentially, the instructions did imply that they do so. Thus, our results are necessarily restricted to situations in which transformations are performed sequentially (or at least to those in which one transformation is started before the other); although the data of Sternberg (1977) and Mulholland et al. (1980) suggest that this may typically be the case. To the extent that subjects performed transformations simultaneously, there should be little, if any, consistency in the ordering of transformations across subjects.

After numbering the transformations, subjects turned the page over and marked the answer they had constructed mentally. Subjects were instructed not to look back. Thus, if the answer they had constructed was not among the alternatives, they were to indicate their best guess. Finally, subjects rated the difficulty of the problem on a 7-point scale (1 = *very easy*; 7 = *very difficult*). This procedure was followed for all 21 problems in the booklet.

The identification condition subjects had only a single page per problem, which was identical to the first page given to the solution condition subjects. Their task was to identify the transformations used to change the A figure into the B figure and number those transformations in the order in which they noticed them. The identification condition subjects did not actually solve the analogies and could ignore the C figure printed on the problem page. However, it was important that the identification condition subjects see the same display as the solution condition subjects because the identification order was hypothesized to be determined by transformation salience, and the presence or absence of the C figure could affect the relative salience of the two transformations.

Results and Discussion

Solution Condition: The Order in Which Transformations Are Performed

The transformation ordering data were analyzed by using a nonparametric scaling procedure based on paired-comparison data (Carroll & Chang, 1964; Chang & Carroll, 1968). The data consisted of a 7 × 7 paired-comparison matrix for each subject, with the cell entries indicating whether the transformation in Row *i* was performed before or after the transformation in Column *j*. A cell entry was coded as missing if the subject marked an incorrect answer alternative or failed to number the correct transformations (inclusion of the former problems does not change the transformation ordering results). Because one of the four analogies from the *rotate/add* problem type was inadvertently misdrawn, the 12 subjects who received that problem were coded as having missing data for this pair of transformations.

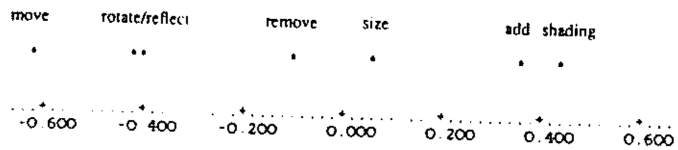
Finding the transformation order that best represents the paired-comparison data is a two-step procedure. In the first step, each subject's paired-comparison data matrix is processed separately. The program determines for each subject the number of times each transformation was performed before the other transformations. From this information, it determines the best overall ordering of the transformations for each subject and assigns each transformation a score that reflects the extent to which the subject performed that transformation first. Thus, the output from this step is a number-of-subjects by number-of-transformations matrix, with each row representing the transformation scores for a particular subject.

In the second step, this matrix is factored to yield two geometric configurations, one of transformations and one of subjects. Although the program is capable of representing objects in a multidimensional space, we specified a one-dimensional solution because it is most appropriate in the present context. In one dimension, the configuration of transformations represents the linear (metric) order that best captures the order in which subjects performed the transformations. The transformations are plotted as coordinates on a line to represent the metric properties of the solution. The configuration of subjects specifies, for each subject, the end of the linear transformation order that best corresponds to the beginning of that subject's order. To the extent that the transformations are performed in a consistent order across subjects, the subjects should cluster at one end of the transformation order.

The resulting transformation order is presented in Figure 3a. *Move*, on the far left, is performed first, *rotate* and *reflect* are next, then *remove*, then *size*, and, finally, *add* and *shading*. There are several ways to assess how well this solution fits the data. First, this order is the best representation of the transformation scores for 47 of the 48 subjects; the remaining subject tended to perform the transformations in exactly the opposite order (i.e., *shading* first and *move* last). The average Pearson correlation between the transformation scores of all pairs of subjects is .45. The average Spearman rank-order correlation is .44. This corresponds to a Kendall's coefficient of concordance (ω) of .45, $\chi^2(6, N = 48) = 129.12$, $p < .001$. As hypothesized, subjects agree on the order in which they perform mental transformations. If the subject whose order is the opposite of everyone else's is removed, the average rank-order correlation increases to .48, $\omega = .49$, $\chi^2(6, N = 47) = 139.25$, $p < .001$. These correlations are quite impressive considering that there are 5040 (7!) different ways subjects could order the transformations.

In the scaling analysis just reported, the data for each transformation pair were collapsed across the four problems representing that pair. To determine whether this was justified, we did follow-up chi-square tests for each transformation pair that explicitly examined performance order as a function of how the transformations were instantiated. For only 2 of the 21 problem

(a)



(b)

move > rotate, reflect > remove > add half, size > shading > add part

Figure 3. (a) Order (from left to right) in which mental transformations are performed in the solution of two-transformation geometric analogies as determined by the scaling analysis and (b) the revised performance order based on follow-up analyses.

types did the ordering of the transformations depend on the particular instantiations. For both the *size/add* and *add/shading* problem types, the order depended on the type of *add* transformation, $\chi^2(1, N = 39) = 14.04$, $p < .001$, and $\chi^2(1, N = 38) = 10.64$, $p < .01$, respectively. These dependencies can be incorporated into the Figure 3a order by putting *add half* next to *size* (those two transformations were not consistently ordered, $p > .11$ by a binomial test) and by putting *add part* after *shading* (and, of course, removing the original *add* transformation). The resulting priority order is shown in Figure 3b. Note that this transformation order is transitive (i.e., there is no triple of transformations for which subjects perform X before Y, Y before Z, and Z before X). This is a particularly strong finding given that no attempt was made to standardize the figures being transformed in the various problem types.

General Constraints on Problem Solving: Transformation Difficulty as a Predictor of the Performance Order

The particular working-memory hypothesis tested was that transformations are performed in decreasing order of difficulty. To assess the difficulty of each of the eight transformations (considering *add half* and *add part* separately), we computed the average number of errors across all problem types involving that transformation. The following order results: *rotate* (7.6 errors out of 48 = 15.8%) > *size* (13.1% errors) > *reflect* (11.0%) > *shading* (10.6%) > *add half* (10.4%) > *move* (7.7%) > *add part* (6.3%) > *remove* (4.2%). A comparison of this order and Figure 3b shows clearly that the relative difficulties of performing single transformations cannot account for the order in which multiple transformations are performed. The Spearman rank-order correlation between the two orders is .20, which is not significantly different from zero. *Move* is a particularly blatant offender. It results in relatively few errors, yet it is most likely to be applied first. *Remove* also presents a problem because it is in the middle of the performance order but results in the fewest number of errors. Thus, there is no evidence for the transformation difficulty version of the working-memory hypothesis.

Task-Specific Constraints: Identification Order as a Predictor of the Performance Order

The identification condition was designed to test the simple perceptual explanation that subjects perform the transformations in the order in which they identify them. This explanation carries with it the implicit assumption that the identification order is influenced primarily by perceptual factors such as transformation salience, although this assumption was not tested. We believe that transformation identification is determined to a large extent by transformation salience, but even if this is not the case, it could still be a reasonable strategy to use the identification order at the application stage of solution. The best-fitting identification order from the scaling program represents the data for only 32 of the 48 subjects (the reverse order is best for the remaining 16 subjects). The average intersubject correlation (Pearson or Spearman) is only .07, which represents a very small, but significant, degree of intersubject agreement, $\omega = .09$, $\chi^2(6, N = 48) = 26.59$, $p < .001$. Although the identification and performance orders are significantly correlated ($r_s = .87$, $p < .05$), suggesting that the order in which transformations are performed might depend on the order in which they are initially identified, this is not likely to be true for other reasons.

First, if subjects performed the transformations in the order in which they identified them, the agreement among subjects in the solution condition should be about the same as in the identification condition. However, it is six to seven times larger. Second, consistent with the hypothesis that identification order is affected by perceptual salience, some of the disagreement among subjects is due to the fact that although all subjects received the same pairs of transformations, the particular instantiations of the transformations differed. For example, *move* tended to be identified first in problems involving a small figure moving from outside a larger figure down inside that figure, but not in problems involving the movement of a small figure from the right side to the left side of the larger figure it is in. This pattern of results obtains when *move* is combined with *size*, *add*, or *shading*. Dependency of the identification order on the particular instantiations of the transformations was also observed for the *rotate/size* problem type. These dependencies were not found in the solution condition data. Thus, to the extent that there is any systematicity at all in the identification order, it tends to resemble the performance order; but some other mechanism must be accounting for the very high degree of consistency for the performance order.

Individual Differences Constraints: Ability as a Predictor of the Performance Order

Descriptive analysis of analogy measures. In addition to the 98 subjects in the experiment reported here, 30 subjects from an unreported rating task and 49 pilot subjects from the same population also completed the timed analogy test prior to the experimental task. The responses from all 177 subjects were included in the item analyses for the test so that more stable correlations could be obtained. Although all 24 items on the test correlate positively with overall performance, for the last item, total score correlates almost as highly with one of the distractors ($r = .124$) as with the correct answer ($r = .176$). Therefore, the responses from this item were removed and the analyses were redone on the 23-item test. The correlations between overall performance on the revised test and individual item scores range from .13 to .54, with a mean of .35. The Cronbach's alpha reliability of the 23-item test is .68 for these subjects. The item difficulties (i.e., the proportion of subjects solving each item correctly) range from .18 to .99, with a mean of .82 and a median of .90. Performance on the analogy test does not differ across conditions, $t(94) = 0.49$, $p > .62$, with an overall mean of 18.7 out of 23 (the scores range from 9 to 23).

We also computed number correct scores for the experimental tasks. In the identification condition, a problem was counted as correct if the transformations that were identified would correctly change the A figure into the B figure. The mean number correct was 17.3 out of 21, with a range of 12 to 21. Solution condition subjects had to mark the appropriate transformations and choose the correct D term. The mean for this condition was 18.8, with a range of 14 to 21. Both identifying and applying transformations are significant sources of individual differences in analogy ability (see also Sternberg, 1977), as the scores on the experimental task for both conditions correlate with performance on the timed analogy test ($r = .43$, $p < .003$ for the solution condition and $r = .35$, $p < .02$ for the identification condition). Although the .43 correlation for the solution condition may seem low given that both tasks involve solving geometric analogies, because the experimental analogies were untimed, there is a restricted range of scores for that task (particularly above the mean because subjects averaged 90% correct). This would tend to attenuate the correlation between the two measures.

Ability hypotheses. This class of hypotheses states that the order in which transformations are performed covaries with ability. As indicated earlier, this general hypothesis can be instantiated in at least two ways. First, high-ability subjects might perform the transformations in a different order than do low-ability subjects. Alternatively, highs and lows generally might perform the transformations in a similar order, but highs might be more consistent in their ordering than lows. To examine these hypotheses, we ran the scaling program on the data from high- and lower ability subjects separately. The 15 subjects in approximately the top third (31%; scores of 21–23) of the analogy test distribution were considered high in analogy ability. The 13 subjects in approximately the bottom third (27%; scores of 12–17) were lower in analogy ability.

The solutions for the two ability groups yield very similar application orders ($r_s = .93$, $p < .02$), which in turn are almost identical to the overall order shown in Figure 3a. The high-ability order is the same as the Figure 3a order except that *add* and *shading* are reversed. The lower ability order is the same except that *rotate* and *reflect* are reversed. The two groups also do not differ in terms of consistency. The average rank-order correlation of the transformation scores among the high-ability subjects is .53, $\omega = .56$, $\chi^2(6, N = 15) = 50.83$, $p < .001$. For the lower ability subjects, the average rank-order correlation is .49, $\omega = .53$, $\chi^2(6, N = 13) = 41.33$, $p < .001$. Contrary to the ability hypothesis, both high- and lower ability subjects are very consistent in their ordering of the transformations.

Thus, in the ability range represented, individual differences do not appear to be related either to the order in which transformations are performed or to the consistency with which the overall preferred transformation order is followed. However, ability differences may appear with more difficult problems and/or a wider range of ability, a possibility examined in Experiment 3.

Experiment 2

Problem solvers clearly have a preferred order for applying mental transformations in the solution of geometric analogies, and this order is consistent across subjects. Does using this order, as opposed to the opposite order, have consequences for task performance? In this experiment, subjects were told the order in which to perform the transformations for each problem, and solution times were recorded. If the order of transformation application reported by subjects in the first experiment is accurate and is determined by cognitive constraints, performance should be faster and better when subjects apply transformations in the preferred order. Thus, this experiment provides important corroboration of the self-report data from Experiment 1. If subjects' reports about the order in which they performed the transformations are not accurate or if the consistency of the reported order reflects some other factor unrelated to the actual order or to any cognitive constraints on that order, then no differences attributable to transformation ordering should be observed in this experiment.

Method

Subjects

The subjects were 32 Stanford University undergraduates (14 female and 18 male subjects) who participated in partial fulfillment of course requirements.

Stimuli

The stimuli were the 84 analogies used in Experiment 1. The 5 alternative answers were reordered for 3 of the 4 problems from each of the 21 problem types so that the correct answer would be in a different position for each of the 4 problems. To indicate the order in which subjects were to perform the transformations, two one-word transformation labels were printed at the top of the problem. For example, if *reflect, add* were printed above the analogy, subjects were supposed to perform the *reflect* transformation first and the *add* transformation second. All of the stimuli were presented on slides.

Design

Two sets of the 84 problems were constructed. They differed only in the order in which subjects were told to perform the transformations. The Set I stimuli are described below; the transformation labels were reversed for each problem in order to get the Set II stimuli. First, two problems from each pair of transformations were assigned to one transformation order and two were assigned to the other order such that all four instantiations of the two transformations would occur in each order for each pair of transformations. For example, in Set I, subjects performed the transformations for the four *move/size* problems in the following orders: *move top–bottom, smaller; move right–left, bigger; bigger, move top–bottom;* and *smaller, move right–left.*

The 84 problems were randomly assigned to four blocks of 21 according to the following four constraints: (a) Each block should contain one problem from each pair of transformations, (b) each transformation should occur approximately equally often first and second in each block, (c) in each block the transformations should be performed in the preferred and unpreferred orders (on the basis of the Experiment 1 results) for approximately equal numbers of problems, and (d) the correct answer should appear approximately equally often in each of the five positions. Four of the random orders for presenting the problems in Experiment 1 were chosen to be used with the four blocks of analogies in this experiment. The four blocks of problems were presented to subjects in four different orders according to a Latin square design.

Procedure

Before beginning the experiment, subjects were familiarized with the two instantiations of each transformation so that they would be able to understand the transformation labels on the problems. The experimenter stressed the importance of performing the transformations in the order indicated for each problem. Subjects were told that the order in which the transformations were presented would vary: Sometimes it would be the same as the order in which they would have chosen to perform the transformations and sometimes it would not, but they were always to perform the transformations in the order specified. Subjects completed 14 practice problems to familiarize themselves with the task. Slide presentation and data collection were performed by an Apple II Plus microcomputer. The slides were projected onto the wall about 6 ft (1.8 m) in front of the subject. Subjects responded by pressing one of six buttons on a response panel that rested on a small table in front of them. Subjects were allowed to use either hand to make their responses as long as they used the same hand throughout the experiment.

The following procedure was followed for all problems: First, a fixation point appeared, indicating the start of a problem. It remained in view until the subject pressed the button on the far right, at which point it was replaced with an analogy after a delay of 1 s. This slide showed the A, B, and C terms of an analogy and contained a question mark in the D box (as in the top part of Figure 2). It remained in view for a maximum of 30 s. Subjects were to mentally solve the analogy and then press the button on the far right again when they figured out the answer. Accuracy rather than speed was stressed. The time from onset of the analogy to the button press was taken as a measure of solution time. Subjects failed to respond within the allotted time on an average of only 1 problem per subject (53% of the missing times were from *rotate/reflect* analogies). The analogy slide was replaced by the alternatives slide 750 ms after the button press (or after the 30-s time limit if the subject failed to respond). When subjects found the answer they had mentally constructed, they responded by pressing one of the five buttons on the left, which were numbered *1* to *5*. The response to this slide caused a rating question to be presented after a 750-ms delay. Subjects were asked to rate on a 5-point scale how easy or difficult it was to perform the transformations in the order indicated. The scale went from 1 (*fairly easy*) to 5 (*fairly difficult*); these data were not informative because of a floor effect and thus will not be discussed. Subjects responded by pressing one of the numbered buttons. The fixation point then reappeared, indicating the start of the next problem.

There was a short break between each of the four blocks of slides, during which the experimenter changed slide trays. Subjects were given a longer break of several minutes between the second and third blocks, that is, about halfway through the experiment. Subjects were tested individually in 1-hr sessions.

Results and Discussion

Correlations With Experiment 1

In all analyses, solution times for problems on which subjects failed to indicate the correct answer were excluded. No subject was missing data for all four analogies from a given problem type. However, for a few subjects and problem types, this procedure did result in missing data for one of the two transformation orders. To facilitate analyses, missing data for one transformation order were replaced with the data from the other transformation order for that problem type. For example, if a subject's *rotate, move* time was missing, it was replaced with the *move, rotate* time. This procedure works against finding the predicted difference between transformation orders.

In order to compare results across Experiments 1 and 2, it is important to verify that the problems behaved similarly in the two experiments. That is, the order of the 21 problem types in terms of difficulty should be similar in the two experiments. We computed the Spearman rank-order correlation between solution times and number correct in this experiment with problem difficulty ratings and number correct in Experiment 1 (these data are shown in Table 1). The time and accuracy data from the present experiment were collapsed across the preferred and unpreferred transformation orders for each problem type. The analyses presented in the next section justify this analysis decision; they indicate that the effect of transformation ordering on both time and accuracy is constant across problem types. Both measures from the current experiment are highly correlated ($r_s = -.68, p < .001$).

The cross-experiment correlations clearly show that there is a stable ordering of transformation pairs. Problems judged to be more difficult in Experiment 1 were correctly solved less often in Experiment 2 ($r_s = -.67, p < .001$) and required more time to solve ($r_s = .94, p < .001$). Similarly, accuracy scores in the two experiments are very highly correlated ($r_s = .79, p < .001$). The only nonsignificant correlation is for accuracy in Experiment 1 with solution times from the present experiment ($r_s = -.27, p > .24$).

Table 1
Number Correct, Time, and Ratings for Each Transformation Pair on the Basis of the Data From Experiments 1 and 2 (Ordered by Decreasing Solution Time)

	Experiment 1		Experiment 2	
Transformation pair	Number correct (out of 48)	Difficulty rating (7-pt. scale)	Solution time (ms)	Number correct (out of 128)
Rotate/reflect	36	4.0	17,327	96
Move/rotate	41	3.2	12,616	112
Rotate/shading	44	3.0	12,530	116
Rotate/add	40	3.5	12,143	110
Move/reflect	41	2.9	11,447	114
Reflect/remove	47	2.6	11,226	123
Reflect/add	45	2.9	11,123	120
Rotate/size	40	2.8	10,774	115
Reflect/shading	44	2.8	10,453	116
Reflect/size	41	2.6	10,241	116
Remove/add	48	2.7	10,191	125
Rotate/remove	42	2.5	10,187	117
Move/remove	45	2.5	9,044	122
Add/shading	41	2.7	8,894	118
Move/add	48	2.6	8,479	126
Size/add	41	2.3	7,623	117
Move/size	43	2.3	7,344	119
Remove/shading	45	2.5	7,269	121
Move/shading	44	2.0	7,134	122
Remove/size	47	2.4	6,922	123
Size/shading	39	2.0	6,188	121

Consequences of Transformation Order

Solution times. Analyses of the consequences of transformation ordering were performed using only those pairs of transformations that were consistently ordered in Experiment 1. Each subject solved analogies using both transformation orders for each of these problem types. However, subjects solved any given analogy using only one order. That is, any particular analogy was solved in one order by subjects who received Stimulus Set I and in the opposite order by subjects who received Stimulus Set II. Because the effect of transformation order is expected to be small relative to other factors known to influence problem solution (e.g., encoding and inference), and because these other factors are very likely to make different problems involving the same transformations differ in difficulty, it is important to compare identical problems solved by subjects using different transformation orders. For example, it might take longer to perform Transformations X and Y (regardless of order) on a complex figure than on a simple figure. Comparing solution times for different transformation orders when the figures differ in this way would not constitute an adequate test of the effect of transformation ordering. Thus, although subjects performed all pairs of transformations in both orders, transformation order must be considered a between-subjects variable in the analyses. For a given analogy, different transformation orders are represented by different stimulus sets.

One further complication is needed. Subjects solved two of the four analogies for a given problem type using one transformation order and two using the other transformation order. For example, subjects who received Stimulus Set I solved *reflect/add* problems 1 and 2 using the preferred order and *reflect/add* problems 3 and 4 using the unpreferred order. In contrast, subjects who received Stimulus Set II solved problems 1 and 2 using the unpreferred order and problems 3 and 4 using the preferred order. It is clear that if we refer to problems 1 and 2 as Problem Set α and problems 3 and 4 as Problem Set β, then the predicted effect of transformation order appears as a Problem Set × Stimulus Set interaction: For Problem Set α, Stimulus Set I times should be shorter than Stimulus Set II times, whereas for Problem Set β, Stimulus Set II times should be shorter.

The data were analyzed by a 2 (stimulus set) × 2 (problem set) × 18 (problem type) analysis of variance (ANOVA). Significance levels for effects that include the subjects factor in the error term (i.e., certain main effects and interactions involving the within-subjects factors: problem set and problem type) are based on the Geisser-Greenhouse correction. There is a main effect of problem type, $F(17, 510) = 32.88, p < .001$. Solution times range from about 6.0 s (*size/shading* problems) to 12.5 s (*move/rotate* problems), with a mean of 9.5 s. The main effect of problem set is also significant, $F(1, 30) = 9.63, p < .005$, as is the Problem Set × Problem Type interaction, $F(17, 510) = 6.85, p < .001$. These results confirm the a priori prediction that the different problems in the two problem sets would not be equivalent in difficulty. The three-way interaction of Problem Type × Problem Set × Stimulus Set fails to reach significance, $F(17, 510) = 1.50, p > .15$, as does the main effect of stimulus set, $F(1, 30) < 1$, and the Stimulus Set × Problem Type interaction, $F(17, 510) = 1.15, p > .33$. Thus, we are free to examine the interaction of interest.