Memory & Cognition
1986, 14 (3), 270-280

Problem representation: The effects of spatial arrays

CAROL McGUINNESS
The Queen's University of Belfast, Belfast, Northern Ireland

Two studies examined how characteristics of spatial arrays contribute to efficient problem representation. Thirty-two adults were presented with information about family relationships in one of two arrays: a hierarchy or a matrix. Their answers to two different sets of questions were timed. The matrix format was superior to the hierarchy for one set of questions only; no differences between the arrays emerged for the other set. The data were interpreted in terms of how the family relationships were mapped onto the arrays; the mapping differences between the arrays affected the number of mental steps needed to solve some questions, but not others. In a second experiment ($N=32$), the same problem information was remapped onto the arrays, with the mapping relations reversed. As predicted, the pattern of response times exactly reflected the change in mapping. It is proposed that, for spatial arrays, efficiency of problem representation is best understood in terms of the number of mental steps in the problem solution.

Theories of problem solving derived from computer simulations have emphasized two important concepts, problem representation and heuristic search. Earlier models (Newell & Simon, 1972) concentrated on heuristic search, and a series of experimental studies examined search processes in well-defined puzzle-type problems (e.g, Jeffries, Polson, Razran, & Atwood, 1977; Simon, 1975; Thomas, 1974). More recent models have focused on problem representation. Hayes and Simon's (1974) program, Understand, mimics how a problem representation is constructed from text and isolates the characteristics of the text which determine the representation. Greeno (1973, 1977) also provided an analysis of how the organization of memory can affect problem representation. This paper addresses the question of problem representation and how it can have important effects on the ease with which problems are solved.

An important experimental strategy for studying problem representation is to use variants of the same problem, all of which describe isomorphs of a single problem. Such studies suggest that isomorphic problems are experienced at different levels of difficulty, which can be attributed to differences in representation (Simon & Hayes, 1976). Other evidence that problem isomorphs yield different representations comes from studies on the transfer of training. Reed, Ernst, and Banerji (1974) showed that solving one problem variant successfully did not necessarily transfer to another variant, indicating that successful solutions may be specific to precise representations. On the other hand, Gick and Holyoak (1980, 1983) reported that if the problem representation was sufficiently abstract or schematic, successful transfer did occur.

Another experimental strategy for examining the effects of problem representation is to restrict observations to a single problem, but to choose one which can be mentally represented in various ways. Problems studied using this strategy tend to be more complex and/or less well-defined than the problems in the isomorph studies just described. Schwartz (1971) and Schwartz and Fattaleh (1972) demonstrated that a matrix representation was superior to other formats (e.g., tree diagrams, sentence rewrites) for solving who-done-it-type deductive reasoning problems and that this superiority became more crucial as problem size increased (Polich & Schwartz, 1974). They suggested that the matrix was superior because it allowed the greatest number of relations to be correctly and simultaneously deduced. Mayer (1976) showed how the organization of a problem statement affected performance through problem representation. Subjects had to answer questions about the outcome of a tournament of games. The problem statement (about the teams which would play one another depending on the outcomes of previous games and the prizes that they would win) was presented to the subjects at different levels of organization, which was varied by using "go to" list structures in different ways or by stating explicit outcomes using an example. The problem statements were also presented either as a series of verbal statements or as a flowchart. Mayer showed that the more integrated the representation, the better the subject's performance on a problem. The flowchart mode of presentation was more helpful than the verbal mode, but only at the poorer levels of organization. Mayer proposed that the degree of structural integration of the problem statement, whether verbal or diagrammatic, was the important factor in the mental representation. Carroll, Thomas,

These studies are part of a PhD thesis submitted to the Queen's University of Belfast, Northern Ireland. The help of W. P. Brown is gratefully acknowledged. Thanks are also extended to G. B. Greer and Ken Brown. Requests for reprints should be addressed to Carol McGuinness, Department of Psychology, The Queen's University, Belfast BT7 1NN, Northern Ireland.

and Malhotra (1980), using isomorphs of an ill-defined design problem, found that the problem was more easily solved when the variant had a spatial theme (about a layout for a business office) than when the variant had a temporal theme (about a schedule for stages in a manufacturing process). The differences in performance were mitigated when the subjects were trained to use a graphic method for representing the problem statement. The graphic representation, which involved using a matrix, appeared to help the solver to order items, to make relations explicit, and to keep track of what information had already been dealt with.

From these studies, certain conclusions can be reached about problem representation. First, people do not spontaneously use the most efficient representation for a problem. Second, as the complexity of problems increases, producing efficient representations becomes more important for successful problem solutions. Third, efficient representations all appear to have a special quality which has been variously referred to as connectedness (Mayer & Greeno, 1972), structural integration (Mayer, 1976), and coherence (Greeno, 1977). Finally, spatial arrays are often, though not necessarily, associated with this special quality.

The experiments reported here attempt to determine the characteristics which make spatial arrays more or less connected, integrated, or coherent. Experimental tasks about family relationships were designed which involved learning a large amount of information about family members and then drawing inferences about them according to certain rules. In order to do the problems, the information had to be organized and systematized. The information was mapped onto two different spatial arrays, a matrix and a hierarchy, and the characteristics of the mapping were considered important for reasoning about the relationships. Figure 1 shows how the family relations are mapped onto the hierarchy (full details of the problems will be given later), and Figure 2 shows the mapping for the matrix. Two dimensions of the family relationships are spatially described, the birth order of the parents and birth order of the children. On the hierarchy, it is easy to check that John is Billy's father or that Tom is Jane's brother. Similarly, it is easy to check these relationships on the matrix. If, however, you want to find the answer to the question, "Who are the oldest children in each family?" you can easily check from the matrix that the answer is Patricia, Peter, Sheila, and Margaret. Answering the same question from the hierarchy is more indirect because it necessitates passing over intervening family members to get from Patricia to Peter to Sheila to Margaret. These simple mapping differences make relationships more or less explicit and may have important performance effects when drawing sequences of inferences. The purpose of this paper is to identify the occasions when the differences in mapping between the two arrays affect reasoning performance and when they do not, and to explain why.

EXPERIMENT 1

Method

Problem materials and subjects. A complex problem was devised about the relationships between a set of 20 family members. The relationships (called the knowledge base) were mapped onto two spatial arrays, a hierarchy and a matrix (Figures 1 and 2). A set of questions was devised which concentrated on the cousin relationships and posed problems about who could go on holiday with whom (holidays questions). These questions were restricted to the children, and there were rules which limited the choice of holiday companion. Two sets of such rules were devised. The sex rule specified whether a child could go on holiday wth someone of the same or different sex. The birth-order rule specified whether a child could choose somebody of a younger or older birth order. A rule was coded as SS OY: SS means same sex (if it is a boy wishing to go on holiday, he must go with a boy; if it is a girl, she must go with a girl). OY means that the person must first choose within his/her

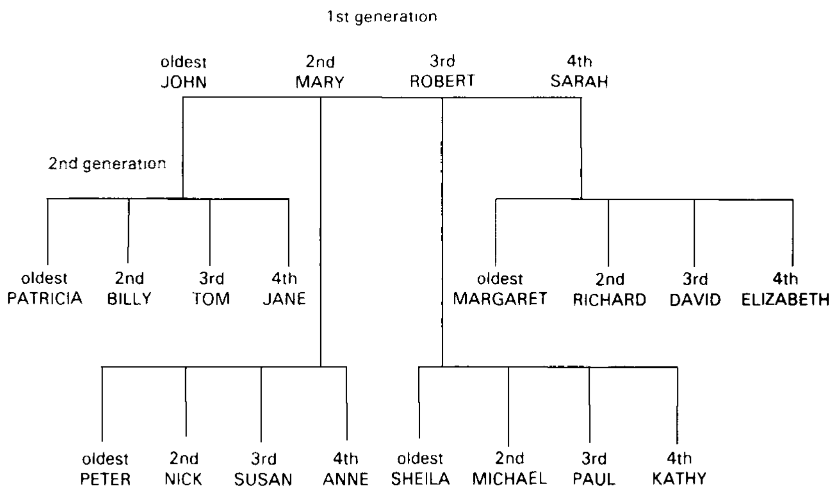

Figure 1. The family relationships arrayed as a hierarchy in Experiment 1. (The version presented to the subjects arranged the four families in full linear sequence from left to right across the page.)

2nd generation	1st generation: oldest JOHN	2nd MARY	3rd ROBERT	4th SARAH
oldest	PATRICIA	PETER	SHEILA	MARGARET
2nd	BILLY	NICK	MICHAEL	RICHARD
3rd	TOM	SUSAN	PAUL	DAVID
4th	JANE	ANNE	KATHY	ELIZABETH

Figure 2. The family relationships arrayed as a matrix in Experiment 1.

own birth order, but always a child of a younger parent; if there is nobody available, he/she can proceed to the next birth order, beginning with the child of the oldest parent and so on. There were four such rules, coded SS OY (same sex, oldest to youngest), DS OY (different sex, oldest to youngest), SS YO (same sex, youngest to oldest), and DS YO (different sex, youngest to oldest). A number of hypothetical situations were then set up in which various family members did not wish to go on holiday at all; they were excluded from the rule. A question asked, "SS OY (Jane Anne) ?Susan." To answer this question, the person had to find Susan's holiday companion given that the SS OY rule was operating and that Jane and Anne did not wish to be considered. The number of family members who were excluded from the rule in any one question varied from one to four. There were 16 holidays questions; each rule combination (e.g., DS YO) was used four times, once for each of the numbers of exclusions. Due to constraints inherent in the rules, not all second-generation family members could be used as targets in the questions. For this question set, four names were not used, eight names were used once, and four names were used twice.

The second set of questions, inheritance questions, concentrated on the within-family relationships and posed problems about who could inherit whose house. The children owned houses which they could inherit from one another. The rules of inheritance varied according to sex; sometimes males inherited before females, and other times females inherited before males. There was also a rule about birth order; sometimes older children inherited before younger ones, and other times younger children inherited before older ones. A rule was coded MI OY: MI means that the males of males inherit before males of females (females do not inherit under this rule). OY means that older boys inherit before younger boys. There were four such rules, MI OY (male inheritance, oldest to youngest), FI OY (female inheritance, oldest to youngest), MI YO (male inheritance, youngest to oldest), and FI YO (female inheritance, youngest to oldest). A number of problems were then posed in which some family members were disinherited; they were excluded from the rule. A question asked, "MI OY (Michael Paul) ?Tom." To answer, a person had to find out who came next in the line of inheritance for Tom's house, given that the MI OY rule was operating and that Michael and Paul were disinherited. There were 16 inheritance questions, and the number of people disinherited in any one question varied from one to four. As for the holidays questions, each rule combination was used four times, once with each of the four numbers of exclusions. Eight second-generation family members were used twice as targets for the 16 inheritance questions.

Thirty-two students from the Queen's University of Belfast, between 18 and 25 years of age, volunteered to solve the problems. Sixteen students (8 males and 8 females) were randomly assigned to either the hierarchy or matrix learning condition. All students answered both sets of questions.

Design. Six variables were manipulated in a mixed-factors design. The main between-subjects variable was array (hierarchy and matrix). The other between-subjects factors were sex of subject and presentation order of question set (inheritance questions first or holidays questions first). Three features of the questions were within-subject variables: the number of exclusions (1-4), the direction of the rule (OY, YO), and the sex rule (SS/DS or MI/FI). The order of questions within a set was counterbalanced. Counterbalancing was achieved by organizing the 16 questions into four blocks. Each block had one instance of each of the four rule combinations and one instance of each of the four possible numbers of exclusions. The order of questions within a block and the order of the blocks were counterbalanced according to a Latin square design. Each question was presented once. Time to answer was the dependent measure. The data for the inheritance questions and the holidays questions were analyzed separately.

Procedure. Subjects were tested individually. A testing session lasted about 2 h and was divided into two parts, learning the arrays and answering the questions. Students were asked to memorize either the hierarchy or matrix version of the knowledge base. They studied it for 20 sec per trial and then attempted to complete a blank array after each trial until they did so successfully for two successive trials. Subjects took a mean of 7.2 trials to learn the hierarchy and 7.9 trials to learn the matrix; there were no significant differences between the two learning conditions. The students then learned the first set of rules, appropriate to their condition, and answered eight practice questions. Sixteen test questions were presented in one of four presentation orders. The questions were answered from memory, and the arrays were not available for prompting. Each section of a question (rule, exclusions, target) was projected on a separate slide. The students self-paced the presentation of the slides, and three response times were recorded. Time 1 was the interval spent processing the rule, Time 2 was the interval spent on the exclusions, and Time 3 was spent processing the target and producing the answer. The procedure was identical for both question sets, whose order of presentation was counterbalanced.

After each session, the students were asked how they answered the questions. It had been anticipated that during the three-slide statement of the question, the students would (a) read and interpret the rule and the number of exclusions, and (b) integrate these details with the array to find the answer. Interest was centered on the time to perform (b) rather than (a), and it was expected that Time 3 would be the best approximation to the desired measure. Initial inspection of the data, together with the students' own reports of their strategies, showed that subjects differed somewhat in their processing of the number of exclusions. One group, the memorizers ($N=15$), performed in the anticipated way; they interpreted the rule, memorized the list of exclusions, and finally began to search the array. Their mean values for Time 2 and Time 3 were 9.2 and 32.2 sec, respectively. For the second group of subjects, the searchers ($N=17$), the mean values for Time 2 and Time 3 showed a different pattern; 19.4 and 23.5 sec, respectively. It was clear from their strategy reports that the searchers were not just slow readers and fast processors of the array. Rather, while reading the list of exclusions, they began searching through the array to locate the exclusions, and so they completed part of the search for an answer during Time 2. Although Time 3 is a good estimate of processing time for the memorizers, it clearly underestimates processing time for the searchers. Consequently, instead of using Time 3 alone, I decided to combine Times 2 and 3 as the measure for analysis. Although the combined time is a less than ideal measure because it included more reading time than was intended, it is considered an acceptable measure for the reasons just outlined. (All mean values quoted in this section are based on raw-response-in-seconds times before adjustments were completed. All mean values quoted for analyses of variances are geometric means in seconds after the adjustments were made.)

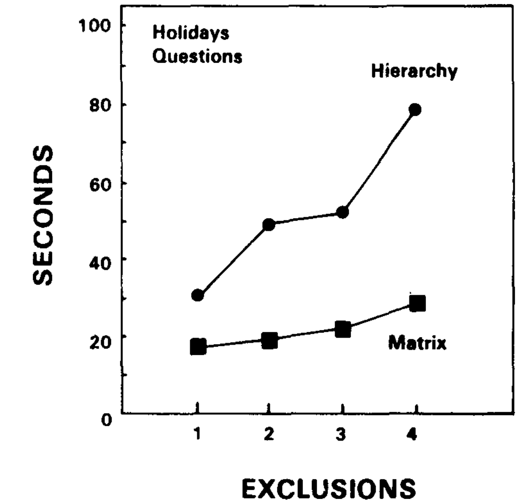

Figure 3. Experiment 1/holidays questions: mean response times to questions with different numbers of exclusions for the two spatial arrays.

Results

The percentage error was 15.1% for the holidays questions/hierarchy and 12.1% for the holidays questions/matrix. The comparable percentages for the inheritance questions were 12.5% and 16%, respectively. Errors were randomly distributed across questions, and there was no indication of a speed/accuracy trade-off. A number of adjustments were carried out on the raw data. Error times were substituted with an average of two estimates (the mean correct response times, RTs, for the appropriate row and column). The data were logarithmically transformed, as they were positively skewed and variances were not homogenous (Winer, 1970). The magnitude of practice effects was statistically ascertained by finding the regression line of RTs on position of question in the sequence of presentation. Each data set was adjusted using the appropriate beta value. The mean values quoted for all conditions are geometric means in seconds.

Holidays questions. The analysis of variance shows that array had a dramatic effect on the time it took students to answer the questions [$F(1,24) = 61.05$, $p < .001$]. Those who learned the hierarchy took more than twice as long (mean = 49.50 sec) to answer than those who learned the matrix (mean = 22.13 sec). Retrieving information about the cousin relationship was substantially easier from the matrix than from the hierarchy. Other variables also affected RTs. As the number of exclusions from the rule increased from one to four, RTs also increased: the respective means were 21.73, 30.27, 34.51, and 52.24 sec. There was a significant interaction between number of exclusions and array [$F(3,72) = 129.13$, $p < .001$] (see Figure 3). It is clear that questions that required more processing became increasingly more difficult on the hierarchy. The effect of number of exclusions occurred for all subconditions except for those DS questions with two and three exclusions, where RTs did not increase [$F(3,72) = 9.59$, $p < .01$]. The direction of the rule (OY/YO) affected RTs [$F(1,24) = 7.92$, $p < .01$]. Inferring about family members from oldest to youngest (mean = 31.84 sec) was easier than making inferences from youngest to oldest (mean = 34.20 sec), and this was true irrespective of the type of array learned. Finding a holiday companion of the same sex was significantly easier (mean = 30.76 sec) than finding one of the opposite sex (mean = 35.60 sec) [$F(1,24) = 37.86$, $p < .001$]. This sex-rule effect was modified by a three-way interaction [$F(3,72) = 4.23$, $p < .01$], which suggests that it was less true for questions with three and four exclusions on the matrix. (Appendix A shows mean values for subconditions.)

Inheritance questions. For this question set, array did not significantly affect RTs [$F(1,24) = .11$]; those who learned the hierarchy took virtually the same time to answer (mean = 31.77 sec) as those who learned the matrix (mean = 32.66 sec). Consistent with results for the holidays questions, increasing the number of exclusions increased the time to answer [$F(3,72) = 89.40$, $p < .001$]; the means for one, two, three, and four exclusions are 21.98, 25.76, 39.81, and 47.10 sec, respectively. The pattern is almost identical for both arrays. On this occasion, the mapping differences between the hierarchy and matrix did not affect performance. No other feature of the questions (OY/YO, MI/FI) directly affected RTs, although a three-way interaction between presentation order, array, and number of exclusions [$F(3,72) = 4.45$, $p < .01$] shows a restricted practice effect; when the inheritance questions came second, easier questions (with one exclusion) were answered faster by those in the hierarchy condition. (Appendix A shows mean values for subconditions.)

Discussion

The pattern of RTs to the two sets of questions clearly shows that array had a differential effect on the processing required to answer one set of questions, but not the other set. Holidays questions are easier on the matrix than on the hierarchy; inheritance questions are equally difficult on both arrays. Why?

The questions about going on holidays involved the cousin relationships, which were directly mapped onto the matrix, but not onto the hierarchy. The students seemed to make use of this mapping to facilitate processing, even though the same basic information was available from both arrays. The differential mapping was not true for the within-family relationships which were central to the inheritance questions; hence no effect of array was observed. There were superficial differences in the way the relationship was mapped—vertically on the matrix and horizontally on the hierarchy—but they did not affect performance.

The format of the spatial arrays is not the only variable that influenced RTs. Increasing the number of exclusions from one to four made the task more difficult for both sets of questions, presumably because more elements in the arrays had to be processed. And, significantly, the rate of increase in difficulty level was greater on the less efficient array (i.e., holidays questions on the hierarchy). It seems that as questions became more difficult, efficient representation became more important. The sex and birth-order rules had differential effects on the two question sets; both affected RTs to the holidays questions but not to the inheritance questions. The sex-rule effect might be accounted for by an appeal to the distinction between matching and mismatching responses. SS, MI, and FI rules all demanded that the sex of the person about whom the question was posed be the same as the sex of the person who constituted the correct answer; for DS rules, the sexes were different. Monitoring different sex elements may have added an additional processing load. The effect of the birth-order rule may be attributed to some reversal of the normal reading pattern; oldest to youngest was arrayed from left to right, whereas the opposite was true for youngest to oldest. Unfortunately, it is not clear why the effect occurred for one set of questions only.

The purpose of varying the rules, of course, had not been for their theoretical significance, but rather to generate sizable and varied question sets. Their real interest lies in how they interacted with the arrays. Although birth-order and sex rules showed interactions with array in a limited number of conditions, the most consistent interaction was between array and number of exclusions. Extra exclusions generally made the questions more difficult, and even more difficult on the less efficient representation. How did the students process the information to produce these effects? At the end of testing, the students described that they did not hold the total array in working memory at any one time (capacity limitations); they processed relevant parts of it and then moved on to the next part. They reported "stepping through" the array, examining each element as the rule prescribed, "skipping over" or "crossing out" elements which were irrelevant (i.e., not allowed by the general rule, or excluded on that occasion). Although the descriptions were informal, they did provide a clue as to how to quantify the difficulty level of the question, and thus accounted for the effects of array, number of exclusions, and their interaction. The notion of a mental step was taken as a basic unit; mental steps were counted for each question on each array. (The step unit was necessarily crude because it could not be assumed that all steps were of equal weight, but it did prove to be remarkably predictive.) The step count assumes that the problem solver processes the array by mentally stepping from one element to the next, in the direction dictated by the rule.

Here is an example of how steps are counted for the holidays question which was described earlier: SS OY (Jane Anne) ?Susan. Ideally, Susan is looking for the third oldest female of a parent who is younger than her own; if this is not possible, she must choose a younger child, beginning her search with the youngest female of the oldest parent. The step count assumes that she proceeds as follows:

Find Susan: female, third oldest, proceed OY	Step 1
To: Paul, male, irrelevant	Step 2
To: David, male, irrelevant	Step 3
To: Jane, female, youngest, excluded	Step 4
To: Anne, female, youngest, excluded	Step 5
To: Kathy, female, youngest, OK	Step 6

To reach the correct answer, Kathy, six mental steps on the matrix must be taken. Answering the same question using the hierarchy takes considerably more steps, precisely because of the inaccessibility of the cousin relationship on that array. For the same question, steps on the hierarchy are counted as follows:

Find Susan: female, third oldest, proceed OY	Step 1
To: Anne, female, youngest, irrelevant	Step 2*
To: Sheila, female, oldest, irrelevant	Step 3*
To: Michael, male, irrelevant	Step 4*
To: Paul, male, irrelevant	Step 5
To: Kathy, female, youngest, irrelevant	Step 6*

(Although Kathy is eventually the correct answer, she is considered irrelevant at this point because other rule-relevant family members are still available and must be searched for.)

To: Margaret, female, oldest, irrelevant	Step 7*
To: Richard, male, irrelevant	Step 8*
To: David, male, irrelevant	Step 9
To: Elizabeth, female, youngest, irrelevant	Step 10*
To: Patricia, female, oldest, irrelevant	Step 11*
To: Billy, male, irrelevant	Step 12*
To: Tom, male, irrelevant	Step 13*
To: Jane, female, youngest, excluded	Step 14
To: Peter, male, irrelevant	Step 15*
To: Nick, male, irrelevant	Step 16*
To: Susan, target element, irrelevant	Step 17*
To: Anne, female, youngest, excluded	Step 18
To: Sheila, female, oldest, irrelevant	Step 19*
To: Michael, male, irrelevant	Step 20*
To: Paul, male, irrelevant	Step 21*
To: Kathy, female, youngest, OK	Step 22

The question takes 22 steps on the hierarchy compared to 6 steps on the matrix, because of the inclusion of 16 rule-irrelevant steps (marked * in the example). A similar step count was carried out for the inheritance questions. An important assumption in the step count is the notion of the "next step." The next step is always to an adjacent element, but the direction of stepping is not arbitrary; it is determined by the rule in operation at the time. The direction of stepping is from left to right when OY rules are operating and from right to left when YO rules are being processed. When stepping reaches the edge of the array, the rule also determines where the next step is. For the matrix example just described, the step from David to Jane (Step 4) follows from the rule requirement that if holiday companions of the same birth order cannot be found, then they must be sought among a younger birth order, first among those of older parents. The same rule requirement accounts for the step from Elizabeth to Patricia (Step 11) in the hierarchy example. Again, the crudeness of the step count as a measure should be appreciated. All types of steps—rule-relevant steps, excluding steps, noisy steps, and off-the-edge steps—are given equal weight when it is unlikely that they make the same processing demands on the subject.

Number of steps per question was then correlated with response times. Table 1 shows the mean number of steps, mean response times, and correlation coefficients for each set of questions on both arrays. First, it is clear that number of steps sharply differentiates the matrix from the hierarchy for the holidays questions but not for the inheritance questions. Second, number of steps increases with number of exclusions and at a much steeper rate for holidays questions on the hierarchy than on the matrix (9-56 steps compared to 3-13 steps). The correlation coefficients between number of steps and response times are remarkably high, considering the possible crudeness of the measure. If these correlation coefficients are taken at their face value, then it can be concluded that mentally stepping through the array accounts for a high proportion of the variance in processing time (82% over all conditions). But, for reasons previously discussed (see Procedure), the response time includes the time required to read the number of exclusions as well as the time needed to search for the answer. As the number of exclusions is correlated with the number of steps, it could be argued that reading time alone accounts for the high correlations observed between number of steps and response time, and, if reading time could be successfully dissociated from searching time, then the steps analysis would not be a plausible account of how the subjects processed the arrays. There are several reasons for supposing that this is not the case: the interaction between number of exclusions and array, and the effect of array on one question set and

Table 1
Experiment 1: Mean Number of Steps, Mean Response Time (in Seconds), and Correlation Coefficients Between Steps and Time

Conditions	Mean Number of Steps	Mean Response Time	*r*
Holidays Questions			
Matrix	7.8	22.13	.91
Hierarchy	29.1	49.50	.93
Inheritance Questions			
Matrix	10.1	32.66	.90
Hierarchy	10.1	31.77	.88

Note—N = *16 for each group.*

Table 2
Experiment 1: Correlation Coefficients Between Number of Steps and Components of Response Time-Reading Time and Searching Time for All Conditions

Conditions	Total Group	Memorizers	Searchers
	Reading Time		
Holidays Questions			
Matrix	.83	.84	.79
Hierarchy	.82	.73	.82
Inheritance Questions			
Matrix	.89	.85	.85
Hierarchy	.76	.77	.70
	Searching Time		
Holidays Questions			
Matrix	.80	.71	.67
Hierarchy	.85	.81	.83
Inheritance Questions			
Matrix	.72	.63	.75
Hierarchy	.80	.66	.71

Note—*For total group,* N = *16; for memorizers in matrix conditions,* N = *8, and in hierarchy conditions,* N = *7; for searchers in matrix conditions,* N = *8, and in hierarchy conditions,* N = *9.*

not on the other. Nevertheless, using a dependent measure which includes time to read (albeit for good reasons) may have produced higher correlations than would have been obtained with a measure reflecting only time to search through the array. To check this possibility within the limits of the present data set, correlation coefficients were calculated between number of steps and the components of response time—reading time (Time 2) and searching time (Time 3)—for the total group of subjects and for the subgroups with different processing styles, memorizers and searchers (see Table 2). As might be expected, the correlation coefficients between number of steps and the presumed reading time are high; number of exclusions is correlated with number of steps (r = .87), and reading time increases with number of items to be read. However, the more informative set of correlations is between number of steps and searching time (Time 3). From these data, it is clear that variations in reading time were not totally responsible for the relationship between number of steps and response times described in Table 1. Although the size of the correlation coefficients fall from an overall mean value of .90 to .79, they remain sufficiently high for the steps analysis to be a reasonable account of how the subjects processed the arrays. Also, although the separation of reading and searching times into Times 2 and 3, respectively, is more legitimate for memorizers than for searchers, there is little difference between the two groups in the size of their steps/searching time correlations. However, the overall distributions of processing resources between Times 2 and 3 would lead one to believe that the postulated steps analysis matches more closely the processing style of the memorizers than that of the searchers. The searchers' searching time (Time 3) is indeed determined by the number of steps in the question, but as the magnitude of these times is smaller than that for the memorizers, it is not clear how much searching (and hence stepping) was completed during so-called reading time (Time 2).

Even allowing for these imprecisions in the steps analysis, it is proposed that the basic assumptions of the mental step as a model for processing are supported. Subjects used the mapping of the relationships onto the array in order to answer the questions. They processed the nearest elements in the direction dictated by the birth rule. For the holidays questions, the mapping onto the matrix maximized the proximity of rule-relevant elements, and processing proceeded with a small number of mental steps. The same processing, implemented on the hierarchy, involved many noisy steps and necessitated processing the same elements more than once, resulting in a large number of mental steps. The mapping effect is not general; it is important only for one set of questions precisely because the critical relationships for the other questions are directly mapped onto both arrays and processing incurs the same number of steps. If the mental step proposal is true, then by altering the mapping of the relationships onto the arrays and posing the same questions, it should be possible to change the pattern of response times. Experiment 2 attempted to do this. With a slight modification to the story line of the problem (including the introduction of the children's birth years), the knowledge base was again mapped onto a matrix and a hierarchy. When compared to the arrays in Experiment 1, the mapping was reversed: the cousin relationships were directly mapped onto both arrays, whereas the within-family relationships were made explicit only on the matrix. If the step analysis is valid, the inheritance questions should be more difficult to process on the hierarchy than on the matrix, and the holidays questions should be equally difficult on both arrays.

EXPERIMENT 2

Method

The same set of family relationships was remapped onto a hierarchy and a matrix (see Figures 4 and 5). Thirty-two students who had not participated in the earlier experiment volunteered as subjects. The design and procedure were the same as in Experiment 1. Subjects learned either the hierarchy or the matrix and answered both sets of questions, holidays and inheritance questions. There were no differences between the groups in the number of trials needed to initially learn the arrays; the mean number of trials was 6.8 for those who learned the hierarchy and 7.5 for those who learned the matrix.

Results

The percentage error was 7.0% for the holidays questions/hierarchy and 8.6% for the matrix. The comparable percentages for the inheritance questions were 11.7% and 12.5%, respectively. As reported for Experiment 1, errors occurred randomly across questions, and there was no evidence of a speed/accuracy trade-off. Adjustments, similar to those described in Experiment 1, were carried out on the raw data. All mean values quoted are geometric means in seconds.

Holidays questions. Contrary to the results obtained in Experiment 1, array had no significant effect on the time taken to answer these questions [$F(1,24) = 1.28$]. The mean RT to the questions was 18.71 sec on the hierarchy and 21.53 sec on the matrix. Increasing the number of exclusions produced the expected effect [$F(3,72) = 51.40$, $p < .001$]. The means for one, two, three, and four exclusions were 16.37, 18.20, 18.97, and 28.71 sec, respectively, which are almost identical on both arrays. The birth-order rule again produced a directional reading effect [$F(1,24) = 20.99$, $p < .001$]; the mean RT

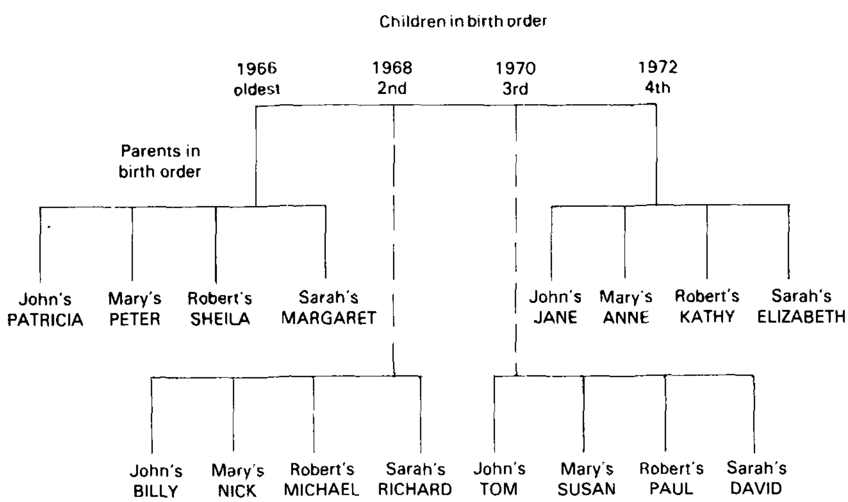

Figure 4. The family relationships arrayed as a hierarchy when the mapping was reversed in Experiment 2. (The version presented to the subjects arranged the four families in full linear sequence from left to right across the page.)

Children in birth order

Parents in birth order	1966 oldest	1968 2nd	1970 3rd	1972 4th
John's	PATRICIA	BILLY	TOM	JANE
Mary's	PETER	NICK	SUSAN	ANNE
Robert's	SHEILA	MICHAEL	PAUL	KATHY
Sarah's	MARGARET	RICHARD	DAVID	ELIZABETH

Figure 5. The family relationships arrayed as a matrix when the mapping was reversed in Experiment 2.

for OY rules was 18.71 sec and 21.53 sec for YO rules. Finding a holiday companion of the same sex was easier (mean = 18.28 sec) than finding one of the opposite sex (mean = 22.03 sec) [$F(1,24) = 42.84$, $p < .001$], although this main effect was modified by two higher order interactions. The first interaction between sex rule and number of exclusions [$F(3,72) = 11.72$, $p < .001$] shows that SS rules were implemented faster for all conditions except for questions with three exclusions, and a further three-way interaction between sex rule, number of exclusions, and birth order rule [$F(3,72) = 4.65$, $p < .01$] shows that the sex rule effect was negligible for oldest to youngest questions with four exclusions. For this experiment, the sex rule did not produce as consistent an effect as in the previous experiment. (Appendix B shows the mean response times for the subconditions.)

Inheritance questions. As had been predicted, times to answer this set of questions were significantly affected by array [$F(1,24) = 26.69$, $p < .001$]. On this occasion, answering the inheritance questions on the hierarchy was much more difficult (mean = 63.83 sec) than on the matrix (mean = 36.56 sec). The main effect of number of exclusions was also significant [$F(3,72) = 139.9$, $p < .001$], as was the interaction between the array and the number of exclusions [$F(3,72) = 2.88$, $.01 < p < .05$], confirming that questions became increasingly more difficult on the hierarchy than on the matrix (see Figure 6). Unexpectedly, the sex rule did have an effect [$F(1,24) = 24.94$, $p < .001$]; answering questions about males (mean = 52.16 sec) took longer than answering questions about females (mean = 44.26 sec). (Appendix B shows the mean response times for the subconditions.)

Discussion

What clearly emerges from Experiment 2 is that reversing the mapping reversed the effects of the arrays. The holidays questions produced an array effect in Experiment 1, but not in Experiment 2; inheritance questions were affected by array in Experiment 2, but not in Experiment 1. When there was an effect of array, there was an interaction between array and number of exclusions, although it shifted from holidays questions to inheritance questions, consistent with the main effect of array. In terms of the steps analysis, the remapping changed the number of steps required to answer questions from both arrays. Table 3 shows the mean number of steps, mean response times, and correlation coefficients between steps and response times for each question set on each array. In Experiment 2, answering the holidays questions involved the same number of steps on both the hierarchy and the matrix (mean = 7.8); in the previous experiment, the step count was 29.1 steps for the hierarchy and 7.8 steps for the matrix. In contrast, the mean step count for the inheritance questions on the arrays differed considerably (matrix = 10.1 steps, hierarchy = 29.9 steps). As in Experiment 1, correlation coefficients between response times and steps were high, even allowing for the inflating effects of reading time. Inspection of the mean response times for the holidays questions for conditions with the same number of mental steps (matrix/Experi-

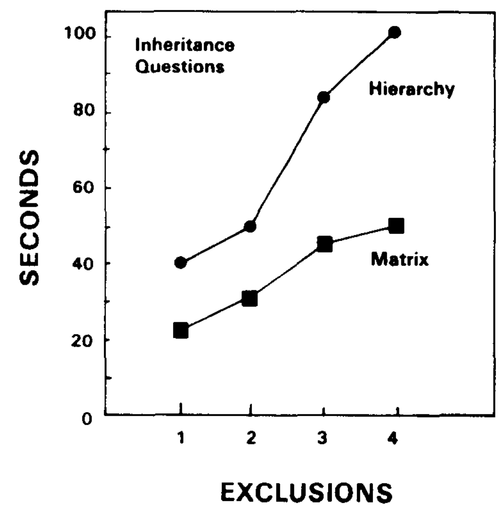

Figure 6. Experiment 2/inheritance questions: mean response times to questions with different numbers of exclusions for the two spatial arrays.

Table 3
Experiment 2: Mean Number of Steps, Mean Response Time (in Seconds), and Correlation Coefficients Between Steps and Time

Conditions	Mean Number of Steps	Mean Response Time	*r*
Holidays Questions			
Matrix	7.8	21.53	.86
Hierarchy	7.8	18.71	.72
Inheritance Questions			
Matrix	10.1	36.56	.87
Hierarchy	29.9	63.83	.96

Note—N = *16*.

ment 1, matrix and hierarchy/Experiment 2, see Tables 1 and 3) shows that the magnitude of processing times were very similar (means = 22.13, 21.53, and 18.71 sec, respectively). The same is true for inheritance questions (see Tables 1 and 3). The effect of mapping, the similarity in the magnitudes of response times for similar step counts, and the correlations between number of steps and response times all point to the conclusion that the mental step analysis is an adequate account of how the students processed the arrays.

GENERAL DISCUSSION

Consistent with more general analyses of representation (Bobrow, 1975; Palmer, 1978), the concept of mapping is considered central to an understanding of problem representation; mappings are efficient only in relation to processing demands. How problem spaces are mentally constructed involves selecting and mapping specific relations from a problem domain; the mapping may facilitate some processing but make other processing more difficult. These experiments have suggested that, for spatial arrays, good representation is realized through a mapping/processing relationship which is implemented in terms of a mental step. Mapping facilitates processing by reducing the number of steps, thus making problem representation more efficient. For these family relations problems (essentially nonspatial problems), such features of the arrays as proximity and dimensionality are used to hold the various family relationships. Making inferences about the relationships are then mimicked by mentally stepping through the arrays. The mental step analysis assumes that a mental representation is constructed which not only encodes the information about the family relationships but does so in a way which mirrors the spatial features of the external array. The mental array is searched systematically: the next step is one of the nearest items and the direction of search is determined by the overall goal of the search. In this way, the spatial array acts as a mental model for reasoning.

In contrast to the studies reviewed earlier (Carroll et al., 1980; Mayer, 1976; Schwartz, 1971), in which spatial arrays were used primarily as external aids to processing, the present experiments show how inferences are drawn and questions are answered from mental spatial arrays. Other reports of mental arrays as the basis for inferencing come from studies on linear syllogisms. De Soto, London, and Handel (1965) and Huttenlocher (1968) reported that in order to answer questions about "Who is the tallest?" subjects construct spatial orderings of the items in the premises which enable them to "read off" the answer. There have been challenges to the spatial ordering interpretation of performance from a psycholinguistic viewpoint (Clark, 1969a, 1969b), but the essential difference between the positions centers on whether the premises in the linear syllogism are integrated into a unified mental representation or whether each premise is represented independently (Johnson-Laird, 1972). Subsequent studies have shown that a unified representation is constructed only under conditions which make large demands on working memory. Potts and Scholz (1975) showed that a unified representation explanation predicts performance when the questions are answered completely from memory, that is, when the problems are no longer visually present. Wood (1969) and Wood, Shotter, and Godden (1974), using three-, five-, and seven-term linear syllogisms, showed that mental spatial arrays of all the information in the premises are constructed when the problems include a large number of premises and when subjects are unpracticed. And, as in the Potts and Scholz study, the spatial strategy is most obvious when the load on memory is greatest, that is, when the problems are presented auditorily rather than visually. With practice and/or with visual presentation of the problems, subjects begin to use the external array to develop shortcuts which avoid the need to construct a mental representation of all the information stated in the premises.

Earlier it was concluded that the subjects studied in the present experiments constructed a mental spatial array through which they searched in a systematic, but exhaustive, fashion. Such a conclusion is consistent with the studies on linear syllogisms. The conditions under which the family relations problems were solved made heavy demands on working memory. The data base to be remembered was large (20 items), and the questions posed were considerably more difficult than even seven-term linear syllogisms. Although the arrays were presented visually, they were removed during testing, and question answering was done completely from memory. There was no evidence that subjects were sufficiently practiced with the problem materials to abandon mental stepping and to devise shortcuts. When considered in conjunction with the studies on linear syllogisms, the findings reported in this paper point to learning conditions where reasoning might be helped by the construction of mental arrays. These conditions may be useful guidelines for those involved in the development of spatial learning strategies. The conditions are (1) if the information to be learned is large and interrelated; (2) if processing has to be completed in working memory, or with minimal external aids; (3) if uncertainty exists about which portion of the information needs to be searched in order to find an answer; (4) if many parts of the data base are involved in making the inference; (5) if the person is relatively unpracticed in the problem domain.

Although these conditions might encourage the construction of mental arrays, they do not determine what array should be constructed in order to best facilitate processing. In many cases, the choice is obvious, as is probably true for a one-dimensional ordering for three-term linear syllogisms. However, as soon as the problem space becomes more complex, there is likely to be a choice about the type of array (matrix, hierarchy, network, graph, etc.) and how the relationships are mapped onto the array. When a choice exists, the problem of achieving a good mapping becomes important. The present studies do not identify the determinants of good mapping, only that mapping is important for processing. Nevertheless, it is likely that choice of mapping will be related to a person's general understanding of a particular problem or level of expertise in a problem domain. To test such speculations, it may be necessary to shift the focus of study from problems which use widely understood relationships (like height and kinship) to problems which require more domain-specific knowledge where expertise might be expected to affect representation through choice of mapping.

In conclusion, the experiments reported in this paper analyzed how spatial arrays can contribute to efficient problem representation. The features of spatial arrays, proximity and dimensionality, are made to hold relationships. Drawing inferences about the relationships are then mimicked by mentally stepping through the arrays. The manner in which the relationships are mapped onto the arrays can reduce the number of mental steps in a problem solution, thus making the problem representation more efficient.

REFERENCES

Bobrow, D. G. (1975). Dimensions of representation. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press.

Carroll, J. M., Thomas, J. C., & Malhotra, A. (1980). Presentation and representation in design problem solving. *British Journal of Psychology*, **71**, 143-153.

Clark, H. H. (1969a). Influence of language on solving three-term series problems. *Journal of Experimental Psychology*, **82**, 205-215.

Clark, H. H. (1969b). Linguistic processes in deductive reasoning. *Psychological Review*, **76**, 387-404.

De Soto, C. B., London, M., & Handel, S. (1965). Social reasoning and spatial paralogic. *Journal of Personality & Social Psychology*, **2**, 513-521.

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.

Greeno, J. (1973). The structure of memory and the process of solving problems. In R. L. Solso (Ed.), *Contemporary issues in cognitive psychology: The Loyola Symposium*. Washington: Winston.

Greeno, J. G. (1977). Process of understanding in problem solving. In N. J. Castellan, D. B. Pisoni, & G. R. Potts (Eds.), *Cognitive theory* (Vol. 2). Hillsdale, NJ: Erlbaum.

Hayes, J. R., & Simon, H. A. (1974). Understanding problem instructions. In L. W. Gregg (Ed.), *Knowledge and cognition*. Hillsdale, NJ: Erlbaum.

Huttenlocher, J. (1968). Constructing spatial images: A strategy for reasoning. *Psychological Review*, **75**, 550-560.

Jeffries, R., Polson, P. G., Razran, L., & Atwood, M. E. (1977). A process model for missionaries-cannibals and other river crossing problems. *Cognitive Psychology*, **9**, 412-440.

Johnson-Laird, P. N. (1972). The three-term series problem. *Cognition*, **1**, 57-82.

Mayer, R. E. (1976). Comprehension as affected by structure of problem representation. *Memory & Cognition*, **4**, 249-255.

Mayer, R. E., & Greeno, J. G. (1972). Structural differences between outcomes produced by different instructional methods. *Journal of Educational Psychology*, **63**, 165-173.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.

Polich, J. M., & Schwartz, S. H. (1974). The effect of problem size on representation in deductive problem solving. *Memory & Cognition*, **2**, 683-686.

Potts, G. R., & Scholz, K. W. (1975). The internal representations of a three-term series problem. *Journal of Verbal Learning & Verbal Behavior*, **14**, 439-452.

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology*, **6**, 436-450.

Schwartz, S. H. (1971). Modes of representation and problem solving: Well evolved is half solved. *Journal of Experimental Psychology*, **91**, 347-350.

Schwartz, S. H., & Fattaleh, D. L. (1972). Representation in deductive problem solving: The matrix. *Journal of Experimental Psychology*, **95**, 343-348.

Simon, H. A. (1975). The functional equivalence of problem solving states. *Cognitive Psychology*, **7**, 268-288.

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

Thomas, J. C. (1974). An analysis of behavior in the hobbits-orcs problem. *Cognitive Psychology*, **6**, 257-269.

Winer, B. J. (1970). Statistical principles in experimental design. London: McGraw-Hill.

Wood, D. J. (1969). Approach to the study of reasoning. *Nature*, **223**, 101-102.

Wood, D., Shotter, J., & Godden, D. (1974). An investigation of the relationships between problem solving strategies, representation and memory. *Quarterly Journal of Experimental Psychology*, **26**, 252-257.

APPENDIX A

Mean Response Times (Geometric Means in Seconds) for Subconditions in Experiment 1

Array		Hierarchy				Matrix			
Exclusions		1	2	3	4	1	2	3	4
Holidays Questions									
Inheritance questions first in order									
Rule OY	SS	21.9	36.1	46.8	65.6	13.6	15.2	25.5	35.4
	DS	39.5	48.1	41.1	66.5	16.1	22.4	23.4	28.3
YO	SS	26.7	44.9	62.8	73.6	16.2	18.7	27.7	25.9
	DS	38.1	62.2	42.2	78.3	16.3	23.4	20.0	35.5
Inheritance questions second in order									
Rule OY	SS	20.7	43.0	43.8	90.8	16.4	15.9	22.7	33.7
	DS	35.5	47.1	56.6	115.1	15.9	27.8	20.8	31.7
YO	SS	23.8	42.3	54.6	90.0	17.6	16.3	22.1	34.8
	DS	34.6	57.3	66.8	101.9	17.6	19.5	25.1	37.6

APPENDIX A (Continued)

	Inheritance Questions							
Inheritance questions first in order								
Rule OY MI	26.4	24.3	40.3	44.5	18.4	28.3	44.7	43.1
FI	31.3	28.7	40.0	43.1	22.1	29.0	33.2	61.5
YO MI	29.9	27.1	45.0	55.1	22.0	30.4	44.7	49.7
FI	34.1	29.6	42.1	40.8	25.0	32.4	45.8	43.2
Inheritance questions second in order								
Rule OY MI	14.1	19.8	29.9	53.7	21.2	30.3	45.1	45.8
FI	16.7	19.8	30.5	40.1	21.6	18.9	34.2	40.4
YO MI	17.9	26.4	50.2	53.3	19.3	25.0	47.3	46.9
FI	17.4	21.4	42.1	48.9	24.1	25.9	37.8	47.3

APPENDIX B
Mean Response Times (Geometric Means in Seconds) for Subconditions in Experiment 2

Array	Hierarchy				Matrix			
Exclusions	1	2	3	4	1	2	3	4
	Holidays Questions							
Inheritance questions first in order								
Rule OY SS	16.0	13.4	22.2	32.7	14.7	17.2	21.8	33.7
DS	27.0	20.9	18.1	30.9	20.0	24.9	19.9	32.7
YO SS	17.8	15.1	21.8	23.6	19.4	23.7	26.6	33.1
DS	20.9	26.5	20.5	45.2	25.3	33.5	20.1	37.5

APPENDIX B (Continued)

Inheritance questions second in order								
Rule OY SS	9.6	10.5	13.5	16.9	13.3	11.8	17.5	25.3
DS	12.9	16.1	13.6	21.3	18.2	22.5	16.6	27.2
YO SS	12.7	16.3	18.2	18.5	15.5	14.7	21.6	27.8
DS	11.0	19.1	19.5	37.7	17.9	18.6	16.8	30.5
	Inheritance Questions							
Inheritance questions first in order								
Rule OY MI	44.3	60.8	78.7	122.2	16.6	27.5	38.7	53.3
FI	28.3	40.2	74.1	93.8	21.0	22.8	32.6	39.4
YO MI	47.2	63.1	93.8	115.6	20.8	26.4	55.3	42.3
FI	32.2	47.0	110.0	105.0	21.3	24.8	45.0	32.4
Inheritance questions second in order								
Rule OY MI	37.2	57.9	91.2	92.5	36.0	34.0	54.2	63.1
FI	30.8	39.0	60.0	100.2	31.0	28.5	51.2	46.9
YO MI	50.8	52.4	88.3	112.7	33.5	36.0	60.1	82.2
FI	36.6	42.2	85.7	78.2	27.6	42.5	50.8	56.1

(Manuscript received May 22, 1985;
revision accepted for publication November 18, 1985.)