British Journal of Psychology (1997), **88**, 473–492 *Printed in Great Britain*

Individual differences and strategy selection in reasoning

Maxwell J. Roberts*
Department of Psychology, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, UK

David J. Gilmore and **David J. Wood**
Department of Psychology, University of Nottingham, UK

Although individuals may use different strategies in order to solve reasoning problems, few attempts have been made to understand the processes that lead to strategy choice. One exception to this is work with the sentence–picture verification task in which it has been found that high spatial ability individuals tend to use a strategy that involves spatial representations while low spatial ability individuals tend to use a strategy that involves verbal representations.

The first study reported here attempted to see whether these findings would generalize to another simple reasoning task with a particularly inefficient spatial strategy. This was found not to be the case; low spatial ability individuals used the spatial strategy while high spatial ability individuals avoided using it. Three explanations were suggested for this based upon (*a*) spatial ability, (*b*) intelligence or (*c*) knowledge. Results of two further studies favoured the spatial ability explanation; individuals do not have explicit prior knowledge of the most effective strategy for this task, and the level of spatial ability determines the degree to which they are able to develop and evaluate the more effective non-spatial strategies.

Current and past emphasis in reasoning research has been on discovering the nature of the underlying fundamental processes; these may consist of either propositions (e.g. Rips, 1994) or mental models (e.g. Johnson-Laird & Byrne, 1991). While the resolution of this issue would be of great importance, Roberts (1993) has outlined a number of problems with this line of research. These follow from the finding that both users of mental models and users of propositions can be identified in most deduction tasks. From this, Roberts suggested that distinguishing the actual fundamental reasoning processes from strategies that overlay and obscure them may be far more difficult than is realized. An alternative, neglected, but potentially fruitful line of research is to attempt to understand why different people use different strategies.

One of the few pieces of work to investigate these issues (MacLeod, Hunt & Mathews, 1978) used the sentence–picture verification task. MacLeod *et al.* outlined two possible strategies for solving these problems; one in which the information in

* Requests for reprints.

the sentence is encoded as a spatial representation which is then compared with the picture (the *spatial strategy*) and one in which the sentence is encoded as a propositional representation which is then compared with a propositional encoding of the picture (the *verbal strategy*). Participants' strategy choice was identified from their response times by using regression and cluster analysis, and it was found that high spatial ability participants were more likely to choose the spatial strategy, while low spatial ones were more likely to choose the verbal strategy (verbal ability being equally high for both strategy groups). This *direct aptitude–strategy relationship* suggests that strategy choice is resource based; people use strategies that make the best use of their abilities.

There are two possible explanations of how people know which is the most appropriate strategy for them; they might perform a cost–benefit analysis comparing the available strategies, choosing the one that their abilities enable them to execute the most effectively. The alternative is that people have distinct cognitive styles; preferences to process information in particular ways. Thus, some people tend to reason with spatial representations, while others tend to use verbal representations (the visualizer–verbalizer distinction—see, for example, Riding, Glass & Douglas, 1993). The basis of this behaviour could be that, with experience, people with low spatial ability assemble a repertoire of verbal strategies that enable them to avoid drawing upon their deficient ability, while high spatial people assemble a repertoire of spatial strategies that enable them to make maximal use of their superior ability.

Overall, MacLeod *et al.* have shown that ability test scores indicate not only the degree of proficiency at reasoning with spatial or verbal representations, but may also indicate the likelihood that these representations will be used for reasoning. However, the actual mechanisms of strategy selection still need to be ascertained and other findings suggest that caution is required even at this point. Several researchers (e.g. Barratt, 1953; French, 1965; Just & Carpenter, 1985) have found that spatial ability tests are also prone to the use of different strategies. Kyllonen, Lohman and colleagues (e.g. Lohman & Kyllonen, 1983) have observed that people who achieve high scores on spatial problems are those who use a larger range of strategies to solve them and are not necessarily using spatial representations. These findings suggest that ability is not necessarily directly related to the type of representation that is used for reasoning, and that more sophisticated models of strategy selection may therefore be necessary.

STUDY 1: THE DIRECTIONS TASK PARADIGM

The studies to be reported are based upon variants of a Compass-point directions task (see Wood, 1978). In its simplest form (the One-person directions task with parallel presentation) participants are given a set of equally sized compass point vectors (e.g. one step east, one step north, one step north) presented simultaneously on one screen. They have to decide at what compass point a person would end up, relative to the starting point, after taking them. With the steps given above, the best answer would be north-north-east. Paths may end precisely on the eight simple compass points (N, NE, E, etc.) but it is not possible for them to end precisely on the complex compass points (NNE, ENE, ESE, etc.). However, in the case of the latter, the error is sufficiently small for it to be clear that, say, NNE is a much more appropriate answer than NE, and participants do not complain about the lack of precision that giving such an answer entails.

There are two distinct strategies for this task; the *spatial strategy* involves generating a spatial representation of the path, step by step, and then reading the compass point of the final bearing of the end point from it. This strategy is particularly demanding, and its accurate use depends upon a person's ability to generate, update and read from an accurate, stable representation of the path. For the *cancellation strategy*, opposite directions are cancelled out with those that remain uncancelled constituting the correct answer. Unlike the spatial strategy, all that is required for its successful execution are simple numerical skills and hence it is far less demanding in terms of spatial ability.

The strategies used for this task may be identified by asking participants to give retrospective reports on completion. Although the use of such reports was thrown into question by Nisbett & Wilson (1977) (see also Ericsson & Simon, 1980), it has never been clear whether retrospective reports will always be unreliable for every task or whether strategies for some tasks may be reported more reliably than others. Many factors will affect this and for the One-person directions task the participants' reports may be validated by seeing whether they perform in line with their strategies; those who claim to use cancellation should be considerably faster and more accurate than those who claim to use the spatial strategy. Exceptions to this rule (e.g. people with exceptionally good spatial ability or exceptionally poor mental arithmetic) should be rare enough not to add an unacceptable degree of error to the validation. If verbal reports are generally reliable, then dividing participants by them is far preferable to performing a median split on performance; this assumes an equal frequency of strategies and can lead to unreliability in the middle of the range of performance. The expected link between performance and reported strategy has been found previously (e.g. Wood, 1978); the difference between the two strategies is very large indeed.

Even when intelligent populations are given the One-person directions task, use of the spatial strategy is very frequent. Hence, it is necessary to know why some people use this highly inefficient strategy at all. A simple cost–benefit analysis account of strategy selection is put in difficulty by this; all participants irrespective of their ability should choose cancellation because of its inherent efficiency and effectiveness. Cognitive-style accounts are not ruled out and individuals with a natural tendency to use spatial representations for reasoning would be expected to use the spatial strategy while those who tend to avoid using them would be expected to use cancellation. Hence, a cognitive-style account of strategy selection derived from MacLeod *et al.*'s findings would predict that spatial strategy users would have high spatial ability, while cancellation users would have low spatial ability. If this were found to be the case, it would be a powerful demonstration of the existence of cognitive styles. For the current study, the chosen measure of spatial ability was the Saville and Holdsworth Advanced Test Battery Spatial Reasoning Test (ST7) (Saville & Holdsworth Ltd, 1979). This requires participants mentally to fold two-dimensional figures in order to make patterned cubes, and then decide which of various options they correspond to (this is similar to tasks investigated by Shepard & Feng, 1972).

The verbal ability test from the same battery was also administered; the Verbal Concepts Test (VA1). This requires words to be chosen to complete analogies. In addition, knowledge of compass points was tested as this could also be associated with strategy selection.

In addition to the One-person directions task, a Two-person directions task with serial presentation was administered. This consists of separate sets of directions for two people; participants have to decide at what compass point one person would end up relative to the other. Each step is presented individually, thus removing the opportunity to cancel by scanning up and down a list of directions. If cancellation is to be used, heavier demands are therefore made upon working memory and hence, when compared with the spatial strategy, some of its advantage is lost. Cancellation is also harder to derive as some directions must be reversed before being cancelled. The initial motive behind giving this task was to discourage the use of cancellation irrespective of previous strategy use and hence to investigate transfer effects. While the cancellation users should be more efficient on the One-person directions task, this strategy is the more task specific and these participants may pay an overall penalty (by using a new, unpractised strategy) for the Two-person directions task when compared with those who used the safer, more general, spatial strategy previously.

Method

Participants

Participants were 78 first-year undergraduate psychology students from the University of Nottingham who were pre-tested with the spatial and verbal ability tests one month before the main experimental sessions. From the test scores, 20 high spatial participants were selected (at least 22 items correct, mean score 27.2, SD 2.9; 10 male, 10 female), and 20 low spatial participants (16 items or fewer correct, mean score 13.0, SD 2.0; three male, 17 female). There was a payment of £1.00 for participation in the main experimental session.

Apparatus

Stimuli were presented by an Apple Macintosh Plus microcomputer running MacLab (Costin, 1988). Response collection was either by keyboard or voice-key.

Materials

The Compass-point test consisted of 24 trials; two of every compass point excluding north, east, south and west. Each trial consisted of a heavy cross in the centre of the screen and a diamond in its vicinity. The location of the diamond for the complex compass point trails was slightly distorted so as to match the corresponding end points for the One-person directions task (see below).

The One-person directions task consisted of one practice trial and two blocks of 18 experimental trials. For each trial, all steps were shown simultaneously on one screen (see Fig. 1). Each block consisted of six each of six-, seven- and eight-step problems. These were devised such that no adjacent steps were opposites and the path would never end at the starting point, either ending on one of the eight simple compass points (a north/south:east/west ratio of 1:1 for NE, NW, SE and SW) or near to one of the eight complex compass points (a north/south:east/west ratio of 1:2 or 2:1).

The Two-person directions task consisted of 12 trails devised using the same criteria. An example is given in Fig. 1 (dashed lines indicate a new screen). All problems involved A taking two steps then B taking two steps. The required answer was always B relative to A. Participants were aware that this was the case.

One step East One step East One step East One step South One step West One step South One step West	A takes: One step North then: One step East B takes: One step East then: One step East end Where is B relative to A?

Figure 1. Example items from the One-person directions task (left) and the Two-person directions task (right).

Procedure

Participants were tested individually in a single session lasting approximately 35 minutes. For all tasks, written instructions emphasizing speed and accuracy were given before each, the order of trial presentation was randomized, the computer recorded response latencies, the experimenter recorded participants' answers, and no feedback as regards accuracy was given.

Participants were first given the Compass-point test. Along with instructions, they were given a diagram to remind them of the 16 compass points. This could also be referred to between subsequent tasks. Participants were to name the compass point of a diamond relative to a cross at the centre of the screen. No practice trial was given. For each trial a blank screen was shown (3.5 s) followed by a cross at the centre (1.5 s). This was then emboldened, and simultaneously the diamond was shown. Naming the compass point activated the voice-key, clearing the screen and beginning the next trial.

Participants were then given the One-person directions task. After the instructions a practice trial was given followed by a diagram showing its path and its answer. This was followed by the two experimental blocks (whose presentation was counterbalanced). For each trial a blank screen was shown (3.5 s) followed by a cross (1.5 s) and then the set of directions. On deciding on an answer, the participant was to speak it, and simultaneously press a key on the keyboard. This cleared the screen and began the next trial. Participants frequently vocalize when solving these problems which precludes the use of the voice-key. The generally long solution times meant that any inaccuracies due to this method of response collection would be proportionately small.

Participants were finally given the Two-person directions task. Written instructions included diagrams showing the paths of an example trial. Each trial began with a blank screen (3.5 s) then a cross (1.5 s) and then the first screen of the problem. Participants advanced to the next screen by pressing a key on the keyboard (a blank screen was shown for 0.5 s between steps). This continued until the fourth screen had been reached and this prompted them for an answer. As before, participants were to speak the answer, simultaneously pressing a key. This cleared the screen and began the next trial.

At the end of this task, participants were asked to describe in writing, separately, the strategies that they had used for each of the directions tasks, including any shifts or mixing of strategies that they had been aware of in each.

Design

The strategy described by participants on the One-person directions task was used as a classification variable, with solution time and accuracy being used to confirm the validity of the reports. Scores on the ability tests and performance at the Compass-point test were to be used to predict strategy selection.

Results

Error rates (the percentage of problems incorrect) are reported throughout (as opposed to scoring the size of the errors). In order to have sufficient data points so that accurate mean solution times could be calculated for every individual, no solution time was excluded for a near-miss answer (no more than one compass point out). Thus, for example, if the correct answer was N, solution times would be included for answers of NNW but not for answers of NW.

For the One-person directions task, key presses were occasionally made before or after the participant spoke the answer. When this occurred, the response time (but not the answer) was excluded. The overall mean solution time was 10.8 s (SD 6.8; individuals ranging from 3.7 to 36.0 s) and the overall mean error rate was 32.8 per cent (SD 25.1 per cent; individuals ranging from 0 per cent to 75 per cent). Fewer than 12 per cent of answers were more than one compass point step out (55 per cent was the highest for an individual, still less than the 81 per cent expected if the individual was guessing).

Participants were able to describe their chosen strategies clearly, with little difficulty or uncertainty. The majority (31 out of 40) either reported one strategy with no mention of the other, or else one strategy was reported as being dominant (i.e. used for the majority of the trials) while the other was very quickly dispensed with or tried only one or twice out of curiosity. The dominant strategy was also the final choice in all cases. For five of the remaining participants, the dominance and final choice of strategy were less clear cut, but there was still a distinct preference. For example, an individual might report using the spatial strategy for every trial, and report using cancellation to double check one-third of the trials. Here the spatial strategy is both dominant and the final choice since the person has failed to appreciate that the slow, inaccurate process of generating a spatial representation for every trial is unnecessary. The four remaining participants reported changing strategies with a clean break from spatial to cancellation no later than one-third of the trials into the task. Thus cancellation was both their dominant and final chosen strategy.

Nineteen participants were classified as using the spatial strategy and 21 as using cancellation. Spatial strategy users were considerably less accurate (mean error rate 52.3 per cent, SD 17.8) than cancellation users (mean error rate 15.1 per cent, SD 15.9); $U = 27.5$, $p < .01$. Spatial strategy users were also considerably slower with a mean solution time of 14.7 s per problem (SD 7.5) versus 7.3 s (SD 3.4) for cancellation users; $U = 41$, $p < .01$. However, Fig. 2 shows that one participant (who reported using cancellation) was clearly out of step with the rest. It was therefore impossible to unambiguously classify this person's strategy choice and these data were not considered further in the analysis. Subsequently, the mean solution time for the cancellation group became 6.8 s (SD 2.5), and the error rate became 12.3 per cent (SD 9.4). Given the otherwise low overlap between strategy groups and the almost perfect agreement between performance and reported strategy, the verbal reports can be taken to be accurate descriptions of strategy choices. It would be hard to account for these findings in any other plausible way.

Means etc. and correlations between the predictor variables for the remaining 39 participants are given in Table 1. Individual logistic regressions (using SPSS) were

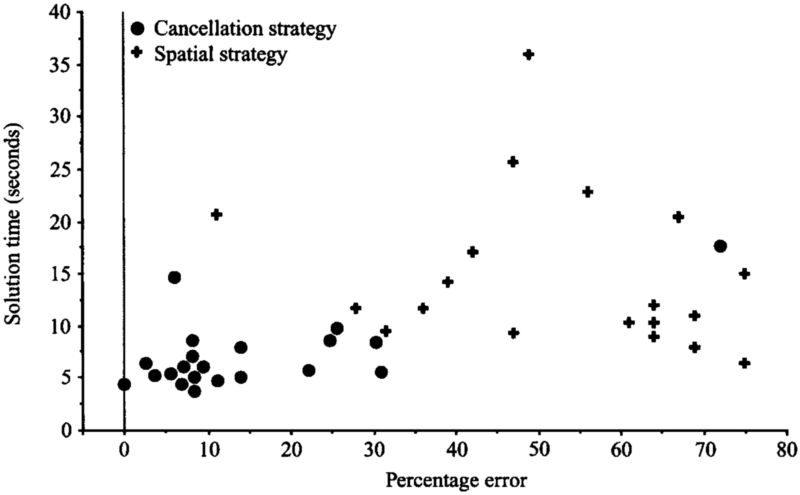

Figure 2. Scattergram of mean solution time vs. error rate for the One-person directions task (Study 1).

Table 1. Overall performance at the predictor variables and correlations between them (Study 1)

	Performance				Intercorrelations			
	Mean	SD	Min.	Max.	CPT RT	CPT % Err.	Spatial	Verbal
CPT response time (s)	2.3	1.0	1.0	5.7	—	0.26	−0.40**	−0.33*
CPT % error	28.0	16.5	0	67		—	−0.56**	−0.20
Spatial ability test score	20.1	7.6	7	32			—	0.48**
Verbal ability test score	26.0	4.0	20	38				—

$^*p < .05$; $^{**}p < .01$.

used to see whether each variable predicted strategy choice by itself over and above the constant. Point biserial correlations are also given in brackets, coding the cancellation strategy as 1 and the spatial strategy as 0. For all χ^2, d.f. = 1. Spatial ability test score was a significant predictor of strategy choice ($\chi^2 = 14.8$, $p < .01$ ($r = 0.59$)) as was error rate at the Compass-point test ($\chi^2 = 5.6$, $p < .05$ ($r = -0.37$)). Compass-point test response time was not a significant predictor ($\chi^2 = 2.0$, $p > .05$ ($r = -0.22$)) nor was verbal ability test score ($\chi^2 = 0.8$, $p > .05$ ($r = 0.15$)). Thus, those who had high spatial ability and/or made fewer compass point errors were more likely to use cancellation, and those who had low spatial ability and/or made more compass point errors were more likely to use the spatial strategy.

Logistic regression was then used to compare the predictive power of spatial ability test score and Compass-point test error rate. When spatial ability test score was entered first the addition of Compass-point test error rate to the model contributed no further significant predictive power ($\chi^2 = 0.3$, $p > .05$); indeed, nor did the addition of any other predictor (no χ^2 exceeded 1.6, $p > .05$). When Compass-point test error rate was entered first, spatial ability test score contributed further significant predictive power to the model over and above that already accounted for ($\chi^2 = 9.5$, $p < .01$). Thus, although error rate for the Compass-point test predicted strategy selection, this was entirely accounted for by spatial ability.

For the Two-person directions task, overall means and error rates were calculated as before, summing the four key press times for each trial to give each solution time. One participant was unable to speak any answers synchronously with the final key press and so only the errors were analysed for this person and not the response times. The mean overall solution time was 12.8 s (SD 4.8, $N = 38$, individuals ranging from 5.5 to 25.8 s) and the mean error rate was 23.5 per cent (SD 22.4, $N = 39$, individuals ranging from 0 per cent to 92 per cent).

Because strategy choice for the One-person directions task has been found to be strongly related to spatial ability, the data here cannot straightforwardly be used to investigate transfer. However, comparing the Two-person directions task performance of high spatials who chose cancellation previously with low spatials who chose the spatial strategy previously would determine whether the high spatial participants pay any penalty for choosing cancellation and having to adapt to a new strategy for this task. Taking the accuracy with which the strategies were described for the One-person directions task as evidence for accurate reporting here, only two people reported attempting any cancellation at all, and these participants were excluded from this analysis.

The difference in error rates was significant ($U = 20.5, p < .01$), with a high spatial mean of 8.4 per cent (SD 8.1, $N = 14$) and a low spatial mean of 33.7 per cent (SD 23.7, $N = 15$). The difference in solution times was also significant ($U = 44$, $p = .01$), with a high spatial mean of 11.5 s (SD 4.5, $N = 13$), and a low spatial mean of 15.3 s (SD 4.6, $N = 15$). Hence, the high spatials were both faster and more accurate than the low spatials.

Discussion

The main finding for this study is an apparently counter-intuitive *inverted aptitude–strategy relationship*. Despite the fact that all of the participants were highly educated (and competence in mathematics was a course entry requirement), those with low spatial ability failed to realize that cancellation was an alternative to the spatial strategy which would yield considerable savings in time and effort as well as increased accuracy. Instead they tended to choose the strategy that made heavy demands on ability that they did not possess. This suggests that they have not developed a repertoire of non-spatial strategies to make better use of their resources. High spatial ability participants tended not to choose the spatial strategy even though they would have been the more suited to using it. Instead, their choice of strategy made far less use of their spatial ability. The second major finding confirms that the high spatial participants were relatively more flexible than the low spatial ones. Not only were the high spatials more likely to use the most efficient strategy for the One-person directions task, but they were also able to change to a spatial strategy and perform well at it when the task facets forced this. Thus they outperformed the low spatial individuals on the Two-person directions task even though the low spatials already had considerable practice at using the spatial strategy for the previous task.

These results, in line with those of Lohman & Kyllonen, demonstrate that a cognitive-style dimension such as visualizer–verbalizer is not able to make general predictions as to how people will reason. Hence, before any questions concerning how people choose between different strategies can be answered, it is necessary to understand why the full set of strategies was not available to all people. Hence, given that the (less efficient and more stressful) spatial strategy was almost certainly available to all, it needs to be explained why cancellation was available mainly to those with high spatial ability and not those with low spatial ability. There are three possible explanations for these findings.

The first explanation is that spatial ability determines a person's ability to develop and evaluate the cancellation strategy. Participants differ in their ability (as measured by the spatial ability test) to generate, manipulate, update and read from accurate, stable spatial representations. These are precisely the processes that must be performed accurately when executing the spatial strategy. It is possible to develop cancellation from the spatial strategy because moving two steps in opposite directions will reveal that they are redundant no matter how many steps intervene, leading to the discovery that the spatial representation is entirely redundant. High spatial individuals, who are more able to execute the spatial strategy accurately, will be in a much better position to discover this than low spatial individuals, who cannot construct sufficiently precise representations.

After developing the cancellation strategy, there may be a need for evaluation in order to ensure that it is valid. One method for this is to compare an answer given by the spatial strategy with the answer given by cancellation. Because low spatial participants are less likely to be able to execute the spatial strategy accurately, they are far more likely to obtain conflict when comparing answers. As the spatial strategy is known to be valid, while the status of the cancellation strategy is unknown (otherwise it would not need to be evaluated to begin with), the outcome for low spatial participants is that the spatial strategy will almost always win, even though it is cancellation that is more likely to be giving the correct answer. Because the high spatials are better able to execute the spatial strategy accurately, they will be more likely to find agreement in answers between the two strategies and hence conclude that cancellation is valid. These processes of strategy development and evaluation need not necessarily occur while the problems are being solved. Participants are likely to think about a task as soon as they read its instructions.

The second explanation is that prior knowledge determines strategy selection. The findings of Lohman & Kyllonen regarding individual differences in strategy use on spatial problems mean that great care must be taken in inferring causal effects from test scores. Although high spatial ability may lead to the development of new strategies, an alternative is that due to past learning experiences, some people may simply have acquired a larger strategy repertoire than others. People with the largest range of strategies would be more likely to possess an effective strategy for any given task including spatial ability tests and the One-person directions task. The problem simply becomes one of recognizing which strategy is appropriate. This position is effectively an extension of findings from the expertise literature (e.g. Ericsson &

Charness, 1994; Ericsson & Smith, 1991) which show that experts use superior strategies to novices not because they have higher ability, but because they have undergone the necessary training to learn the strategies.

The third explanation is that intelligence determines strategy selection. Hence, despite the intentions of the spatial ability test constructors and the outward similarity of this test to others designed to measure spatial ability, the test is really acting as a measure of intelligence (*g*), being highly discriminative even amongst university students. It would not be surprising that intelligent people perform more intelligently on the One-person directions task. However, the lack of any prediction of strategy selection by the verbal ability test should sound a note of caution for this explanation, as this test should also load onto *g*.

Both the knowledge and intelligence explanations deny that people differ in their ability to represent spatial information, or if they do differ, they deny that these differences constrain the development of strategies, either because no development takes place at all (the knowledge explanation) or because intelligence is a far more important determinant. Although there are problems with both of these interpretations (which will be discussed later) two further studies were devised in order to see which was the best of the three alternatives.

STUDY 2: THE EFFECT OF THE AUGMENTATION OF SPATIAL ABILITY ON STRATEGY SELECTION

If ability to form spatial representations determines whether or not cancellation will be developed for solving One-person directions task problems, then improving this ability (for example by giving people the option of using paper and pencil) should increase the likelihood that cancellation will be developed, and enable it to be evaluated more effectively, particularly for low spatial individuals. Work by Larkin & Simon (1987) suggests that it is valid to treat external and internal spatial representations as equivalent. Hence, they define mental imagery as '...the uses of diagrams and other pictorial representations that are not stored on paper but are held in human memory' (p. 97) and state that '...the creation of a mental imagine [*sic*] (for instance from a verbal description) employs inference processes like those that make information explicit in the course of drawing a diagram' (p. 98).

If giving participants the opportunity to use pencil and paper increases the incidence of cancellation, the spatial ability explanation would be supported; the addition of these aids would not be expected to have any effect on either knowledge or intelligence. If, despite being given pencil and paper, low spatials persisted in using the spatial strategy and high spatials still used cancellation, then this would be evidence against the spatial ability explanation.

Method

Participants

Participants were 164 students at the Universities of Nottingham and Newcastle upon Tyne who were pre-tested with the same ability tests as for Study 1. Using the same criteria, 20 high spatial participants were selected (mean score 26.7, SD 3.4; six male, 14 female), and 20 low spatial ones (mean score 13.7, SD 2.5; four male, 16 female). There was a payment of £1.00 for participating in the experimental session.

Apparatus and materials

These were identical to Study 1, but excluded the Two-person directions task.

Design

The participants' choice of strategy and their spatial ability group were classification variables. It was intended to compare the strategy selection and performance of those who had taken part in Study 2 with those who had taken part in Study 1. Solution time and accuracy at the One-person directions task were dependent variables for the purpose of investigating task performance.

Procedure

Participants were tested individually, performing both tasks sequentially in the same session. The procedure for the Compass-point test and One-person directions task were identical to Study 1 except for the following alteration to the latter: after this task was described to participants, they were informed: 'You are provided with a pen and sheets of paper, if at any point during the trials you feel that you would like to use them to help you solve the problems, then feel free to do so.' The instructions emphasized that the problems should be solved as quickly and as accurately as possible and that the use of pencil and paper was entirely optional. Before the commencement of the trials, participants were provided with a pencil and sheets of blank paper. The experimenter recorded which trials these had been used for. Afterwards participants were asked to describe in writing the strategies that they had used for the directions task.

Results

Data from the Compass-point test were not analysed because this task was included to ensure that the Study 2 participants attempted the One-person directions task with similar prior experience to those in Study 1. Classifying participants by their verbal reports was straightforward in almost every instance, with the same patterns of strategy description being observed as for Study 1. All who used pencil and paper used them to externalize the spatial strategy, none externalized cancellation for even a single trial. Where participants reported using pencil and paper, this was found to match records kept by the experimenter. For one person, the final choice of strategy was made comparatively late after much experimentation (within the last six trials of the second block). Hence, although the dominant strategy was spatial, the final choice of strategy was cancellation. Where making comparisons of strategy choice (as opposed to investigating strategy performance where the dominant strategy should be used), final chosen strategy is the most appropriate classification; it is better to choose a more efficient strategy late than not at all. The final choices of strategy are shown in Table 2. Significantly fewer individuals chose cancellation compared with Study 1 ($\chi^2 = 11.8$, $p < .01$), particularly for the high spatials ($\chi^2 = 12.1$, $p < .01$). In the following comparisons, there were no significant differences in verbal and spatial ability test scores between subgroups (no t exceeded 1.5, all $p > .1$). Means and error rates were calculated identically to Study 1.

Table 2. Comparison of the final chosen strategies for Study 1 and Study 2

	Study 2 (Option to use pencil and paper)		Study 1 (No pencil and paper)	
	Spatial	Cancellation	Spatial	Cancellation
High spatial	15[a]	5	4	16
Low spatial	19[b]	1	15	4
Total	34	6	19	20

[a] 12 participants used pencil and paper, three used the spatial strategy without these aids.
[b] All 19 participants used pencil and paper.

Looking at the low spatial individuals, the 19 in Study 2 who used the spatial strategy aided with pencil and paper were significantly more accurate but *not* significantly faster than the 15 in Study 1 who used the spatial strategy unaided. The mean solution time was 14.4 s (SD 4.2) in Study 2 and 15.9 s (SD 7.9) in Study 1; $U = 136.5, p = .84$. The mean error rate was 42.8 per cent (SD 15.8) in Study 2 and 54.7 per cent (SD 16.6) in Study 1; $U = 77, p < .05$. Hence, although using pencil and paper appeared to confer a slight advantage for low spatial participants, gains were far smaller than would have been expected had they used cancellation (Roberts, 1991).

Looking at the high spatial individuals, the 13 in Study 2 who were spatial strategy dominant and aided with pencil and paper were significantly less accurate and significantly slower than the 16 in Study 1 who used cancellation. The mean solution time was 9.8 s (SD 2.0) in Study 2 and 6.4 s (SD 2.6) in Study 1; $U = 21, p < .01$. The mean error rate was 22.2 per cent (SD 13.8) in Study 2 and 9.8 per cent (SD 7.7) in Study 1; $U = 35, p < .01$. Hence, choosing the spatial strategy externalized on pencil and paper is a poorer option for high spatials than choosing cancellation.

Discussion

Overall, augmenting spatial ability by giving the opportunity to use pencil and paper did not help the low spatial participants to develop cancellation (the most effective strategy of all for them). However, because far fewer high spatials used cancellation in Study 2, the spatial ability explanation of strategy development has not necessarily been ruled out. Instead, it can be argued that this is evidence against the knowledge-based explanation. If it is to be surmised that a person brings a repertoire of strategies to a task (rather than derives and evaluates new ones while performing it) then it would have to be explained how the option of using pencil and paper eliminated the most efficient strategy from most of the repertoires. However, ruling out the knowledge explanation on this basis assumes that pencil and paper reduced the effort required to execute the spatial strategy to the extent that the *need to develop* more efficient alternatives was suppressed. Knowledge would still be implicated if pencil and paper suppressed the *need to use* more efficient strategies, so that participants still knew that cancellation was possible, but did not adopt it because, despite being encouraged to solve the problems as quickly and as accurately as possible, they were not sufficiently motivated to do so. Although this would be difficult to disprove empirically, it should be noted that every participant who used pencil and paper did so to augment the spatial strategy. Not one single participant used pencil and paper to externalize cancellation (e.g. by copying down the initial letter of each compass point step, and then crossing out opposites) despite the fact that this is likely to be one of the most effective, and least effortful strategies of all. Had cancellation been in the repertoire of most of the high spatial individuals before the beginning of the study, then it would have been reasonable to expect *some* of them to use pencil and paper to aid this strategy. The absence of this suggests that most participants do not possess explicit knowledge that cancellation is possible for this task, and hence overall that the opportunity to use pencil and paper suppresses the reasoning processes leading to strategy development rather than suppressing strategies that already exist in people's repertoires.

Knowledge-based explanations of strategy selection have several theoretical difficulties. At an extreme, they are tautological in that they effectively assert that some people choose better strategies *because* they know better strategies while simultaneously using observations of better strategy choice as *evidence* for better strategy knowledge. Quite apart from denying the importance of reasoning as a means of developing strategies, the claim is difficult to falsify because it is difficult to distinguish between a newly developed strategy and a piece of knowledge acquired at some unspecifiable point in the past and possibly hidden from view until triggered by the right circumstances. Thus, while this position is not necessarily false, care must be taken not to over-apply it in its current form. It is also necessary to consider the extent to which findings from the expertise literature may be generalized to a domain where some people choose superior strategies but have not obviously devoted a greater period of time specifically to learning about it. In order to generalize to such domains (including the One-person directions task), it is necessary to assert that the most successful performers have superior knowledge of related domains. The problem with this assertion is that expert skills transfer with notorious difficulty. Hence, while expert chess players may have a superior memory for chess positions, this skill does not even transfer to memory tasks where the chess pieces have been placed on a board at random, and hence the assertion becomes a self-contradiction.

The knowledge-based approach is strictly a strategy-possession account of reasoning and problem solving. Experts possess relevant effective strategies while novices do not. Clearly, those who are lucky enough to know many strategies and hence who are more likely to know the correct strategy in any given situation will have an advantage over those who do not. However, where some people use an effective strategy, while other people who are likely to have similar domain-specific knowledge do not, then expertise/knowledge-based explanations are not enough to account for strategy selection.

STUDY 3: COMPARING SPATIAL ABILITY AND INTELLIGENCE AS PREDICTORS OF STRATEGY SELECTION

In Study 3, participants were drawn from a panel of elderly people in order to replicate the findings of Study 1 on a non-student population and extend the range of intelligence tested. If intelligence determines strategy selection then increasing the range should increase the level of prediction. Three intelligence test scores were available; both parts of the AH4 test (Heim, 1970) and the Culture-Fair Intelligence Test (Cattell & Cattell, 1959). The AH4 part 1 consists of verbal and mathematical reasoning items while part 2 consists of non-verbal and spatial reasoning items. The Culture-Fair Intelligence Test consists solely of non-verbal reasoning items similar to the AH4 part 2, and is designed to load onto the fluid intelligence factor. A less demanding spatial ability test was selected than has been previously used; the ASE (NFER-Nelson) General Ability Tests Spatial Test (Smith & Whetton, 1988). This requires participants mentally to fold various two-dimensional patterns and then compare these with three-dimensional drawings.

The Compass-point test was administered to ensure equivalent experience prior to the One-person directions task. However, it was anticipated that it would be difficult to use the voice-key with this population; instead the experimenter controlled presentation. The One-person directions task was redesigned to be less demanding and consisted of five- and six-step problems only. It was expected that it would be highly unlikely that cancellation would lose its advantage when compared with the spatial strategy.

Method

Participants

These were 64 people (26 were male and 38 female) from the University of Manchester Age and Cognitive Performance Research Centre subject panel. They were aged between 60 and 71 years inclusive. Thirty-two high IQ participants were recruited from those who had scored at least 40 out of 65 correct on the AH4 part 1 test. Thirty-two mid-IQ participants were recruited from those who had scored between 25 and 39 correct inclusive. All had corrected-to-normal vision where necessary, were tested at the centre and were paid £3.00 for participation.

Apparatus

Stimuli for the Compass-point test and One-person directions task were presented using an Apple Macintosh SE microcomputer running MacLab.

Materials and procedure: Psychometric tests

The AH4 tests had been administered two to three years previously. The Culture-Fair Test (scale 2, form B) had been administered one to two years previously. The spatial ability test was administered at the beginning of the main experimental session before the computer tasks.[1] One modification to this

[1] Test–retest reliability coefficients show that the time delay between the administration of the intelligence tests and the One-person directions task is unlikely to be a problem. The second psychometric testing session was after the One-person directions task and at least three years after the first psychometric testing session. The test–retest reliability for the AH4 part 1 was 0.87 and for the AH4 part 2 was 0.85 ($N = 1135$). In addition, means and standard deviations were virtually unchanged between testing sessions, showing not only that the rank ordering of scores has been reliably measured but also that there has been no decline in intelligence nor compression of the range with time.

test was that participants were shown cut-out models of the practice items if they made any errors in order to ensure their understanding of the task.

Materials and procedure: Computer-presented tasks

For the experimental session, participants were tested individually. The same 24 Compass-point test trials as before were used, the only change to the procedure was that the voice-key was not used. When the participants announced the answer, the experimenter advanced to the next trial. The final task was the One-person directions task; there were 32 trials in two blocks of 16 with a practice trial. Each block consisted of eight five-step and eight six-step trials, created using the same criteria as before. This task was otherwise identical to the task in Study 1. After performing both computer tasks, participants were asked to describe in writing their choices of strategy for the One-person directions task.

Design

There were five variables that were to be used to predict the final choice of strategy of the participants. In addition, error rate and mean solution time on the One-person directions task were to be used to verify that they had given accurate accounts of their strategy choices.

Results

For the One-person directions task, means and error rates were calculated as before. Overall, less than 23 per cent of answers were wrong by more than one step overall (59 per cent being the highest recorded for an individual). The overall mean solution time was 14.8 s (SD = 6.2, ranging from 6.1 to 31.6 s) and the error rate was 52.9 per cent (SD = 22.8, ranging from 3 per cent to 91 per cent).

All participants were able to give classifiable reports describing their choice of strategy. Patterns of strategy mixing were similar to those found in Study 1, although the use of cancellation was less frequent in Study 3. Where people changed from the spatial strategy to cancellation, this tended to take place later than has been observed before. For this reason, dominant strategy will be used for validating the verbal reports, but final choice of strategy will be used when predicting strategy selection.

Fifty-three participants were spatial strategy dominant, nine were cancellation strategy dominant and two were mixed (claiming to have changed strategy exactly half-way through the trials at the break). The final choice of strategy was spatial for 47 and cancellation for the remaining 17. Looking at the dominant strategy: for solution times, the spatial strategy users had a mean of 15.9 s (SD 6.2); this was significantly slower than the cancellation and mixed strategy users who together had a mean of 9.7 s (SD 2.6); $U = 109.5$, $p < .01$. For error rates, the spatial strategy users had a mean of 58.2 per cent (SD 19.5); this was significantly worse than the cancellation and mixed strategy users who together had a mean of 27.6 per cent (SD 21.7); $U = 84$, $p < .01$. Thus, the verbal reports again appeared to be an accurate reflection of strategy choice as is shown in Fig. 3.

Means etc. and correlations between the predictor variables are given in Table 3. High inaccuracy at naming compass points was mainly due to confusing east with west. Individual logistic regressions were used to see whether each variable predicted strategy choice by itself over and above the constant. Point biserial correlations are also given, coded as for Study 1. For all χ^2, d.f. = 1. Spatial ability test score was a

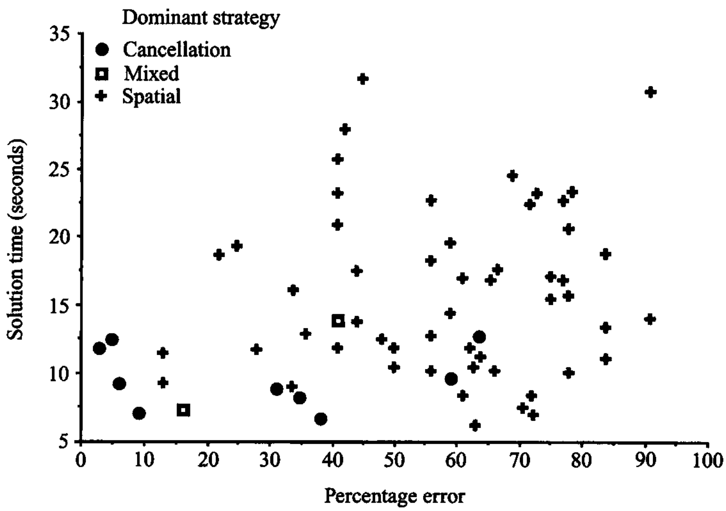

Figure 3. Scattergram of mean solution time vs. error rate for the One-person directions task (Study 3).

Table 3. Overall performance at the predictor variables and correlations between them (Study 3)

	Performance				Intercorrelations				
	Mean	SD	Min.	Max.	CPT % Err.	Spatial	AH4/1	AH4/2	CFT
CPT % error	38.7	21.5	0	100	—	−0.30*	−0.37**	−0.38**	−0.48**
Spatial score	53.5	11.2	24	74		—	0.32**	0.59**	0.60**
AH4 part 1	39.5	7.7	26	61			—	0.55**	0.58**
AH4 part 2	37.3	9.2	21	60				—	0.59**
Culture-Fair Test	31.4	4.8	16	41					—

$*p < .05$; $**p < .01$.

significant predictor of strategy choice ($\chi^2 = 7.8, p < .01$ ($r = 0.34$)) as was Culture-Fair Test score ($\chi^2 = 5.8, p < .05$ ($r = 0.29$)). Compass-point test error rate was not a significant predictor ($\chi^2 = 0.1$, $p > .05$ ($r = -0.04$)), nor was AH4 part 1 score ($\chi^2 = 3.0$, $p > .05$ ($r = 0.22$)) nor AH4 part 2 score ($\chi^2 = 1.9$, $p > .05$ ($r = 0.17$)). Again, high spatial ability was more likely to be associated with cancellation than low spatial ability.

Logistic regression was then used to compare the predictive power of spatial ability test score and Culture-Fair Test score. When spatial ability test score was entered first, the addition of Culture-Fair Test score to the model added no significant further predictive power ($\chi^2 = 0.6$, $p > .05$); indeed, nor did the addition of any other predictor (no χ^2 exceeded 0.9, $p > .05$). When Culture-Fair Test score was entered first, spatial ability test score also failed to contribute significant further predictive power to the model over and above that already accounted for ($\chi^2 = 2.6$,

$p > .05$), but when the AH4 part 1 score was entered first, spatial ability test score did contribute significant further predictive power ($\chi^2 = 5.7$, $p < .05$).

Discussion and comparison with the results of Study 1

Taking these results by themselves, spatial ability test score was a moderately successful predictor of strategy selection, with only the Culture-Fair Test of the intelligence tests also predicting. It could be suggested that the spatial ability explanation has only a slight advantage over the intelligence explanation because the spatial ability test does not have any advantage over the Culture-Fair Test as a predictor. However, had intelligence rather than spatial ability determined strategy selection, then all three intelligence tests should have predicted strategy selection. The selection of participants on the basis of their AH4 part 1 score ensured that the range of intelligence sampled in Study 3 had been extended downwards substantially compared with Study 1 but without sacrificing the top of the range. Had intelligence determined strategy selection, then extending the sampled range of intelligence and using tests specifically designed to measure it should have resulted in far more convincing predictions of strategy choice. The overall level of prediction should have increased in Study 3; it certainly should not have declined to the extent found.

When comparing spatial ability levels between Study 1 and Study 3, published general population norms may be used in order directly to compare scores on the two tests. A median split performed on the Study 3 spatial scores produces a group exactly equivalent in spatial ability to the low spatial group of Study 1 (ranging from the 1st to the 31st percentile for the former and from the 3rd to the 30th for the latter). Not only do these groups not differ in spatial ability ($U = 246.5$, $p > .05$) but Table 4 shows that they do not differ appreciably in strategy selection. At the other end of the scale, the high spatial ability group from Study 1 (ranging from the 70th to the 97th percentile) are significantly more proficient than the high spatials from Study 3 (who range from the 35th to the 88th percentile) ($U = 33$, $p < .01$). These groups also differ in the incidence of cancellation, a significantly lower proportion of high spatials from Study 3 adopted cancellation (see Table 4). Hence, when comparing Study 3 with Study 1, the range of spatial ability scores has been reduced due to the lack of people at the top of the range. If spatial ability is the best explanation of strategy selection, then the observed reduction in predictive power by spatial ability score for Study 3 is precisely what would have been expected. *Hence, the spatial ability explanation is a better account of strategy selection in these studies than either intelligence or knowledge.*

Past attempts to show a relationship between intelligence test score and strategy choice have been unconvincing. For example, although Haygood & Johnson (1983) and Ippel & Beem (1987) have found a relationship, Alderton & Larson (1994) have failed to replicate these findings. Theoretically, explaining performance by intelligence level also runs into problems due to the lack of current understanding of the nature of *g* (see Rabbitt, 1988). Because of this, in the absence of any attempt to suggest a mechanism, invoking intelligence would not satisfactorily explain the findings in any case, instead it would merely redescribe the observed patterns of behaviour (see Howe, 1988; Sternberg, 1988).

Table 4. Comparison of the final chosen strategies for Study 1 (university students) and Study 3 (elderly people)

	High spatial groups		Low spatial groups	
	Study 3	Study 1	Study 3	Study 1
Cancellation	13	16	4	4
Spatial	19	4	28	15

For high spatials: $\chi^2 = 7.7, p < .01$.

GENERAL DISCUSSION

Study 1 found a counter-intuitive inverted aptitude–strategy relationship for the One-person directions task and also found that high spatial ability individuals were more flexible overall than low spatial ability ones. This showed that cognitive-style accounts of strategy selection do not necessarily apply even to straightforward reasoning tasks. Hence, it is not possible to have a useful theory of strategy selection based upon the properties of strategies and the abilities of people unless the account takes note of *strategy availability* in the form of a theory of how people develop strategies or why people are in possession of them. The results also highlight the difficulties entailed in using psychometric tests to identify strategy choice. It is simplistic to assume that high spatials will always prefer to use spatial representations; they can dispense with them spontaneously when needed. Conversely, low spatials are not always more likely to dispense with spatial strategies; they may lack the necessary ability to do so unaided.

The results of Studies 2 and 3 show that for the One-person directions task, an account of strategy selection based upon spatial ability as a limiting factor in strategy development is the best explanation of the pattern of strategy selection obtained in Study 1. The superior ability of the high spatial group to reason with spatial representations meant that they were better able to develop and evaluate cancellation than the low spatial group. It is therefore suggested that in order to understand strategy selection, it is necessary to consider what a person can learn about a task from the way in which information is encoded and manipulated, and also how individual differences in the ability to carry out these processes can affect what is learned.

There are several computer modelling approaches that are intended to account for changes in human reasoning performance with practice. Examples include SOAR (Laird, Rosenbloom & Newell, 1986; Newell, 1990) and ACT* (Anderson, 1983). With these approaches, performance may be speeded when, for example, solutions to impasses are incorporated into a knowledge base, and/or when redundant steps are deleted. However, the findings of the current studies highlight several difficulties with these approaches which suggest that they are currently incomplete.

The first difficulty is that the only way in which these systems can model individual differences is by the modifying the knowledge base. The underlying assumption is that this is the only important way in which individuals differ. With these approaches, it is difficult to see how the efficiency with which the architecture itself performs could be modified in order to model differences in more general, wide-ranging abilities such as spatial ability and intelligence.

The second difficulty is that, although these systems are intended to model improvements in performance, this manifests itself as an attempt to model learning curve type improvements. While the extent to which the cancellation strategy is a compiled or a chunked spatial strategy is open to debate, there remains the problem that, for an individual, the transition from the spatial strategy to cancellation is characterized by a large increase in solution times (when strategies are being compared) followed by a sudden decrease (when cancellation is chosen). This would not resemble a learning curve by any stretch of the imagination.

The third difficulty is with the suggestion that knowledge and behaviour are modified as a result of resolving an impasse (which is a failure of the current strategy to give a satisfactory next step, e.g. Schank, 1982). This suggestion does not correspond with the results obtained here. Although no feedback as regards accuracy was given, the low spatial individuals were in general acutely aware that they were performing badly, and hence they experienced a great need to change strategy. It could be argued that every single problem was acting as an impasse to them and yet they remained with the spatial strategy. This suggests that either impasses do not drive strategy change or else that, although these persons experienced more impasses than the high spatials, they were not equipped to resolve them. Other researchers (e.g. Siegler & Crowley, 1991; VanLehn, 1991) have also found that impasse-based theories of strategy acquisition can run into difficulty and that other processes may need to be considered. It may be the case that the importance of impasses has been overemphasized in the literature as it is these situations that are most easily observed and manipulated. Setting up a problem with an impasse at which the participants will trip up—with the consequence that they must by definition change strategy at that point otherwise problem solving will fail—is a relatively straightforward task for an experimenter. By contrast, strategy development processes are completely under the control of the participants and although problems can be modified to help these processes along, their observation will be far more difficult.

Acknowledgements

This work was supported by a research studentship from the Science and Engineering Research Council while the first author was at the University of Nottingham. Study 3 was also supported by an Experimental Psychology Society study visit grant. The author would like to express his gratitude to the ACPRC at the University of Manchester for permission to use their facilities.

References

Alderton, D. L. & Larson, G. E. (1994). Cross-task consistency in strategy use and the relationship with intelligence. *Intelligence*, **18**, 47–76.

Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Barratt, E. S. (1953). An analysis of verbal reports of solving spatial problems as an aid in defining spatial factors. *Journal of Psychology*, **36**, 17–25.

Cattell, R. B. & Cattell, A. K. S. (1959). *The Culture-Fair Test*. Champaign, IL: IPAT.

Costin, D. (1988). MacLab: A Macintosh system for psychology labs. *Behavior Research Methods, Instruments and Computers*, **20**, 197–200.
Ericsson, K. A. & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, **49**, 725–747.
Ericsson, K. A. & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, **87**, 215–251.
Ericsson, K. A. & Smith, J. (1991) (Eds). *Toward a General Theory of Expertise*. Cambridge: Cambridge University Press.
French, J. W. (1965). The relationship of problem-solving styles to the factor composition of tests. *Educational and Psychological Measurement*, **25**, 9–28.
Haygood, R. C. & Johnson, D. F. (1983). Focus-shift and individual differences in the Sternberg memory-search task. *Acta Psychologica*, **53**, 129–139.
Heim, A. W. (1970). *Manual for the AH4 Group Test of General Intelligence*. Windsor: NFER-Nelson.
Howe, M. J. A. (1988). Intelligence as an explanation. *British Journal of Psychology*, **79**, 349–360.
Ippel, M. J. & Beem, A. L. (1987). A theory of antagonistic strategies. In H. DeCorte, H. Lodelwijks, R. Parmentier & P. Span (Eds), *Learning and Instruction*, vol. 1. Oxford: Pergamon.
Johnson-Laird, P. N. & Byrne, R. M. J. (1991). *Deduction*. Hove: Erlbaum.
Just, M. A. & Carpenter, P. A. (1985). Cognitive co-ordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, **92**, 137–172.
Laird, J. E., Rosenbloom, P. S. & Newell, A. (1986). Chunking in Soar: The anatomy of a general learning mechanism. *Machine Learning*, **1**, 11–46.
Larkin, J. H. & Simon, H. A. (1987). Why a diagram is (sometimes) worth more than ten thousand words. *Cognitive Science*, **11**, 65–99.
Lohman, D. F. & Kyllonen, P. C. (1983). Individual differences in solution strategy on spatial tasks. In R. F. Dillon & R. R. Schmeck (Eds), *Individual Differences in Cognition*, vol. 1. New York: Academic Press.
MacLeod, C. M., Hunt, E. B. & Mathews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior*, **17**, 493–507.
Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.
Nisbett, R. E. & Wilson, T. DeC. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, **84**, 231–259.
Rabbitt, P. M. A. (1988). Human intelligence. *Quarterly Journal of Experimental Psychology*, **40A**, 167–185.
Riding, R. J., Glass, A. & Douglas, G. (1993). Individual differences in thinking: Cognitive and neuropsychological perspectives. *Educational Psychology*, **13**, 267–280.
Rips, L. J. (1994). *The Psychology of Proof*. Cambridge, MA: MIT Press.
Roberts, M. J. (1993). Human reasoning: Deduction rules or mental models, or both? *Quarterly Journal of Experimental Psychology*, **46A**, 569–589.
Roberts, M. J. (1993). Human reasoning: Deduction rules or mental models or both? *Quarterly Journal of Experimental Psychology*, **46A**, 569–589.
Saville & Holdsworth Ltd. (1979). *Advanced Test Battery: Manual and User's Guide*. Thames Ditton, Surrey.
Schank, R. (1982). *Dynamic Memory: A Theory of Learning in Computers and People*. Cambridge: Cambridge University Press.
Shepard, R. N. & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology*, **3**, 228–243.
Siegler, R. S. & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, **46**, 606–620.
Smith, P. & Whetton, C. (1988). *General Ability Tests User's Guide*. Windsor: NFER-Nelson.
Sternberg, R. J. (1988). Explaining away intelligence: A reply to Howe. *British Journal of Psychology*, **79**, 527–533.
VanLehn, K. (1991). Rule acquisition events in the discovery of problem-solving strategies. *Cognitive Science*, **15**, 1–47.
Wood, D. J. (1978). Problem solving—the nature and development of strategies. In G. Underwood (Ed.), *Strategies in Information Processing*. London: Academic Press.

Received 27 July 1995; revised version received 22 April 1996