3

Analyses of Spatial Aptitude and Expertise

James W. Pellegrino
Randall J. Mumaw
Valerie J. Shute

INTRODUCTION

The research discussed in this chapter relates to aptitude analysis and to the analysis of technical skill competence. In both cases, the focus is on the general area of spatial cognition and processing. Spatial ability represents a major individual differences factor (see, e.g., Lohman, 1979; McGee, 1979). However, such abilities are only weakly related to typical academic achievement. There is no major curriculum area that focuses primarily on spatial cognition and processing. Nevertheless, spatial ability measures, like other aptitude and intelligence measures, have served as useful predictors of success in other environments. McGee (1979) identified two major areas where spatial tests have been utilized for prediction. One area is industry and the other involves certain academic settings and vocational–technical training programs. In the job performance area, spatial ability tests have been most useful for predicting success in engineering, drafting, and design. In academic settings and training programs, spatial tests have been most strongly correlated with performance in mechanical drawing, shop courses, art, mechanics, and to some extent mathematics and physics.

Although correlations exist between spatial aptitude measures and job or technical course performance, there is little substantive theoretical basis for explaining and understanding such relationships. Our research on spatial cognition and processing, as well as that of others (e.g., Cooper, 1982) has

ISBN 0-12-238180-7

been directed at establishing such a base. In the remainder of this chapter we first illustrate some initial attempts to understand the cognitive skills underlying both spatial aptitude and achievement. The next section focuses on efforts to design and use a systematic set of problems to analyze components of processing in a complex spatial aptitude task. The task is a standard index of spatial visualization ability that is predictive of success in technical skills courses such as engineering design and graphics. The following section discusses some initial work that examines dimensions of performance in the engineering design and graphics area. These two focuses are largely disconnected at present, although at the end of the chapter we consider some directions for linking them together.

SPATIAL APTITUDE ANALYSIS

One approach to studying the nature of aptitude is to apply cognitive process theory and methodology to the analysis of performance on tasks found on various aptitude test batteries. This so-called cognitive components approach (Pellegrino & Glaser, 1979) does not presuppose that aptitude is uniquely defined by the circumscribed performances required by tests. Aptitude or ability covers a much wider range of knowledge and skill. However, the cognitive components approach recognizes that various tests have been devised that reliably assess individual differences in cognitive abilities and that these differences are predictive of success and achievement in diverse real-world settings. The question is, then, what are the skills that are being assessed by such instruments and how can the basis for individual variation be understood? The goal is to treat the tasks found on aptitude tests as cognitive tasks (Carroll, 1976) that can be approached and analyzed in the same way that cognitive and developmental psychologists have approached and analyzed other cognitive tasks.

An initial step in a systematic analysis of individual differences in spatial aptitude is identification of the domain of tasks that serve to define it. This involves identifying a core or prototypical set of tasks that frequently occur across many widely used spatial aptitude tests and that have a history of consistent association with the spatial aptitude construct. Such an initial step delineates the task forms that should serve as the target for rational, empirical, and theoretical analysis. A multitask approach is important because an adequate understanding of individual differences in spatial ability cannot be based on an intensive analysis of only a single task with a high loading on the spatial aptitude factor(s). Rather, it is necessary to conduct analyses that consider the various tasks yielding correlated performance, and in so doing specify a set of performances that define the aptitude construct.

A successful analysis of multiple tasks should provide a basis for understanding the patterns of intercorrelations among tasks. More importantly, the analysis of multiple, related tasks should permit the differentiation of general and specific cognitive processes and knowledge. This differentiation can lead to a level of analysis where research can be pursued on the feasibility of process training and transfer.

Spatial aptitude has remained a nebulous psychometric construct even after 70 years of psychometric research. There appears to be little agreement among major studies about the number of distinct spatial abilities that may exist and how best to characterize each one. Lohman (1979) has provided an overview of some of the problems encountered in trying to integrate the major factor-analytic work that has been done on spatial aptitude. First, identical tests appear with different names in different studies, and tests given the same name are often quite different in appearance and cognitive demands. A second problem is that subtle changes in test format and administration can have major effects on the resultant factor structure. A typical change that can produce such an effect is the use of solution time as opposed to number correct as the measure of performance. Finally, perhaps the most important difference relates to procedural variation in factor extraction and rotation.

To correct for some of the problems just described, Lohman (1979) reanalyzed the data from several major studies in an attempt to isolate a common set of spatial factors. The result of these efforts was the delineation of three distinct factors. One factor, labeled *Spatial Orientation,* appeared to involve the ability to imagine how a stimulus or stimulus array would appear from another perspective. Typically, such tasks require the individual to reorient himself or herself relative to the array, as when in a plane or boat that shifts heading relative to some land mass. The other two factors were labeled *Spatial Relations* and *Spatial Visualization.* The spatial relations factor appears to involve the ability to engage rapidly and accurately in mental rotation processes that are necessary for judgments about the identity of a pair of stimuli. Spatial relations tasks can be found in test batteries such as the Primary Mental Abilities test (PMA; Thurstone & Thurstone, 1949). The spatial visualization factor is defined by tests that are relatively unspeeded and complex. Such tasks frequently require a manipulation in which there is movement among the internal parts of the stimulus configuration or the folding and unfolding of flat patterns. Spatial visualization tasks can be found in test batteries such as the Differential Aptitude Test (DAT; Bennett, Seashore, & Wesman, 1974) or as separate tests such as the Minnesota Paper Form Board Test (Likert, 1934; Likert & Quasha, 1970).

The differences between and among spatial relations and visualization tasks seem to reflect two correlated dimensions of performance (Lohman, 1979). One of these is the speed–power dimension. Individual spatial relations problems are solved more rapidly than spatial visualization problems, and the tests themselves are administered in a format emphasizing speed in the former case and both speed and accuracy in the latter case. The second dimension involves stimulus complexity. A gross index of complexity is the number of individual stimulus elements or parts that must be processed. Spatial relations problems, although varying among themselves in complexity, involve less complex stimuli than do spatial visualization problems. In terms of a process analysis of spatial aptitude, the important question is whether individual differences in performance on these various tasks reflect differential contributions of the speed and accuracy of executing specific cognitive processes.

Considerable attention has been given to the analysis of performance on spatial relations tasks (see, e.g., Pellegrino & Kail, 1982). Studies have examined sources of gender, individual, and developmental differences in performance on simple mental rotation problems. These studies have applied a process model originally developed by Cooper and Shepard (1973) for mental rotation tasks. The results are quite consistent in showing that substantial speed differences exist in the encoding and comparison of unfamiliar stimuli and in the execution of a rotation or transformation process that operates on the internal stimulus representation. Adult individual differences exist in all these components of processing and they are mirrored by overall developmental trends. The limited analyses of age changes further suggest that individual differences initially relate to encoding and comparison processes and that the rotation process subsequently becomes an increasingly important source of individual differences. A further potential source of individual differences, one that needs further analysis, involves the strategy for task execution. Systematic individual differences may also exist in the speed and criteria for judging the mismatch between stimuli in different orientations.

The differences in encoding, comparison, and rotation that exist for simple spatial relations tasks are of even greater magnitude in complex spatial relations tasks employing more abstract stimuli. The complexity and abstractness of the stimuli lead to substantial errors on such problems that are also related to individual differences in reference test scores. The particular errors that seem most important for differentiating among individuals involve the processes associated with comparing figures for differences. Latency data for different judgment performance also contribute to predicting individual differences in reference test performance. The data indicate that individuals experience considerable difficulty in establishing the correspondences between common segments of complex stimuli, leading to several

iterations through a sequence of processes and often culminating in an incorrect evaluation or guess.

The tasks associated with spatial visualization have received considerably less attention than spatial relations tasks. Relatively little has been done to develop and validate information-processing theories and models for such tasks. There are two major exceptions; these include the early work of Shepard and Feng (1972) and our own recent work (Mumaw, Pellegrino, & Glaser, 1980; Pellegrino, Cantoni, & Solter, 1981). Shepard and Feng (1972) studied performance in a mental paper-folding or surface development task. In the Shepard and Feng study, individuals were presented a representation of a flat, unfolded cube. Two of the surfaces had marked edges and the task was to decide if the marked edges would be adjacent when the pattern was folded to form the cube. The items that were used varied in the number of 90-degree folds required to bring the two marked edges together. Items were also classified by the number of surfaces that had to be carried along with each fold (i.e., the number of surfaces that had to be mentally moved to complete each new fold). Ten different stimulus values were obtained and decision times for items showed a general linear trend consistent with the total number of folds and surfaces that had to be processed to solve a problem. Shepard and Feng were not explicit about a model of performance for this task. Thus, the component processes and their sequencing are not well understood at present and no systematic process analysis of individual differences has been conducted.

ANALYSIS OF A SPATIAL VISUALIZATION TASK

Our analyses of spatial visualization involved performance on the Minnesota Paper Form Board (Likert & Quasha, 1970). Figure 1 illustrates a typical problem from this test. The individual is presented with an array of pieces and five completed figures. The task is to determine which of the five alternative choices is the correct figure that can be constructed from the particular set of pieces. Our analysis of the form board task began with a rational task analysis of the types of problems presented on this test and the dimensions that seem to underly task difficulty and errors. Items on form board tests vary in the number of individual stimulus elements that must be processed, their similarity, and the number of mismatching pieces for incorrect solutions.

From the standpoint of the cognitive processes necessary to solve an item, we hypothesized that the elementary processes included encoding, compari-

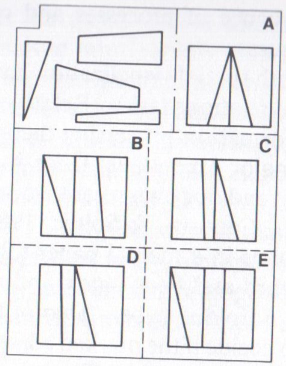

FIGURE 1. Sample paper form board item.

son, search, rotation, and decision processes. Validation of such assumptions could not be done in the context of the problems present on the actual test. This was a function of the unsystematic nature of the problems themselves. To circumvent this, we developed a variant of this task that emulated the problems and processing required by the psychometric test items. The item type that was created for our studies is shown in Fig. 2. Individual stimulus pairs were constructed consisting of a complete figure and an array of individual pieces. The stimuli were both selected from psychometric tests and constructed so they would permit the evaluation of several models of performance.

Figure 2 also shows a process model for performance on an item of this type. We assume that there is an initial encoding of one of the pieces followed by a search for a potentially corresponding piece. Given the identification of a possible match there is rotation to bring the two stimuli into congruence so that a comparison process can be executed. If the two pieces correspond and all pieces have been examined then a positive response is executed. If all pieces have not been examined then the entire process recycles for examination of another stimulus element. There are three required processes and two optional processes that depend on the nature of the stimulus type. The example problem is one presumably requiring all five processes. The search process is required because the pieces are randomly arranged and have been displaced relative to each other given their position in the completed figure. The rotation process is required because each piece has also been rotated in the picture plane in addition to being spatially displaced. Both rotation and displacement characterize items on psychometric tests. The appropriate general latency equation for such items is also shown in Fig. 2. By varying the number of stimulus elements

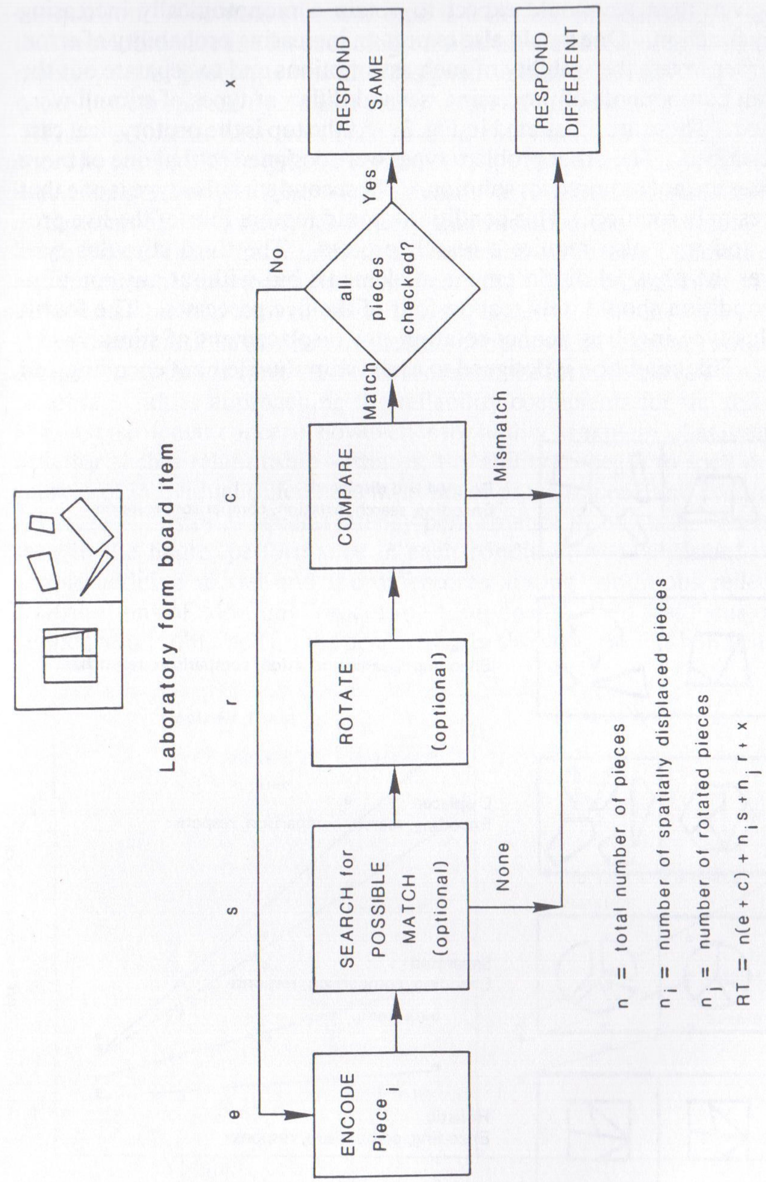

FIGURE 2. Laboratory form board item and process model for item solution. $RT = n(e + c) + n_i s + n_j r + x$, where n = total number of pieces, n_i = number of spatially displaced pieces, and n_j = number of rotated pieces.

for a given item we would expect to obtain a monotonically increasing latency function. One would also expect an increasing probability of error.

In order to test the viability of such assumptions and to separate out the different components of processing, several different types of stimuli were designed. These are illustrated in Fig. 3. At the top is the prototypical case just described. The other problem types were designed so that one or more processes are not required for solution. The second stimulus type is one that involves only rotation. This condition should require four of the five processes and may also require a search process. The third stimulus type involves the physical displacement of elements but without any rotation. This condition should only require four of the five processes. The fourth stimulus type involves neither rotation nor displacement of stimulus elements. This condition is designed to assess stimulus element encoding and

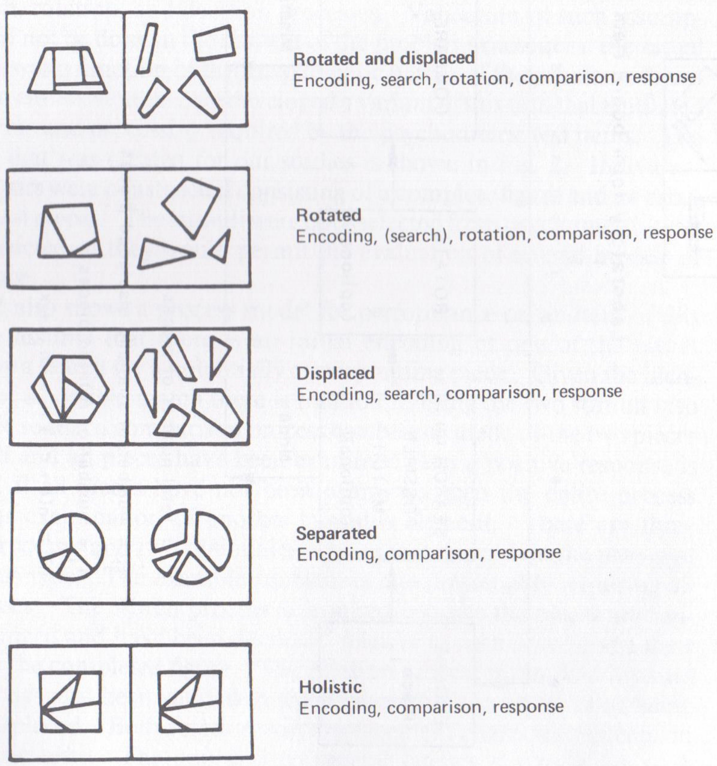

FIGURE 3. Examples of positive match items for experimental conditions differing in process complexity.

comparison. The final stimulus type is a holistic presentation condition that provides a baseline for encoding, comparison, and response. A complementary set of problems was also designed so that a mismatch existed between the completed figure and one or more of the pieces in the array.

These problems have now been used in two studies of individual differences in spatial ability (Mumaw *et al.*, 1980; Pellegrino *et al.*, 1981). The individuals tested were selected to represent varying levels of spatial aptitude as determined by a reference battery that included the Minnesota Paper Form Board Test. Each individual was tested on several hundred of the individual problems that have been illustrated. Before describing some of the results obtained with this task, it is important that we indicate that our experimental task has external validity. When overall accuracy within the experimental task is correlated with scores on the form board reference test, we obtain values approaching the reliability coefficients for the reference test. Of particular concern, however, is the ability to explain what is responsible for such a relationship—that is, the ability to begin to spell out the sources of individual differences in terms of process speed and accuracy.

Figure 4 illustrates general latency performance in our task. As can be seen in the figure, performance in each condition was consistent with a simple additive model and the differences among conditions reflect the contributions of additional processing components. We first focus on the latency results obtained for the positive trials, shown in the left-hand panel of

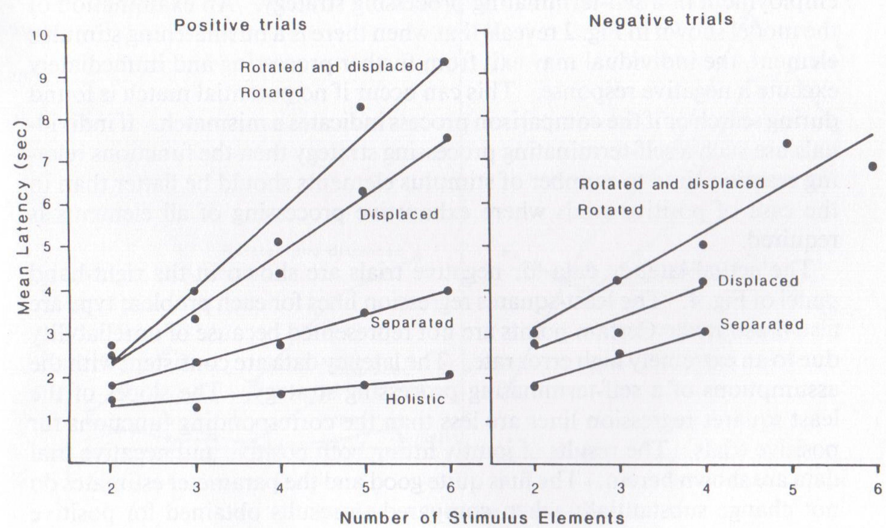

FIGURE 4. Group mean latency data for positive and negative match items.

Fig. 4. The data for the rotated and the rotated and displaced problem types have been combined because they did not differ. The linear functions shown in the figure represent the least-squares regression lines for each of the four problem types. As expected, the condition with the steepest slope was the one requiring search and rotation in addition to encoding, comparison, and response. The next-steepest slope occurred in the condition that required only search in addition to encoding, comparison, and response. The shallowest significant slope occurred in the separated condition. This condition presumably required only encoding, comparison, and response. Finally, the baseline holistic condition showed a basically flat function, as expected.

The adequacy of the model shown earlier and the assumptions about processing for each problem type were tested by simultaneously fitting the data from all conditions. When group mean data were used, the overall fit of the model was quite good and it accounted for over 96% of the variance. The values obtained for each of the individual parameters were plausible and there were no major deviations from the model. Model fitting was also done for each individual subject. Almost all subjects had R values above .90 and only three subjects had poor model fits. Thus, the model was not only representative of the group data but it also provided a good characterization of the performance of each individual.

The latency data for negative trials complement the data for positive trials and permit a test of whether performance in the task is consistent with the employment of a self-terminating processing strategy. An examination of the model shown in Fig. 2 reveals that when there is a mismatching stimulus element, the individual may exit from further processing and immediately execute a negative response. This can occur if no potential match is found during search or if the comparison process indicates a mismatch. If individuals use such a self-terminating processing strategy then the functions relating reaction time to number of stimulus elements should be flatter than in the case of positive trials where exhaustive processing of all elements is required.

The actual latency data for negative trials are shown in the right-hand panel of Fig. 4. The least-squares regression lines for each problem type are also illustrated. Certain points are not represented because of unreliability due to an extremely high error rate. The latency data are consistent with the assumptions of a self-terminating processing strategy. The slopes of the least squares regression lines are less than the corresponding functions for positive trials. The results of jointly fitting both positive and negative trial data are shown herein. The fit is quite good and the parameter estimates do not change substantially when compared to results obtained for positive trials only.

Results of Model Fitting for Positive and Negative Trials ($N = 41$)

Parameters:

Encode and compare = 556 msec/element
Rotate = 299 msec/element
Search = 689 msec/element
Preparation–response = 624 msec
Index reset (negation) = 859 msec

$R^2 = .94$
$RMSD = 545$

The error data for both positive and negative trials were also systematic and had considerable importance relative to individual differences. The error data for positive trials are shown in the left-hand section of Fig. 5. As can be seen in the figure, positive trial errors were related to the presence of the rotation component. There was a significant increase in overall errors as a function of the number of times that the rotation process needed to be executed. The other processing components, with the possible exception of search, did not systematically contribute to errors for positive trial types. Individual subjects differed substantially in error rates with an overall range

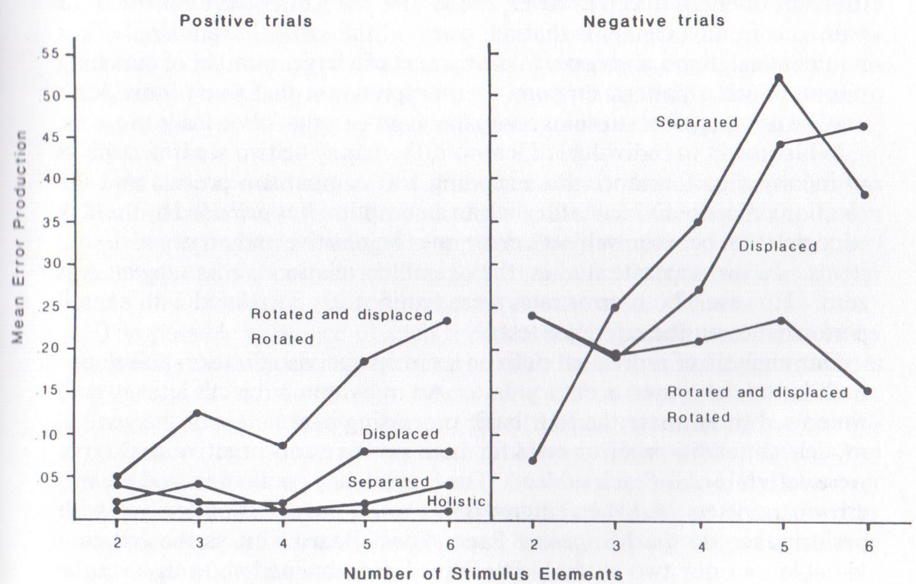

FIGURE 5. Group mean accuracy data for positive and negative match items.

of 0–23% for all positive trial types. For the problem types involving rotation, the range was 0–43% errors.

Of particular interest is the different patterning of error data for the negative trial types. The highest error rates were obtained for the conditions that did not require rotation. This is not to say, however, that the presence of rotation did not lead to errors. Rather, errors were highest when the ratio of matching to mismatching pieces was high and when processing could proceed rapidly because of the absence of a rotation component. It appears as if individuals may have processed elements superficially for comparison purposes and thereby failed to detect differences in size and shape for globally corresponding elements. The individual subject error rates on the negative trials ranged from 5% to 55%.

The error data for the positive and negative trials support the notion of different mechanisms contributing to incorrect final decisions. In the case of positive trials, errors seem to result from the inability to determine the correspondence between two stimulus elements that must be rotated into actual congruence. In such cases, either rotation is incorrectly executed or the resultant representation following rotation is imprecise, leading to a rejection of a matching element. Errors resulting from execution of the rotation process also appear in the negative trials where there is a tendency to accept the match between two similar but nonidentical pieces that are in different orientations. However, the largest error rates were obtained for pairs of stimulus elements that are oriented the same, have a similar but nonidentical shape, and occur in the context of a larger number of matching pieces. Such a pattern supports the interpretation that some individuals may be using a global stimulus comparison process that often leads to errors.

With respect to individual differences, there may be two separate aspects of incorrect performance, the encoding and comparison process and the rotation process. Evidence for such an assumption was provided by the lack of correlation between subject's error rates on positive and on negative trial items. In two separate studies, the overall correlation across subjects was zero. However, both error rates were significantly correlated with overall performance on the reference test.

Our analysis of individual differences in spatial visualization ability utilized both latency and accuracy data. An individual subject's latency data were used to estimate the four basic processing parameters of the general model. In addition, error rates for both positive and negative trial types were determined for each subject. The four latency parameters and the two error parameters were then entered into a multiple regression analysis with performance on the Minnesota Paper Form Board Test as the criterion variable. In our two studies, data have been obtained showing that the combination of both error data and latency parameters accounts for over

60% of the variance in spatial ability as defined by reference test performance. Based on individual subject model fitting and error data we have reached the following conclusions about ability differences. First, skilled individuals make fewer errors on problems involving stimulus rotation or transformation. Second, they are also more accurate in detecting mismatches between similar stimuli independent of the occurrence of rotation. Third, skilled individuals are faster at searching through an array to find corresponding stimulus elements. Fourth, they are faster in encoding and comparison processes. A general conclusion based on these types of findings is that skill in a visualization task such as the form board is related to the speed and quality of stimulus representation processes. A more precise representation of stimulus elements permits more rapid search for a corresponding element and a faster and more accurate decision about their identity.

Because ability appears to be associated with several aspects of task performance involving both speed of processing and representation, we explored the possibility that subjects also differed in their approach to processing different problem types. In particular, we were intrigued by the possibility that some subjects may adopt a very precise analytic mode of processing consistent with the general model shown earlier. Other subjects may be performing the task in a way that represents an analytic or semianalytic mode of processing on some item types and a more global, holistic mode of processing on other item types. The particular condition of interest relative to identifying such a mixture of strategies is the separated condition. It is possible to imagine a strategy in which the individual attempts to merge together or fuse the individual stimulus elements into a whole figure, which is then matched against the completed stimulus. Such a strategy should have two consequences. First, the latency function for such items should be relatively flat. Second, on negative separated trials the fusion process may result in forcing all the pieces into a whole, even though one of the elements is incorrect with respect to specific features such as size or angle. The result would be a higher error rate on such items.

The possible existence of such a processing approach was explored by separating the subjects into two groups solely on the basis of their error rates on the separated different item types. The 12 low-error subjects had fewer than 20% errors on these problems. The 18 high-error subjects had considerably more than 20% errors on these problems. This partition revealed an interesting pattern of skill differences. As shown in Table 1, the low-error group was primarily composed of high-skill individuals as determined by reference test performance. In contrast, the high-error group was primarily composed of medium- and low-skill individuals.

Of interest is whether these two groups of subjects also show different

TABLE 1
SUBJECT DISTRIBUTION AS A FUNCTION OF ABILITY LEVEL AND ERROR RATES

	Spatial ability		
Separated item errors	High	Medium	Low
High	1	8	9
Low	9	2	1

latency patterns for the positive trial data. The left-hand panel of Fig. 6 shows the data for the low-error group. An important point to note is the fanning of the slopes consistent with the general model of processing, reflecting an analytic-processing approach. Second, the slope for the separated condition is relatively steep and its value is approximately 600 msec per stimulus element. The right-hand panel of Fig. 6 shows the data for the high-error group. There are two important points to notice. First, the conditions involving rotation show no differentiation from the condition that does not involve rotation. Second, the slope for the separated condition is shallower, with a value of 375 msec.

We think that these results, when combined with the error data on negative items, support the hypothesis that skill differences are also associated with differences in processing strategy. The major strategy differences may involve a precise analytic mode of processing versus a less analytic and partly holistic or global processing mode for certain item types. With respect to the latter, less-skilled subjects show shallow latency functions for separated item types with very high error rates when an incorrect element is embedded in this stimulus format. They also fail to show evidence of executing a rotation process, with substantially higher error rates on rotation items when an incorrect or mismatching element is in the array. Our work to date on individual differences in spatial visualization supports the existence of speed and accuracy differences associated with encoding and representational processes. In addition, it suggests strategy differences reflecting less analytic modes of processing in lower-skill individuals. Further, validation of such strategy differences requires more precise methods of analysis of the type possible with eye-movement data.

The preceding overview of some of our spatial aptitude research emphasizes that it is possible to construct and validate cognitive process models for performance in complex spatial visualization tasks. Such models not only provide a good characterization of group performance but they also capture

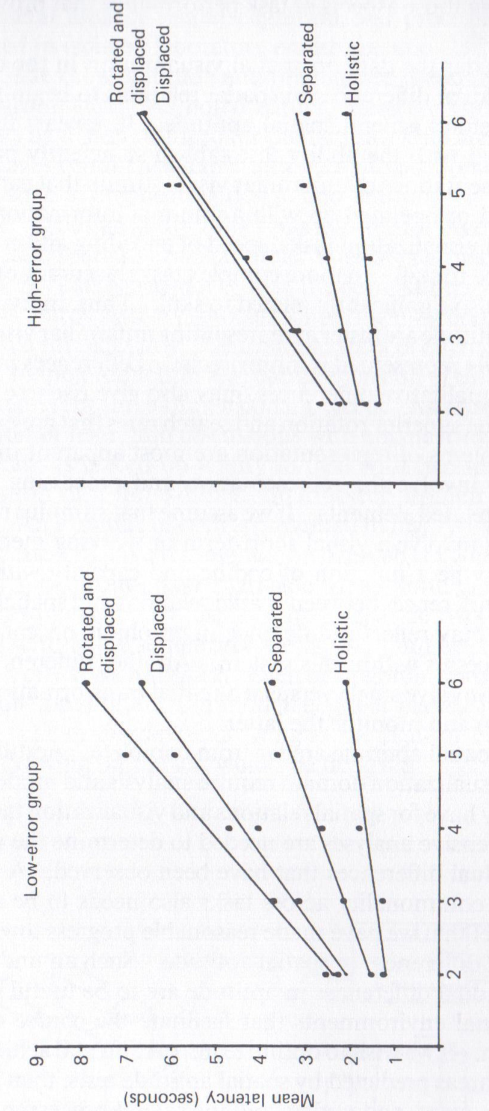

FIGURE 6. Positive trial latency data for low- and high-error subjects.

the performance of individual subjects. The problem variants provide a means of determining those aspects of task performance that prove difficult for a given individual.

We can also consider the data on spatial visualization in the context of other data on individual differences in spatial relations to begin to address the issue of what defines general spatial aptitude. It appears that spatial aptitude is associated with the ability to establish sufficiently precise and stable internal representations of unfamiliar visual stimuli that can be subsequently transformed or operated on with a minimal information loss. In spatial relations and visualization tasks, speed of encoding and comparison is significantly related to skill. In more complex tasks, accuracy of encoding and comparison is also significantly related to skill. Thus, individuals who are high in spatial aptitude are faster at representing unfamiliar visual stimuli and what is ultimately represented is more precise. Differences in representation, most likely qualitative differences, may also give rise to other speed differences such as the superior rotation and search rates that are exhibited in various tasks. Problems of representation are most apparent in the more complex tasks that involve the representation and processing of stimuli having several interrelated elements. If we assume that stimulus representation and processing involve a visual short-term or working memory, then skill differences may be a function of coding and capacity within such a memory system. Differences between spatial relations and spatial visualization tasks (factors) may reflect a difference in emphasis on coding versus transformation processes within this system. Another difference between the two factors may involve single versus sequential transformations and the ability to coordinate and monitor the latter.

Our analyses of spatial aptitude are far from complete. Additional tasks within the spatial visualization domain require analysis and modeling. The models we currently have for spatial relations and visualization tasks require refinement, and intensive analyses are needed to determine the underlying bases of the individual differences that have been observed. A systematic analysis of process commonality across tasks also needs to be attempted. Nevertheless, we feel that we have made reasonable progress toward understanding individual differences in spatial aptitude. Such an understanding is essential if individual differences in aptitude are to be useful in creating adaptive instructional environments that facilitate the course of learning and skill acquisition. If we wish to optimize instruction and achievement in the technical skills areas predicted by spatial aptitude tests, then we need to have a better way to assess and understand the spatial-processing skills that individuals bring to the instructional setting, and the impact that these skills can have on the design of instruction and the acquisition of skill.

ENGINEERING DESIGN AND GRAPHICS

The general area of spatial cognition and processing has been largely investigated in isolated laboratory environments. However, spatial ability measures, like other aptitude and intelligence measures, predict success in other environments. As noted earlier, there are two major areas where spatial tests have been utilized for prediction. One area is industry and the other involves certain academic settings and vocational–technical training programs such as engineering design and drafting.

Engineering design and graphics was selected for investigation because it is an area where spatial visualization abilities appear necessary for achieving competence. This is true not only in the correlational literature, but also in individuals' retrospective protocols for performing certain tasks, which will be illustrated subsequently. One major aspect of engineering design courses is training in the production and comprehension of different types of drawings that represent three-dimensional objects. Analyses of course content, examination of texts, and discussions with engineering instructors indicate that emphasis is placed on ability to deal with two major forms of visual representation of a three-dimensional object: isometric and orthographic drawings. An example of each is presented in Fig. 7. An isometric drawing is shown at the left; it is a representation of an object where the viewing angle shows the top-, front-, and right-side views simultaneously. An orthographic drawing generally displays three separate two-dimentional projections representing the same three views, as would be seen by an individual looking directly at each surface. Each orthographic projection also includes information about internal or hidden edges or planes, indicated by the dotted lines. The orthographic drawing at the right of Fig. 7 shows how the isometric drawing on the left would be depicted.

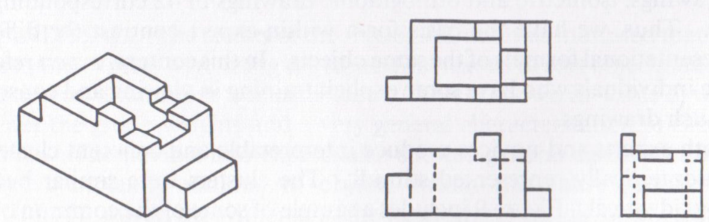

FIGURE 7. Examples of isometric and orthographic drawings of a three-dimensional object.

Cognitive Representation of Engineering Drawings

Two preliminary studies have been conducted that examined issues of representation and comprehension of isometric and orthographic drawings. Because these were initial, probing studies, the questions we sought to answer were relatively simple and were not process oriented. Rather, we focused on trying to understand more about these two different forms of representation and stimulus complexity within and across representations. Initially, an engineering instructor helped us develop a set of stimuli he deemed representative of materials that individuals deal with in such courses. The set of stimuli he created varied considerably but not systematically in complexity, ranging from easy to medium to difficult in ease of processing. Figure 8 contains a representative sample of the orthographic and isometric drawings in the total stimulus set.

The two studies employed similar data collection procedures and analyses. In both cases, a set of drawings was given to individuals who were instructed to do two things. First, they sorted them into groups according to level of object complexity. There were no limitations placed on the number of groups to be formed, yet most of the individuals created four or five groups. Second, they also rank-ordered all objects within and across groups with respect to object complexity. Thus, we could examine the features that contribute to perceived complexity and the groups or clusters that were formed. In the first study, two different subject groups were compared on their ranking and sorting behavior for a set of 48 isometric drawings. The subjects were 29 individuals without any experience in engineering design and graphics courses and 21 students who were at the end of a one-semester course in this area. This provides the basis for a novice–expert contrast. In the second study, a different group of 26 students was tested who were also nearing completion of a one-semester engineering design and graphics course. These subjects performed the ranking and sorting task with two sets of drawings: isometric and orthographic drawings of 42 corresponding objects. Thus, we have the basis for a within-expert contrast for different representational formats of the same objects. In this context, *expert* refers to those individuals who have some explicit training in viewing and constructing such drawings.

Both experts and novices produce interpretable and coherent clusters of the isometrically represented stimuli. The clusters were similar but not always identical. Figure 9 includes a sample of some of the common object clusters formed by both groups. For some clusters, the specific characteristics defining the group can be readily identified and labeled. Multidimensional scaling analysis of the novice and expert sorting data was conducted by

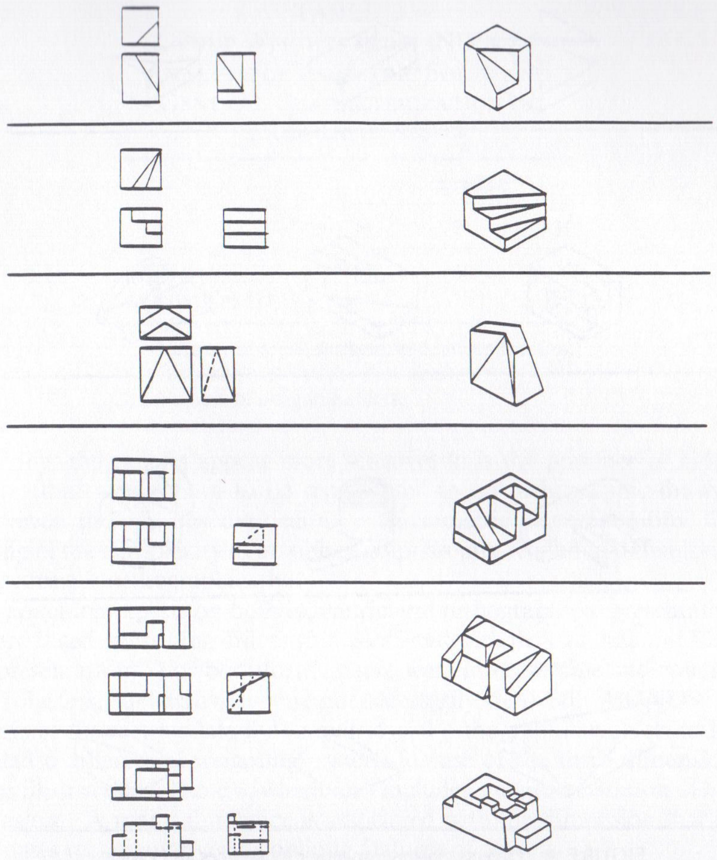

FIGURE 8. Representative stimuli used in sorting and ranking studies.

using INDSCAL. The results of the INDSCAL analysis indicated that both novices and experts were using the same dimensions to sort objects, and the weighting on dimensions was equivalent. This is shown in Table 2, which provides the group weights and a very general characterization of each dimension. The novices and experts also showed general agreement in the average complexity ranking of objects within the entire problem set ($r = .70$). However, they differed in the weighting of certain object features with respect to the evaluation of object complexity. The nature of this difference is systematic. The coordinate values of objects along each of the three dimensions of the INDSCAL analysis were used in multiple regression

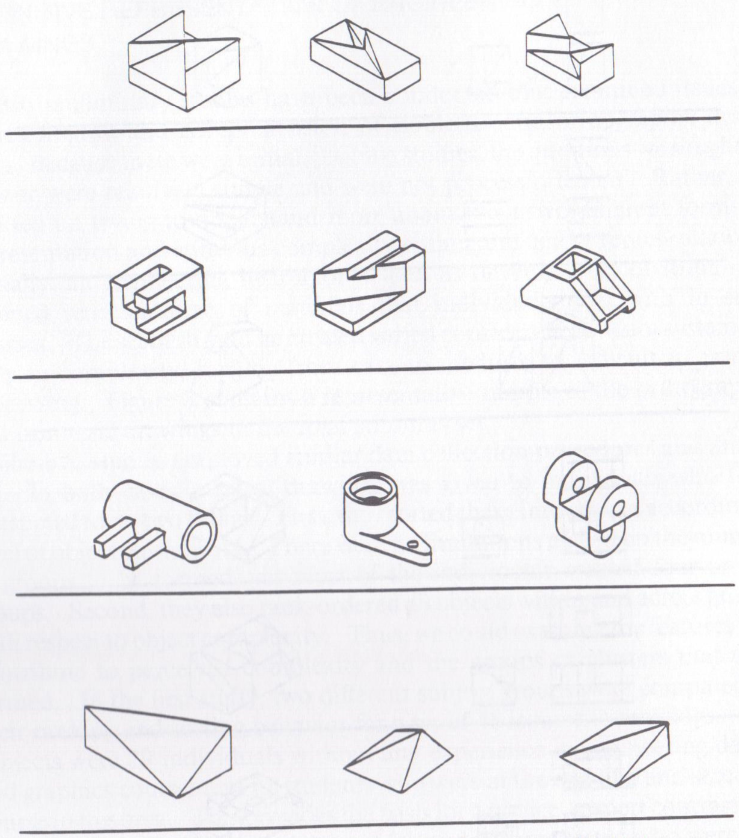

FIGURE 9. Objects clustered together by novices and experts.

analyses to predict rated object complexity for each subject group. These multiple regression analyses produced different patterns for the two groups, shown in Table 3. The major difference between novices and experts was the relative importance of Dimension 2 in predicting object complexity. This dimension reflects a partition of objects such that those with curved and oblique surfaces are at one extreme and those with simple rectangular and right-angle features are at the other extreme. Objects that have multiple oblique surfaces and those with curved surfaces are systematically rated more complex by the experts than novices. This can be understood by considering that these features of an object are the most problematic in creating an orthographic representation of an isometric drawing. A related

TABLE 2
GROUP WEIGHTS FROM INDSCAL ANALYSIS OF ISOMETRIC SORTS AND GENERAL CHARACTERIZATION OF EACH DIMENSION

	Dimension		
	1[a]	2[b]	3[c]
Novices ($N = 29$)	.61	.50	.38
Experts ($N = 21$)	.54	.53	.37

[a] Number of edges, surfaces, and hidden features.
[b] Right angles (blocklike structures).
[c] Number of oblique surfaces.

factor that the experts appear more sensitive to is the presence of hidden features that would have to be represented in an orthographic drawing. Thus, even though they are dealing with isometric representations, their ranking of the complexity of an object is apparently related to difficulties of representing it orthographically.

The data for experts on both isometric and orthographic representations also produced interesting differences associated with the two different forms of representation. For both forms, there were interpretable and coherent object clusters, but again they were not necessarily identical. An INDSCAL analysis of the sorting data for isometric and orthographic representations revealed a differential weighting pattern for use of the three dimensions. This is illustrated in Table 4, which also includes a characterization of each dimension. A major difference is associated with the dimension that discriminates rectangular versus oblique features of objects.

TABLE 3
MULTIPLE REGRESSION ANALYSIS OF NOVICES' AND EXPERTS' RANKING DATA

	Novices		Experts	
Predictor	F	β	F	β
Dimension 1	290.5	.88	61.5	.67
Dimension 2	2.8	.08	34.2	.50
Dimension 3	21.6	.24	4.4	.18
	$R^2 = .89$		$R^2 = .69$	

TABLE 4
DIMENSION WEIGHTS FROM INDSCAL ANALYSIS OF ISOMETRIC AND ORTHOGRAPHIC SORTING DATA AND GENERAL CHARACTERIZATION OF EACH DIMENSION

	Dimension		
	1[a]	2[b]	3[c]
Isometric	.36	.33	.39
Orthographic	.86	.31	.10

[a] Number of edges, surfaces.
[b] Number of hidden and internal features.
[c] Rectangular versus oblique features.

A subsequent correlational analysis using the INDSCAL dimension coordinates to predict rated complexity within each type of representation confirmed the fact that these dimensions of objects were also differentially salient in judging object complexity. These data are shown in Table 5. Hidden features and oblique surfaces are weighted more highly in the rankings for the isometrically represented stimuli. This suggests that for orthographic drawings, the presence of these features may not be detected because of a failure to integrate or difficulty in integrating all three projections to create a composite mental image of the actual object. The predominant factor for making judgments about objects represented orthographically seems to be the number of surfaces or edges that are illustrated in the three separate views.

Our data analyses suggest several directions for process- and strategy-oriented research. One direction focuses on the process of mapping across different forms of representation. The speed, accuracy, and strategy of determining correspondence between an isometric and orthographic representation should vary with object complexity. Of interest is how specific types of features affect the mapping process. In addition, the orthographic representations are of interest by themselves because there should be systematic differences in the speed and accuracy of constructing composite mental images of objects given the information contained in the individual projections. A related issue involves the processes employed to determine consistency and correspondence of features represented in each of the three separate orthographic views. This seems to involve various spatial comprehension and inference skills.

TABLE 5
MULTIPLE REGRESSION ANALYSIS OF ISOMETRIC AND ORTHOGRAPHIC RANKING DATA

	Isometric		Orthographic	
Predictor	F	β	F	β
Dimension 1	29.0	.48	314.6	.90
Dimension 2	11.3	.30	5.8	.12
Dimension 3	36.7	.51	3.4	.09
	$R^2 = .73$		$R^2 = .91$	

STRATEGIES FOR PROCESSING ENGINEERING DRAWINGS

An initial study has been conducted that explores some of the issues discussed (Mumaw, Cooper, & Glaser, 1982). Three forms of representation—isometric drawing (ID), orthographic drawing (OD), and verbal description (V)—were combined to produce three problem types that were used in a problem-solving task. The subject's goal in the task was to determine if the representations shown on two separate slides depicted the same three-dimensional object or were compatible in some way. One problem type was created by pairing orthographic and isometric drawings in the serial order OD–ID. A second problem type was created by pairing a verbal description with an orthographic drawing in the serial order V–OD. For the third problem type (OD), subjects were first shown a slide containing three orthographic projections with one view replaced by an empty frame. The second slide showed a possible third view for that orthographic drawing. The subject had to determine whether the third view given was compatible with the first two views.

Thirty problems were presented for each of the three engineering drawing problem types and given to 28 subjects (14 high spatial aptitude and 14 low spatial aptitude). The subjects were given a maximum of 60, 80, or 100 seconds for the OD–ID, OD, and V–OD tasks, respectively. In addition, each subject responded to questions about solution strategy and problem difficulty after each set of 15 problems. Several other data sources were available for the subjects in this experiment. These included measures of course performance—both an overall course grade and a separate laboratory grade based on drawing assignments—and standardized measures of spatial aptitude. The latter were the spatial subtests of both the Primary

Mental Abilities battery (PMA; Thurstone & Thurstone, 1949) and the Differential Aptitude Test (DAT; Bennett *et al.*, 1974). As noted earlier, the PMA test emphasizes spatial relations ability and the DAT emphasizes spatial visualization. In addition, both verbal and quantitative Scholastic Aptitude Test (SAT) scores were available for all subjects.

The more detailed aspects of performance on the three problem types will be examined, but first it is important to consider the general relationships obtained among aptitude, course grades, and task performance. Correlations were computed between each of the various data sources available for each subject. These results are summarized in Table 6. There are three important groups of correlations that bear consideration. First, the spatial aptitude measures are significantly correlated with both overall course performance and laboratory grades. These data replicate typical findings showing that spatial aptitude predicts performance in engineering design and graphics courses. Of interest is that the correlations are higher for the measure of spatial visualization ability. Second, performance on each of the three engineering drawing problem types is significantly correlated with measures of course performance. These results provide an external validation of the experimental problem types. The differential pattern of correlations across problem types is also of significance and will be discussed subsequently. Third, the correlational data reveal a consistent pattern that addresses the relationship between performance on the engineering drawing tasks and spatial aptitude. First, OD and V–OD task accuracy show significant correlations with both aptitude tests, the OD task at the .05 level and the V–OD task at the .01 level. Accuracy on the OD–ID task, however, is not significantly correlated with either of the aptitude tests. In addition both the OD and the V–OD tasks were significantly correlated with the Quantitative SAT scores of the subjects, but not the Verbal SAT scores. The V–OD task was much more strongly related to spatial and mathematics test scores than to verbal test scores, suggesting that reading and verbal comprehension skills are not as important for this task as are spatial and quantitative reasoning skills.

The experimental procedure provided a rich data base for investigating aspects of problem-solving performance. The subject was allowed to control alternation between the two problem slides until the correct response was determined. Thus, a number of dependent measures were obtained for each trial, including number of alternations between slides, initial viewing time for Slide 1, initial viewing time for Slide 2, total solution time, and accuracy. These data together with retrospective protocols were extremely useful in identifying the solution strategy each subject used for each problem type. Because complex spatial tasks are often susceptible to several solution strategies, both spatial and nonspatial (Lohman, 1979), each subject was

TABLE 6
INTERCORRELATIONS OF SUBJECT PERFORMANCE INDICES

	Course grade	Laboratory grade	Task accuracy		
			OD–ID	OD	V–OD
PMA	.40*	.30	.11	.34	.40
DAT	.58**	.48**	.23	.41*	.54**
Course grade			.40*	.49**	.59**
Laboratory grade			.32	.28	.45*

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

asked to describe the strategy that had been used. These retrospective protocols provided the first clues to the types of strategy differences that exist between and within subjects. One general difference found in solution strategies can be characterized as the need to construct an isometric representation mentally to mediate problem solving. For instance, a subject may read an entire verbal description and try to imagine the three-dimensional object mentally before viewing the second slide. The need to construct this mediating representation to integrate information may have a large role in determining the spatial requirements of a task. The alternative solution strategy for many subjects was an analytic feature-matching strategy, which requires identifying and comparing local features of the representations. Tasks that are more susceptible to this feature-match strategy may be less related to spatial aptitude because less spatial integration is demanded.

An analysis of the engineering drawing task accuracy data as a function of aptitude and strategy provided data about the spatial demands of each task, which complement conclusions drawn from the correlational analysis. Two levels of both spatial aptitude (high and low) and solution strategy (constructive and analytic) were used to examine task accuracy. The determination of solution strategies was based on specific patterns in the dependent measures and retrospective verbal protocols. As stated earlier, some subjects claimed to construct the isometric representation mentally to mediate problem solution for each of the three problem types. The alternative strategy involved the comparison of local features extracted from the two-dimensional orthographic projections. The number of alternations between slides and the ratio of the time for the first viewing of Slide 1 to total solution time were measures that best indicated a subject's strategy. Subjects who reported a constructive strategy spent a large percentage of their total time on the first viewing of Slide 1. This was assumed to reflect construction time, and they subsequently required few alternations between slides prior to final solution. The analytic or feature-match subjects, on the other hand, had a lower time ratio and used more alternations between slides in order to make several local feature comparisons. The general distinction allowed the classification of most of the 28 subjects for each task. The remaining subjects were either inconsistent in using a certain strategy or used a strategy combining aspects of both construction and feature matching.

For the OD problem type, 24 of the 28 subjects consistently fit one of the two solution strategy patterns. Thirteen subjects were classified as constructive and 11 as analytic. Comparisons of problem-solving accuracy for these two groups and for aptitude groups revealed no main effects. However, when subjects were grouped by both aptitude and strategy, an interaction was obtained whereby low-aptitude subjects using the constructive strategy showed the poorest performance. This pattern was emphasized more strongly and became more interpretable when subject groups were further broken down based on the total time to solution, as shown in Table 7. Subjects who had an average solution time of less than 50 seconds were placed in the fast cells. Those subjects who took longer to solve the items were placed in the slow cells.

There are two important findings to note in Table 7. The first concerns the relationship between aptitude level and strategy selection: 7 of the 11 high-aptitude subjects chose the analytic strategy and 9 of the 13 low-aptitude subjects chose the constructive strategy. This trend is in the opposite direction to what might be expected. The second important result is the occurrence of a significant interaction revealing an incompatibility between aptitude and strategy. Contrasting the performance of the constructive, high-aptitude subjects with the low-aptitude subjects using the same strategy, the constructive, low-aptitude subjects, who worked at the same speed, had an accuracy level that was half that of the constructive, high-aptitude subjects. The constructive, low-aptitude subjects, who took twice as long to solve the problems, however, obtained the same level of performance shown by the constructive, high-aptitude subjects. However, these trends were not found for those subjects using an analytic strategy. The performance of the low-aptitude subjects was identical to the high-aptitude subjects in accuracy and solution time. This pattern of results suggests that there are some limits on the efficiency with which low-aptitude subjects can use the constructive strategy. For these subjects to perform as accurately as the high-aptitude subjects, they must use twice as much time. In addition, when low-aptitude subjects use a strategy that does not seem to depend on mentally constructing an isometric representation, they are able to perform as efficiently as the high spatial aptitude subjects.

The results for the V–OD problems were more straightforward. The analysis of the dependent measures showed that 18 of the 28 subjects consistently fit one of the two strategies. As Table 8 shows, the distribution was quite similar to the OD task data. Five of the nine high-aptitude subjects chose an analytic strategy and six of the nine low-aptitude subjects chose a constructive strategy. The accuracy data also show a pattern similar to the OD task data. The subjects showing the poorest performance were the constructive, low-aptitude subjects. The subjects who performed best were the constructive, high-aptitude subjects. In addition, the accuracy difference between the high- and low-aptitude individuals using the analytic strategy was not substantial. Therefore, like the OD task data, the strategy-by-aptitude interaction, though not significant, was in a direction suggesting that aptitude differences are stronger for the subjects using a constructive strategy. However, unlike the OD task data, total solution time was unimportant and there was a significant main effect due to spatial aptitude. This main effect concurs with the results of the correlational analyses.

The OD–ID task data were the least systematic of the three tasks in two respects. First, solution strategies reported by the subjects were more varied and less consistently used for this task. Only five subjects consistently used a constructive strategy and eight subjects used an analytic strategy. In addi-

TABLE 7
PERFORMANCE ON OD PROBLEMS AS A FUNCTION OF APTITUDE AND STRATEGY

		Constructive		Analytic	
		%	*N*	%	*N*
High spatial	Slow	—		70	1
	Fast	62	4	55	6
Low spatial	Slow	59	4	—	
	Fast	41	5	56	4

TABLE 8
PERFORMANCE ON V–OD PROBLEMS AS A FUNCTION OF APTITUDE AND STRATEGY

	Constructive		Analytic	
	%	*N*	%	*N*
High spatial	73	4	64	5
Low spatial	54	6	61	3

TABLE 9
PERFORMANCE ON OD–ID PROBLEMS AS A FUNCTION OF APTITUDE AND STRATEGY

	Constructive		Analytic	
	%	N	%	N
High spatial	82	3	76	6
Low spatial	77	2	83	5

tion, three subjects used what might be called a reversed strategy. This strategy was very similar to an analytic strategy except that these subjects moved very quickly to the isometric drawing on the second slide and imaged a single projection for comparison with the corresponding projection on Slide 1. Secondly, as Table 9 shows, neither strategy nor spatial aptitude was related to task accuracy. When task accuracy was averaged for the three levels of strategy and two levels of aptitude, no main effects or interactions were found.

These task-specific findings corroborate the results found in the correlational analyses: The V–OD task data show a strong aptitude effect, the OD–ID task data show no aptitude effect, and the OD task data represent a complex interaction in which the effects of aptitude are dependent on the solution strategy a subject chooses. These results, though based on small sample sizes and post hoc determinations of strategy, suggest that the "spatialness" of a task may depend on the need to mediate problem solving with a representation not directly available in the stimuli. The V–OD task seems most dependent on this mediating representation and the OD–ID task seems least dependent on other representations. This interpretation implies that a solution strategy utilizing only information from the two-dimensional representations should not be successful for the OD and V–OD tasks. To accommodate this implication, it may be necessary to recast the analytic and constructive solution strategies in terms of predictive versus facilitative use of an isometric representation. The subject using an analytic strategy may consider parts of the isometric representation to compare single features but never consider the entire object at once. A constructive strategy, on the other hand, may imply that the subject considers the entire isometric representation to make predictions about what will be present on the second slide. In this latter case, the isometric representation is used to predict details of the second representation. In the case of the analytic strategy, the

subject only needs to construct the key local parts of the object to facilitate feature comparison.

Another goal of this work is to understand the task domain better, which requires, in part, discovering the characteristics of items that affect the difficulty and/or spatial demands of the task. The preceding analysis suggests that making an item more dependent on mediation through an isometric representation should make the item more spatial. Other data from this experiment also address this question. By isolating OD items on which the constructive strategy group outperformed the analytic group, commonalities were found suggesting manipulations that can increase an item's dependence on mediation. The constructives were more accurate on items in which there was little overlap of local features between Slide 1 and Slide 2. (Figure 10 provides an example.) That is, it is necessary to construct the isometric drawing to determine the compatibility of the three views. Note that the second slide reveals an angled plane that is not shown on Slide 1.

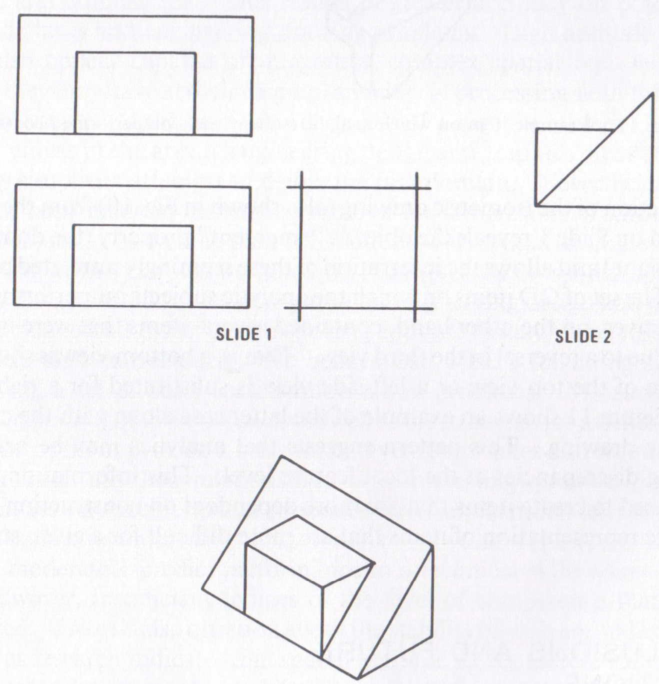

FIGURE 10. Example item on which constructive strategy users outperformed analytic strategy users.

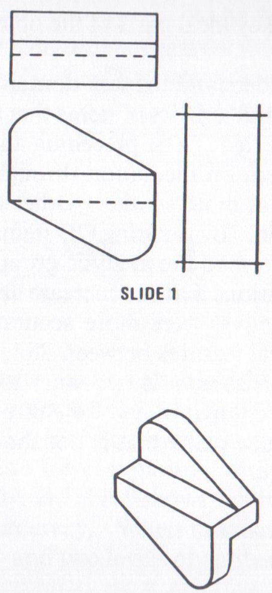

FIGURE 11. Example item on which analytic strategy users outperformed constructive strategy users.

Construction of the isometric drawing (also shown in Fig. 10) from the views provided on Slide 1 reveals the object's "emergent" property (the dominant angled plane) and allows the integration of these seemingly unrelated projections. The set of OD items on which the analytic subjects outperformed the constructives, on the other hand, contained several items that were incompatible due to a reversal in the third view. That is, a bottom view is shown in the place of the top view or a left-side view is substituted for a right-side view. Figure 11 shows an example of the latter case along with the correct isometric drawing. This pattern suggests that analytics may be better at detecting discrepancies at the local feature level. This information, then, can be used to create items that are more dependent on construction of the isometric representation of items that are more difficult for a given strategy user.

CONCLUSIONS AND FUTURE DIRECTIONS

In the preceding two sections we have provided an overview of efforts to analyze spatial aptitude and competence. Our efforts in this area are only at

the initial level of analysis and understanding. In the area of individual differences in spatial aptitude, process-oriented analyses of performance on spatial relations and visualization tasks have taken us beyond the stage of simple test score differences. We are beginning to localize the specific sources of individual differences in reference test scores. Doing so depends on being able to create systematic sets of problems that can be related to a theory or model of task performance. Individuals differ in the speed and accuracy of executing specific mental processes associated with visual–spatial stimuli. An important issue is that they also differ in the strategies they use to solve simple and complex problems. There are both between-individual and within-individual strategy differences. One interesting but very tentative result is that high-spatial individuals appear to be more precise and analytic in solving problems found on aptitude tests and this is also true in solving engineering drawing problems. *Analytic* does not imply use of a nonspatial strategy or process. Rather, it appears to involve a spatial–analytic mode of processing in which spatial detail and precision are maintained and emphasized. This results in greater accuracy on both spatial aptitude tasks and engineering drawing problems. High aptitude individuals also appear capable of integrating complex spatial representations. Thus, they may have at their disposal a variety of processing skills that can be flexibly adapted to the particular demands of a spatial problem.

Our efforts in the area of engineering design and graphics must be viewed as very simplistic attempts to define the task domain. Nevertheless, progress has been made in identifying some of the dimensions of stimulus and task complexity as well as their interactions with spatial aptitude. At the beginning of this chapter we noted that the research on spatial aptitude is largely disconnected from the research on engineering design and graphics. A major concern for the future is linking aptitude research with research on technical skill competence. We believe that a basis for doing so lies in specifying a set of basic spatial information processes. Measures of the speed and power of these processes as well as the strategies for assembling and monitoring them provide a basis for analyzing aptitudes and technical task performance and their relationships. An interesting question is whether deficiencies on specific spatial processes have implications for the acquisition of certain technical skill competencies. As we have shown, aptitude test scores moderately predict performance in a technical skills course. They are, however, insufficient indices of the level of competence that can be achieved. There is also question about the stability of such aptitudes. Some previous research indicates that spatial aptitude scores increase after taking engineering design, mechanical drawing, and drafting courses. It is possible that individuals often have little experience with the type of spatial information processing examined on aptitude tasks. They can, however, gain proficiency in such skills through practice and training. Some of our own very recent research has shown that low spatial ability individuals who participate in several spatial-processing sessions show dramatic increases in performance on standardized aptitude tests (Pellegrino, 1984). Thus, although aptitude may predict achievement it also does not necessarily preclude achievement in the spatial domain. These and other issues can be addressed by continuing to develop and apply a process-oriented approach in the area of spatial cognition.

REFERENCES

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1974). *Differential aptitude test.* New York: The Pyschological Corporation.

Carroll, J. B. (1976). Psychometric tests as cognitive tasks: A new "structure of intellect." In L. B. Resnick (Ed.), *The nature of intelligence.* Hillsdale, NJ: Erlbaum.

Cooper, L. A. (1982). Strategies for visual comparison and representation: Individual differences. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, NJ: Erlbaum.

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing.* New York: Academic Press.

Likert, R. (1934). A multiple-choice revision of the Minnesota Paper Form Board. *Psychological Bulletin, 31,* 674.

Likert, R., & Quasha, W. H. (1970). *Manual for the Revised Minnesota Paper Form Board Test.* New York: The Psychological Corporation.

Lohman, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature* (Tech. Rep. No. 8). Stanford, CA: Aptitude Research Project, School of Education, Stanford University.

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, generic, hormonal, and neurological influences. *Psychological Bulletin, 86*(5), 889–918.

Mumaw, R. J., Cooper, L. A., & Glaser, R. (1982, May). *Individual differences in complex spatial problem solving: Aptitude and strategy effects.* Paper presented at annual meetings of the Midwestern Psychological Association, Minneapolis, MN.

Mumaw, R. J., Pellegrino, J. W., & Glaser, R. (1980, November). *Some puzzling aspects of spatial ability.* Paper presented at annual meetings of Psychonomic Society, St. Louis, MO.

Pellegrino, J. W. (1984, April) *Information processing and intellectual ability.* Paper presented at the annual meetings of the American Educational Research Association, New Orleans, LA.

Pellegrino, J. W., Cantoni, V. J., & Solter, A. (1981, November). *Speed, accuracy and strategy differences in spatial processing.* Paper presented at the annual meetings of the Psychonomic Society, Philadelphia, PA.

Pellegrino, J. W., & Glaser, R. (1979). Cognitive components and correlates in the analysis of individual differences. *Intelligence, 3,* 187–214.

Pellegrino, J. W., & Kail, R. V. (1982). Process analyses of spatial aptitude. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, NJ: Erlbaum.

Shepard, R. N., & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology, 3,* 228–243.

Thurstone, L. L., & Thurstone, T. G., (1949). *Manual for the SRA Primary Mental Abilities.* Chicago: Science Research Associates.