Spatial Ability and G

David F. Lohman
The University of Iowa
Iowa City, Iowa
USA

July 8, 1993

Paper presented at the first Spearman Seminar, University of Plymouth, July 21, 1993.

Abstract

Spatial abilities have long been relegated to a secondary status in accounts of human intelligence. Tests of spatial abilities are viewed as measures of practical and mechanical abilities that are useful in predicting success in technical occupations, but not as measures of abstract reasoning abilities (Smith, 1964). This conflicts with the important role afforded to spatial imagery in accounts of creative thinking (Shepard, 1978), and with the observed correlations between spatial tests and other measures of intelligence. In fact, Spearman (see Spearman & Wynn Jones, 1950) considered spatial tests merely as unreliable measures of G. Hierarchical factor analyses generally support Spearman's conclusion, especially for complex spatial tests. Such tests are primarily measures of G, secondarily measures of something task-specific, and thirdly, measures of something that covaries uniquely with performance on other spatial tasks (Lohman, 1988). Simpler, speeded spatial tasks show lower G loadings, higher task specific loadings, and higher spatial factor loadings. In this paper, I first summarize and then attempt to explain these findings. The relationship between spatial task performance and G may reflect both statistical artifacts and psychological factors. Psychological factors include the attentional demands of maintaining and transforming images in working memory (Kyllonen & Christal, 1990) and the importance of mental models in reasoning (Johnson-Laird, 1983). Indeed, one can turn Spearman's conclusion around and with equal conviction conclude that measures of G are by and large unreliable measures of the ability to generate and coordinate different types of mental models in working memory. Evidence that supports and challenges such a conclusion is reviewed.

Introduction

There is a paradox in the literature on human spatial abilities. Indeed, many of those who have studied spatial abilities have noted it with reactions that range from amusement to annoyance (Paivio, 1971; Smith, 1964). It is this: On the one hand, tests of spatial abilities -- especially performance tests that use blocks or form boards or pieces of paper that must be folded and unfolded -- such tests are among the best measures of G (or Gf). Furthermore, spatial abilities are routinely implicated in accounts of creative and higher-order thinking in science and mathematics (Shepard, 1978; West, 1991). On the other hand, spatial abilities are often equated with concrete, lower-level thinking. Thus, they are used to predict success in various practical and technical occupations, such as carpentry, auto mechanics, and the like.

The source of the paradox is this. Hierarchical models of human abilities and the cannon of parsimony give G logical and statistical priority over measures of spatial ability. Therefore, we first account for the effects of G, and then examine correlations between residual scores and other variables. To be sure, there is something left. But the majority of the systematic variance is gone. It has already "been accounted for" by G. Or has it? Is G a psychological entity? Or is it primarily a statistical dimension? I would like to use modern theories of cognition to reexamine the ancient debate between Spearman and Thorndike, a debate later joined by Thomson and Thurstone on Thorndike's side, and by Burt and Vernon on Spearman's side (see Thorndike & Lohman, 1990). I believe that we know some things now that Spearman, Thomson, and Thorndike did not know that may allow us to see this controversy in a new light.

Importance of Spatial Abilities

Spatial ability may be defined as the ability to generate, retain, retrieve, and transform well-structured visual images. It is not a unitary construct. There are, in fact, several spatial abilities, each emphasizing different aspects of the process of image generation, storage, retrieval, and transformation. Spatial abilities are pivotal constructs of all models of human abilities. For example, Guilford (1967) devoted one slice of his Structure of the Intellect model to them. Hierarchical models (e.g., Vernon, 1950) place broad verbal-educational and spatial-visualization factors immediately below general ability since, after general ability, the verbal-spatial dimension captures more variance than any other dimension in large, representative batteries of ability tests [see, e.g., Eysenck's (1939) reanalysis of Thurstone's (1938) data]. Similarly, research on hemispheric specialization suggests that the difference between verbal-sequential processing and spatial-analog processing is a fundamental dichotomy in human cognition. Paivio (1971) has long argued for a dual code theory of memory in which verbal and spatial information are stored in different codes. More recently Anderson (1983), a long-standing opponent of this view, proposed a multicode theory of memory, with separate codes for temporarily ordered strings, spatial images, and abstract propositions.

High levels of spatial ability have frequently been linked to creativity, not only in the arts, but in science and mathematics as well (Shepard, 1978; West, 1991). For example, on several occasions Albert Einstein reported that verbal processes seemed not to play a role in his creative thought. Rather, he claimed to achieve insights by means of thought experiments on visualized systems of waves and physical bodies in states of relative motion. Other physicists (such as James Clerk Maxwell, Michael Faraday, and Herman Von Helmholtz), inventors (such as Nikola Tesla and James Watt), and generalists (such as Benjamin Franklin, John Herschel, Francis Galton, and James Watson) also displayed high levels of spatial abilities and reported that they played an important role in their most creative accomplishments.

In psychology, Shepard (1978, 1990) has given particularly lucid accounts of the role of spatial imagery in his own thinking. Involuntary dream images were the source of many of his most creative and influential contributions, including the idea for his experiment with Metzler on mental rotation, the first method of nonmetric multidimensional scaling, and the computer algorithm underlying additive nonhierarchical cluster analysis. However, in spite of the prominent role of spatial abilities both in models of human abilities and in models of cognition, tests of spatial abilities are not widely used, except as tests of "performance" or "nonverbal" intelligence, a role they have fulfilled since the introduction of the Army Beta exam during the first world war. Smith (1964) and Ghiselli (1973) summarize studies in which spatial tests have been used to predict job performance. Spatial tests add little to the prediction of success in traditional school subjects, even geometry, after general ability has been entered into the regression (Bennett, Seashore, & Wesman, 1974; McNemar, 1964). Predictive validities are somewhat higher for trade school courses (Bennett et al., 1974; Newman, 1945), and engineering school courses, particularly engineering drawing (e.g., Holliday, 1943). Tests of spatial and mechanical abilities are the best predictors of successful completion of training for machine workers and bench workers (Ghiselli, 1973) and for success in training courses for air crew positions (Guilford & Lacey, 1947). Vernon (1950) claims that validities are generally higher for younger and female populations than for older and male populations.

Combining Vernon's (1950) suggestion that the predictive value of spatial tests depends on other characteristics of subjects with the general finding that spatial abilities are more predictive in some courses of instruction than in others leads to the sort of hypothesis common in the Aptitude x Treatment Interaction (ATI) literature. The hypothesis that "verbalizers" and "visualizers" would profit from different instructional methods stretches from Galton (1880) to the present. It has been one of the most popular, yet one of the most elusive ATI hypotheses. Interactions between verbal and spatial abilities and instructional methods designed to require differential amounts of verbal and spatial processing are few, usually small, and inconsistent (Cronbach & Snow, 1977; Gustafsson, 1976), for both statistical (Gustafsson, 1989) and psychological reasons (Cronbach & Snow, 1977).

There are several possible reasons for the gulf between the theoretical importance of spatial abilities and their practical utility in predictive studies or in ATI studies. First, it may be that, beyond some minimum level of competence, spatial abilities are simply not that important for success in school or work. Second, the strength of spatial ability relative to other abilities, particularly verbal and phonemic fluency abilities, may be more important for predicting how problems are represented and solved rather than whether they can be solved. Third, the criterion measures used in most studies may be biased in favor of other abilities, such as verbal or reasoning skills. Fourth, existing tests may not be very good measures of spatial abilities. Fifth, the practice of first entering G into the regression may distort more than it reveals. In other words, the epistemological decision to give parsimony priority over psychological meaningfulness may leave us trying to explain a dimension that is statistically optimal, but psychologically obscure. This last argument will be a major focus of this paper. Before embarking on that discussion, however, it may be helpful briefly to review what is included under the heading of "spatial abilities," beginning with how these abilities are measured.

Measuring Spatial Abilities

Spatial abilities have been measured using four different types of tests: performance tests, paper-and-pencil tests, verbal tests, and film or dynamic computer-based tests. Performance tests were the earliest. Form board, block manipulation, and paper-folding tasks were among the items Binet and Simon (1916) used to measure the intelligence of children. Others created entire tests of a particular item type, such as form boards (Itard, 1801, cited in Spearman & Wynn-Jones, 1950; Paterson et al., 1930) or blocks (Kohs, 1923). Many of these tasks are used in contemporary intelligence tests as measures of performance or nonverbal intelligence (e.g., Wechsler, 1955). Another type of performance test seeks to estimate the ability to function in large-scale space. However, measures of the ability to orient oneself, to find efficient routes between locations, etc., show at best moderate correlations with other measures of spatial abilities (Allen, 1982; Lorenz & Neisser, 1986), perhaps in part because such tasks may be solved in ways that do not demand analog processing.

Paper-and-pencil tests of spatial abilities have an even more extensive history. Many such tests have been devised over the years. Eliot and Smith (1983) give directions and example items for 392 spatial tests, most of which were used in factorial investigations of abilities. Early factor analyses sought to demonstrate the existence of one or more spatial factors (Kelley, 1928, El Koussy, 1935). Some researchers, particularly those in Britain, were satisfied when they showed that a single spatial factor could be identified in the correlations among spatial tests once the general factor had been removed. These researchers tended to construct tests of spatial abilities that contained several different types of items and to study young, age-heterogeneous samples. American researchers used different methods of factor analysis, more homogeneous tests, and older, more homogeneous subject samples, and therefore, identified many different spatial factors. French (1951) made an early attempt to catalog these factors. Others (Guilford, 1967; Eysenck, 1967) proposed rational models that classified existing factors and suggested how others might be identified. Recent efforts to understand the dimensions of spatial abilities have moved away from these rational schemes and have sought instead to reanalyze old data sets using modern factor analytic methods and a hierarchical factor model in which factors are organized according to breadth from G, to broad group factors, to narrow group factors, to specifics (Carroll, 1993; Lohman, 1979). Table 1 lists the five major spatial factors identified in these reviews, a brief definition of each, and the name of a test that commonly loads in the factor. However, the specific variance in such tests is large, and so attempts to measure these abilities should always employ multiple tests for each factor (see Carroll, 1993, and

Lohman et al., 1987 for additional recommendations on tests, and Eliot & Smith, 1983, for examples of test items).

Insert Table 1 here

Verbal tests of spatial abilities have received much less attention, despite the fact that they often show high correlations with other spatial tests and various criterion measures (Guilford & Lacey, 1947; Ackerman & Kanfer, 1993). In this type of test, examinees must listen to a problem, presumably one that requires construction of a mental image, and then answer one or more questions. For example, "Imagine that you walk north for a while, then take a right turn, then walk further and take another right turn. In what direction are you facing?" Such tests require subjects to use spatial abilities in a way that is probably more representative of the manner in which such abilities are used in everyday life than do the items on most paper-and-pencil tests. Many cognitive tasks often require -- or at least benefit from -- the ability to construct a mental image that can be coordinated with linguistic inputs (Baddeley, 1986; Johnson-Laird, 1983; Kintsch, 1986).

Although spatial tests may require subjects to transform objects (such as by rotating or transposing them mentally), they typically present static objects. Some have hypothesized that the perception of dynamic spatial relationships involves somewhat different abilities. Gibson (1947) and later Seibert and Snow (1965) developed a variety of motion picture tests that were designed to measure these dynamic spatial abilities. However, individual differences on most dynamic spatial tasks appear to be well accounted for by performance on factors defined by paper-and-pencil tests. One exception is an ability factor called Serial Integration that appeared in several studies. Tests defining this factor required subjects to identify a common object from a series of incomplete pictures presented successively. More recently, Pellegrino and Hunt (1989) have devised several computer-administered tests of dynamic spatial abilities. Results to date show that individual differences in the ability to predict object trajectories and arrival times can be reliably measured and that individual differences on these tasks load on different factors than performance on paper-and-pencil tests.

Response Mode and Speededness

In addition to differences in presentation format (e.g., performance, paper-and-pencil, verbal, dynamic), spatial tests also differ in the type of response required (such as the selection of an alternative, construction of a response, or a verbal statement). For paper-and-pencil tests, there is some evidence that constructed-response tests are somewhat better measures of spatial ability (Lohman, 1988), and for this reason have long been preferred by British psychometricians (Eliot & Smith, 1983). Smith (1964) argues that spatial ability is best measured when subjects are required to maintain an image in its correct proportions. This is often done by presenting well-structured but fairly simple geometric designs which subjects must remember and then reproduce. As with all constructed-response tests, however, scoring is more difficult than when forced-choice tests are used. On the other hand, much additional information about subjects' abilities and test strategies can be had from a careful analysis of the type errors they make in drawing or constructing their answers (e.g., Kyllonen, 1984). Another important aspect of spatial tests is the relative emphasis placed on speed versus level. Tests administered under highly speeded conditions tend to measure more specific aspects of spatial ability than do tests administered under relatively unspeeded conditions. Altering the complexity of test items generally results in a change in the factorial loading of a test. Simple items must be administered under speed conditions in order to generate individual differences, and so changes in task complexity usually mean a change in test speededness as well. Computer-based tests offer the opportunity to gather both error and latency scores, which can then be combined to predict criterion performances with greater precision than from either measure considered separately (Ackerman & Lohman, 1990). However, performance on such tests is more influenced by the speed or accuracy emphasis subjects adopt than is performance on time-limited tests.

Practice and Training Effects

Spatial tests often show substantial practice effects. Retest gains range from .2 to 1.2 SD, effects being somewhat larger for simpler tests, shorter retest intervals, and subjects who are given feedback (Krumboltz & Christal, 1960; Lohman, 1988; 1993). Effects of this magnitude can seriously compromise interpretations of test scores if examinees are differentially familiar with test problems. However, transfer to nonpracticed tests that load on the same spatial factor is typically much smaller, often nonexistent. Several studies now suggest that the key variable in predicting transfer is similarity of procedures employed rather than of stimuli used, at least when the subjects are young adults and the stimuli are regular polygons (Lohman, 1993).

Spatial abilities can be improved with practice and training, even though particular courses of instruction (such as engineering drawing) have inconsistent effects. This may in part reflect the fact that treatments designed to improve performance on spatial tasks often are disruptive for high verbal ability subjects. One possibility is that these treatments impose or induce external regulation of performance. External regulation may compensate for the inadequate self-regulation activities of low verbal subjects, but interfere with the self-regulation activities of high-verbal subjects (Lohman, 1986a).

Although short-term studies often produce small or conflicting findings, Balke-Aurell's (1982) study of the effects of tracking in the Swedish secondary school system suggest that the cumulative effects of differential educational and work experiences can be quite large. Students educated in schools using a verbally-oriented curricula showed greater growth in verbal abilities than spatial abilities, whereas those educated in schools using a technical curricula showed greater growth in spatial abilities.

Personality Correlates

The relative ease with which we can create imagistic versus semantic elaborations also correlates with personality constructs (Smith, 1964). One of the clearest demonstrations of this comes from the work of Riding (1983). Riding (1983) was interested in children's habitual modes of thinking. He developed a task in which he read a short story to a child and then asked a series of questions about the passage, all of which required some inference. Questions were of two types, those that depended on imagery and those that depended on semantic elaboration. For example, the story may have mentioned the fact that someone knocked on the door of a cottage. The question might be "What color was the door?" Of course, where was no right answer, since color of the door was not specified. Response latency was recorded. However, the dependent variable of interest was an ipsative score that compared latencies on semantic and imagery questions. The idea was to identify children who were much quicker to answer one type of question than the other. Correlations were then computed between this ipsative score and a personality scale. Children who showed a preference for imagistic processing were much more likely to be introverted, whereas those who showed a preference for verbal elaboration were more likely to be extroverted.

Contributions of Cognitive Research

Spatial Cognition

Cognitive psychology has contributed importantly to our understanding of how subjects encode, remember, and transform visual images, and thus to what spatial abilities might be. Seminal research here was that of Roger Shepard and his students. Shepard (1978) poses an interesting challenge tocognitive scientists:

> Suppose that we do not start by asking what kinds of thought processes are most accessible to empirical study, are most conveniently externalized in the form of discrete symbols,words, or sentences, or are most readily described by existing models imported into cognitive psychology from linguistics or computer science. Suppose, instead, that we first ask what sorts of thought processes underlie human creative acts of the highest and most original order. Perhaps we shall come to be less than fully satisfied with research that is exclusively motivated by current theories of linear sequential processing of discrete symbolic or propositional structures. (p. 134)

The challenge to propositional theories of cognition was made most forcefully in an early series of experiments on mental rotation (see Shepard & Cooper, 1982, for a summary). The basic finding was that the time required to determine whether two figures could be rotated into congruence was a linear function of the amount or rotation required. On the basis of this and other evidence, Shepard claimed that mental rotation was an analog process that showed a one-to-one correspondence with physical rotation. The second claim was that this rotation process was performed on a mental representation that somehow preserved information about structure at all points during the rotation transformation.

Others have been more explicit about the nature of this representation. Most agree that spatial knowledge can be represented in more than one way. One representation (sometimes called an image code) is thought to be more literal (Kosslyn, 1980) or at least more structure- or configuration- preserving (Anderson, 1983). Another representation is more abstract and is more meaning- or interpretation-preserving (Kosslyn, 1980; Anderson, 1983; Palmer, 1977) and is usually modeled by the same propositional structures used to represent meaningful verbal knowledge. Much of the confusion in understanding spatial abilities can be traced to whether spatial abilities are restricted to image-coded memories and the analog processes that operate on them or whether proposition-coded memories and the general procedural knowledge that operate on them are also considered part of the term. In other words, much of the confusion lies in whether abilities are defined by performance on a certain class of tasks or by skill in executing certain types of mental processes.

Individual Differences in Spatial Cognition

Although research and theory in cognitive psychology and artificial intelligence suggest much about the nature of spatial knowledge and processes, it does not explicitly address the source of individual differences in spatial processing. Research on this question has followed four hypotheses: that spatial abilities may be explained by individual differences in (a) speed of performing analog transformations, (b) skill in generating and retaining mental representations that preserve configural information, (c) the amount of visual-spatial information that can be maintained in an active state, or (d) the sophistication and flexibility of strategies available for solving such tasks.

Transformation hypothesis. The most popular hypothesis has been that spatial abilities may be explained by individual differences in the speed with which subjects can accurately perform analog mental transformations, particularly rotation. However, correlations between rate of rotation (estimated by the slope of the regression of latency on angular-separation) and spatial ability vary from highly negative (e.g., Lansman, 1981) to moderately positive (e.g., Poltock & Brown, 1984). Correlations are generally higher for three-dimensional rotation problems than for two-dimensional problems (Pellegrino & Kail, 1982; Cooper & Regan, 1982), and for practiced than for nonpracticed subjects (Lohman & Nichols, 1990). However, even moderate correlations between the slope measure and other variables are difficult to interpret. The slope is heavily influenced by the amount of time taken on trials requiring the most rotation. It may thus better reflect the number of attempts made to solve the problem or simply the time taken to solve these most difficult problems, not rate or rotation. More importantly, it can be shown that slopes and other component scores are incapable of explaining individual differences on tasks that are consistent across trials. Instead, these individual differences are captured in individual mean or intercept scores (Lohman, in press). On the other hand, correlations between overall error rates and spatial reference tests are often quite high. Indeed, although rate of information processing on rotation tasks and accuracy levels achieved under liberal time allotments are necessarily confounded, differences between high and low spatial subjects are much greater on the accuracy score than on a rate of information processing score (Lohman, 1986b). One interpretation of this finding is that the amount of information that can be maintained in an active state of working memory is more important than rate of processing that information in accounting for individual differences in spatial ability.

Superior spatial working memory. Baddeley's (1986) model of working memory hypothesizes a central executive and two slave systems: an articulatory loop and a visual-spatial scratch pad. Perhaps high spatial subjects can maintain more image-coded information in this scratch pad. Kyllonen's (1984) study of ability differences in types and number of errors made on a paper folding task supports this hypothesis. His study showed that high and low spatial subjects differed not so much in the type of error committed but in the number or errors committed. Because of this, Kyllonen (1984) concluded that the main difference between the performance of high- and low-spatial participants was that lows were more likely to forget a fold and then either not perform it or substitute an incorrect fold for the forgotten one. Other theorists (e.g., Just & Carpenter, 1992) emphasize the tradeoff between storage and transformation functions in a unitary working memory system. By this account, mental rotation problems are good measures of spatial ability because they place substantial demands on both storage and transformation functions, and require subjects to manage the tradeoff between them. Other evidence in support of this view comes from verbal problems that require imagery for their solution, such as Binet's "It is 12:15. If we switch the hands on the clock, what time will it be?" Tests constructed of such problems often show high predictive validities, but are not factorially pure (Ackerman & Kanfer, 1993).

Nature of the representation. Several investigators have sought to determine whether high and low spatial subjects differ in the type of mental representations they create (see, e.g., Cooper, 1982; Lohman, 1988). Individual differences in memory for random forms show no relationship with performance on other spatial tests (Christal, 1958). Thus, it is not so much the ability to remember stimuli but the ability to remember systematically structured stimuli that distinguishes between subjects high and low in spatial ability. Low-spatial subjects seem to have particular difficulty in constructing systematically structured images. High spatial subjects appear to be able to construct images that can be compared holistically with test stimuli. Differences between high and medium spatial subjects are often small in this respect. It is the very low spatial subjects who appear qualitatively different (Pellegrino & Kail, 1982; Lohman, 1988).

Strategies. It has long been noted that spatial tasks may be solved in more than one way, with some strategies placing greater demands on analog processing than others. Several studies have now shown that the strategies subjects employ on form-board type tasks are systematically related to their ability profiles (Kyllonen, Lohman, & Woltz, 1984; Lohman, 1988). The major distinction is between spatial and non-spatial strategies. Subjects using the spatial strategy remember complex polygons by decomposing them into simpler geometric shapes. When required to assemble figures mentally, their performance is more influenced by the characteristics of the to-be-assembled figure than by that of the component figures. Time to perform this assembly operation is usually negatively correlated with reference spatial tests. On the other hand, subjects using the non-spatial strategy try to remember complex polygons by associating the figure with another concrete, easily labeled object. When asked to assemble figures mentally their performance is strongly influenced by the complexity of the component figures rather than of the to-be-assembled figure. Further time to perform the assembly often shows higher correlations with tests of verbal ability than with tests of spatial ability.

Rotation tasks are also solved in different ways by different subjects. Bethell-Fox and Shepard (1988) found that rotation times for unfamiliar stimuli were generally influenced by the complexity of the stimulus. With practice, most subjects learned to rotate all stimuli at the same rate. However, some subjects continued to show effects for stimulus complexity even after much practice. Bethell-Fox and Shepard (1988) argued that these subjects rotated stimuli piece by piece whereas after practice others rotated them holistically. Carpenter and Just (1978) argued that even practiced subjects do not rotate an image of an entire three-dimensional object, but rather only a skeletal representation of it. In experiments on a cube rotation task, they found that subjects used different strategies, presumably related to the coordinate system the subject adopted. Low spatial subjects appeared to rotate the cube iteratively along standard axes whereas high spatial subjects were able to use the shorter trajectory defined by a single transformation through both axes (Just & Carpenter, 1985).

Thus, subjects of different ability levels and profiles often solve spatial tests in predictably different ways. However, flexibility of strategy in solving such tasks seems to be more related to G or Gf than to spatial ability (Kyllonen et al., 1984). Indeed, subjects high in spatial but low in verbal abilities have been found to apply the same "spatial" strategy to all problems. Perhaps they have no need to switch to other strategies.

Spatial Ability and G

Factorial studies of spatial ability routinely show that spatial tests are good measures of G. Investigations of individual differences in spatial processing suggested that the major reason for overlap is that spatial tests place extraordinary demands on working memory. If Kyllonen and Christal (1990) are correct, then this would explain the high correlation between them.

It is useful to extend this reasoning further. Although individuals are sometimes required to solve the sorts of problems presented on tests (particularly paper and pencil tests) of spatial abilities, everyday thinking for the most part requires a different use of spatial abilities. One example comes from research on reading comprehension. For years, reading comprehension was modeled as the process of creating an internal model of the text that mimicked its logical structure. In other words, to comprehend something meant to construct an internal outline or summary of it. Then Kintsch and Greeno (1985) joined forces to understand how children solve -- or better, why they fail to solve -- word problems in mathematics. What they discovered was that a text model was not enough. Children needed also to construct a visual mental model that could be coordinated with the text model. Further, the visual model (or analog) became increasingly important as problem complexity increased. It provided a way to integrate and coordinate concepts and the relationships among them. A good example of what this means for comprehension comes from the experience of assembling a toy from printed directions that only look like English. I can read the words "Put hex nut K and lock washer Q on spindle d-1, and tighten loosely." I may be able to repeat them, to paraphrase them, even to summarize them. But if I cannot visualize what I must do, then I do not understand. Similarly, children need pictures to help them understand stories. Indeed, the first text without pictures they read is an important milestone. Thus, by this model, understanding means using linguistic clues to construct a text model and imagery clues (or analogy or metaphor) to construct an image model, and then coordinating the two. These two aspects of working memory are nicely depicted in Baddeley's (1986) theory. Baddeley claims that working memory contains a central executive (whose functioning remains somewhat mysterious. In fact, he calls it "the area of residual ignorance") and two slave systems: a phonological loop and a visual-spatial scratch pad. As I see it, the most important function of imagery in the visual-spatial scratch pad is simply to help us keep track of what we are doing, to see relationships among concepts we have represented either literally or metaphorically by our images. Thought without imagery would be like prose without metaphor. Indeed, one indication of the importance of these models in our thought is the pervasiveness of metaphor in our speech.

Thus, one can turn Spearman's conclusion around (to use a spatial metaphor) and, with equal conviction, conclude that measures of G are largely unreliable measures of the ability to generate and transform different types of mental models in working memory. Alternatively, G may represent the crucial but poorly understood process of coordinating different types of mental models represented in linguistic and imagistic slave systems. It would, to use Baddeley's (1986) words, represent the "area of residual ignorance," or what he also calls the *Supervisory Attentional System.*

Although the mental models view of Johnson-Laird (1983) and Kintsch and Greeno (1985) is useful, another perspective is provided by Anderson's (1983) ACT* theory. Here the crucial distinction is between different types of memory codes. The first set of memory codes contains several perception-based codes (such as the image code and the linear-order code) that preserve important structural aspects of a percept, such as the configuration of its elements (image) or their temporal sequence (linear-order). Interestingly, these two codes correspond rather nicely to Baddeley's two slave systems. The second type of code -- called the abstract proposition -- is not tied to any particular perceptual process. Rather, it is said to preserve meaning, which may be extracted from perception-based codes or from other proposition-coded memories. I believe that evidence for the separability of meaning-based and perception-based codes is compelling. Anderson (1983) reviews some of it. Other evidence comes from differential psychology, particularly studies of savants. Indeed, the most striking characteristic of most savants is their remarkable ability to retain and manipulate information in a perception-based code without being able to extract more abstract meanings from these perceptual experiences.

On this view, then, particular abilities represent the effectiveness (relative to other individuals) with which an individual can create, transform, and retain different types of mental codes. The dominance of G would reflect the importance of the abstract proposition code in human cognition, that is, with individual differences in the ability to discover and retain meaningful relationships. Particular verbal and spatial abilities would be relegated to a secondary status, although one would expect them to be more important early in development, as some have hypothesized. For example, Bruner (1973) hypothesized such a role for imagery abilities. Others, though, have noted a similar phenomenon in children's acquisition of language. Vygotsky (in Wertsch, 1985), borrowing from the great French psychiatrist Janet put it this way:

> One of the mechanisms that makes possible the cognitive development and general acculturation of the child is the process of coming to recognize the significance of the external sign forms (e.g., words) that he or she has already been using in social interaction. [In other words] children can say more than they

> realize and it is through coming to understand what is meant by what is said that their cognitive skills develop (p.).

The suggestion that G is closely tied to the ability to create, transform, and retain meaning-coded memories may in fact overlap with the mental models view. Perhaps what was earlier called the Supervisory Attentional System, that is, the ability to choose among different mental models and especially, to coordinate linguistic-based [or what Kintsch (1986) calls the text-based model] and image-based models, perhaps this is the function and meaning of general intellectual ability.

I hasten to add, however, that the psychological basis of a general cognitive function has no necessary relationship with a general dimension of individual differences. The former refers to generalization across cognitive functions or tasks, whereas the latter refers to generalization across individuals. Just as printers must learn not to confuse their p's and q's, so too, must we learn not to confuse generalization across the rows and columns of our data matrices. In other words, even though the ability to create, transform, and retain meaning-based memories may be the most important source of cross-task consistency for a particular individual, individual differences may derive only in part from this process. Indeed, the literature I reviewed earlier suggests that general working memory or attentional resources may account for even more of the individual difference variance.

Conclusions

Spatial ability may be defined as the ability to generate, retain, retrieve, and transform well-structured visual images. Distinguishably different aspects of spatial ability can thus be measured by constructing tests that emphasize different aspects of this process. Once relegated to lower-order processing and concrete thought, spatial abilities are now understood as important for higher-order thinking in science and mathematics, for the ability to generate and appreciate metaphor in language, and for creativity in many domains. The ability to generate visual-spatial models that can be coordinated with linguistic inputs has a pervasive impact on all of cognition. Individuals who excel in the ability to create and manipulate such models are not only more likely to succeed in occupations that require spatial abilities, but are also more likely to generate such models spontaneously when thinking and, especially when verbal fluency abilities are relatively low, to be more introverted. Spatial transformations place heavy demands on working memory and so spatial tests often show high correlations with tests of general fluid ability.

Table 1

Factor	Definition[a]	Example Test[b]
Visualization	Ability in manipulating visual patterns, as indicated by level of difficulty and complexity in visual stimulus material that can be handled successfully, without regard to the speed of task solution	Paper folding
Speeded Rotation	Speed in manipulating relatively simple visual patterns, by whatever means (mental rotation, transformation, or otherwise).	Cards
Closure Speed	Speed in apprehending and identifying a visual pattern, without knowing in advance what the pattern is, when the pattern is disguised or obscured in some way.	Street Gestalt
Closure Flexibility	Speed in finding, apprehending, and identifying a visual pattern, knowing in advance what is to be apprehended, when the pattern is disguised or obscured in some way.	Concealed Figures
Perceptual Speed	Speed in finding a known visual pattern, or in accurately comparing one or more patterns, in a visual field such that the patterns are not disguised or obscured.	Identical Pictures

[a] From Carroll, 1993, pp. 362-363.

[b] See Eliot and Smith (1983).

References

Ackerman, P. L., & Kanfer, R. (1993). Integrating laboratory and field study for improving selection: Development of a battery for predicting air traffic controller success. *Journal of Applied Psychology, 78,* 413-432.

Ackerman, P. L., & Lohman, D. F. (1990). *An investigation of the effect of practice on the validity of spatial tests.* Final Report (NPRDC Contract N66001-88C-0291). Minneapolis, MN: Personnel Decisions Research Institute.

Allen, G. L. (1982). *Assessment of visuospatial abilities using complex cognitive tasks.* Norfolk, VA: Department of Psychology, Old Dominion University.

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

Baddeley, A. (1986). *Working memory.* Oxford, UK: Clarendon Press.

Balke-Aurell, G. (1982). *Changes in ability as related to educational and occupational experience.* Goteborg, Sweden: Acta Universitatis Gothoburgensis.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1974). *Manual for the Differential Aptitude Test* (5th ed.). New York: The Psychological Corporation.

Bethell-Fox, C. E., & Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity, familiarity, and individual differences. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 12-23.

Binet, A., & Simon, T. (1916). *The development of intelligence in children* (E. S. Kite, Trans.). Baltimore: Williams & Wilkens.

Brinkman, E. H. (1966). Programmed instruction as a technique for improving spatial visualization. *Journal of Applied Psychology, 50,* 179-184.

Bruner, J. S. (1973). *Beyond the information given.* New York: W. W. Norton & Company.

Carpenter, P. A., & Just, M. A. (1978). Eye fixations during mental rotation. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.). *Eye movements and the higher psychological functions.* Hillsdale, NJ: Erlbaum.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Chrystal, R. E. (1958). Factor analytic study of visual memory. *Psychological Monographs, 72,* (13, Whole No. 466).

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, pp. 77-124). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, L. A., & Regan, D. T. (1982). Attention, perception, and intelligence. In R. J. Sternberg (Eds.) *Handbook of human intelligence* (pp. 123-169). Cambridge: Cambridge University Press.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions.* New York: Irvington.

Eliot, J. C., & Smith, I. M. (1983). *An international directory of spatial tests.* Windsor, England: NFER-Nelson.

Eysenck, H. J. (1939). [Review of the book *Primary mental abilities* by L. L. Thurstone]. *British Journal of Psychology, 9,* 270-275.

Eysenck, H. J. (1967). Intellectual assessment: A theoretical and experimental approach. *British Journal of Educational Psychology, 37,* 81-98.

French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monographs,* No. 5.

Galton, F. (1880). Statistics of mental imagery. *Mind, 5,* 300-318.

Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology, 26,* 461-477.

Gibson, J. J. (Ed.) (1947). Motion picture testing and research. *Army air forces aviation psychology program,* (Report No. 7). Washington, DC: GPO.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Guilford, J. P., & Lacey, J. I. (Eds) (1947). Printed classification tests. *AAF aviation psychology research program reports* (No. 5). Washington, DC: GPO.

Gustafsson, J. E. (1976). Verbal and figural aptitudes in relation to instructional methods: Studies in aptitude-treatment interaction. *Gotenborg Studies in Education Sciences,* No. 17, Gotenborg, Sweden.

Gustafsson, J. E. (1989). Broad and narrow abilities in research on learning and instruction. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota Symposium on learning and individual differences* (pp. 203-237). Hillsdale, NJ: Erlbaum.

Holliday, F. (1943). The relations between psychological test scores and subsequent proficiency of apprentices in the engineering industry. *Occupational Psychology, 17,* 168-85.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review, 92*, 137-172.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

Kelly, T. L. (1928). *Crossroads in the mind of man.* Stanford: Stanford University Press.

Kintsch, W. (1986). Learning from text. *Cognition and Instruction, 3,* 87-108.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*, 109-129.

Kohs, S. C. (1923). *Intelligence measurement: A psychological and statistical study based upon the block designs test.* New York: MacMillan.

Kosslyn, S. M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Koussy, A. A. H. El (1935). The visual perception of space. *British Journal of Psychology, 20.* (Monograph supplement)

Krumboltz, J. D., & Christal, R. E. (1960). Short-term practice effects in tests of spatial aptitude. *Personnel and Guidance Journal,* __, 384-391.

Kyllonen, P. C. (1984). Information processing analysis of spatial ability. (Doctoral dissertation, Stanford University) *Dissertation Abstracts International, 45,* 819A.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence, 14,* 389-433.

Kyllonen, P. C., Lohman, D. F., & Woltz, D. J. (1984). Componential modeling of alternative strategies for performing spatial tasks. *Journal of Educational Psychology, 76,* 1325-1345.

Lansman, M. (1981). Ability factors and the speed of information processing. In M. P. (Friedman, J. P. Das and N. O'Connor (Eds.), *Intelligence and learning* (pp. 441-457). New York: Plenum Press.

Lohman, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature* (Tech. Rep. No. 8), Stanford, CA: Stanford University, Aptitude Research project, School of Education. (NTIS NO. AD-A075 972).

Lohman, D. F. (1986a). Predicting mathemathanic effects in the teaching of higher-order thinking skills. *Educational Psychologist, 21,* 191-208.

Lohman, D. F. (1986b). The effect of speed-accuracy tradeoff on sex differences in mental rotation. *Perception and Psychophysics, 39,* 427-436.

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 40, pp. 181-248). Hillsdale, NJ: Erlbaum.

Lohman, D. F. (1993). *Effects of practice and training on the acquisition and transfer of spatial skills:* Two speed-accuracy studies. (Final Report Grant AFOSR-91-9367). Iowa City, IA: Lindquist Center for Measurement.

Lohman, D. F. (in press). Component scores as residual variation (or Why the intercept correlates best). *Intelligence.*

Lohman, D. F., & Nichols, P. D. (1990). Training spatial abilities: Effects of practice on rotation and synthesis tasks. *Learning and Individual Differences, 2,* 69-95.

Lohman, D. F., Pellegrino, J. W., Alderton, D. L., & Regian, J. W. (1987). Dimensions and components of individual differences in spatial abilities. In S. H. Irvine & S. N. Newstead (Eds.), *Intelligence and cognition: Contemporary frames of reference* (pp. 253-312).

Lorenz, C. A., & Neisser, U. (1986). *Ecological and psychometric dimensions of spatial ability.* Atlanta, GA: Department of Psychology, Emory University, Report #10.

McNemar, Q. (1964). Lost: Our intelligence? Why? *American Psychologist, 19,* 871-882.

Newman, J. (1945). The prediction of shopwork performance in an adult rehabilitation program: The Kent-Shakow Industrial Formboard Series. *Psychological Record, 5,* 343-352.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart and Winston.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9,* 441-474.

Paterson, D. G., Elliot, R. M., Anderson, L. D., Toops, H. A., & Heidbreder, E. (1930). *Minnesota mechanical ability tests.* Minneapolis: University of Minnesota Press.

Pellegrino, J. W., & Hunt, E. B. (1989). Computer-controlled assessment of static on dynamic spatial reasoning. In R. F. Dillon & J. W. Pellegrino (Eds.), *Testing: Theoretical and applied perspectives* (pp. 174-198). New York: Praeger.

Pellegrino, J. W., & Kail, R. (1982). Process analyses of spatial aptitude. In R. J. Sternberg (Ed.). *Advances in the psychology of human intelligence.* (Vol. 1, pp. 311-366). Hillsdale, NJ: Lawrence Erlbaum Associates.

Poltrock, S. E., & Brown, P. (1984). Individual differences in visual imagery and spatial ability. *Intelligence, 8,* 93-138.

Riding, R. J. (1983). Extraversion, field independence, and performance on cognitive tasks in twelve-year-old children. *Research in Education, 29,* 1-9.

Seibert, W. F., & Snow, R. E. (1965). *Studies in cine-psychometry I: Preliminary factor analysis of visual cognition and memory.* Lafayette, IN: Audio Visual Center, Purdue University.

Shepard, R. N. (1978). Externalization of mental images and the act of creation. In B. S. Randhawa & W. E. Coffman (Eds.), *Visual learning, thinking, and communication* (pp. 133-190). New York: Academic Press.

Shepard, R. N. (1990). *Mind sights.* New York: W. H. Freeman & Company.

Shepard, R. N., & Cooper, L. A. (1982). *Mental images and their transformations.* Cambridge, MA: MIT Press.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701-703.

Smith, I. M. (1964). *Spatial ability*. San Diego: Knapp.

Spearman, C., & Wynn Jones, L. L. (1950). *Human ability.* London: MacMillan.

Thorndike, R. M., & Lohman, D. F. (1990). *A century of ability testing.* Chicago: The Riverside Publishing Co.

Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs, 1*.

Vernon, P. E. (1950). *The structure of human abilities.* London: Methuen.

Wechsler, D. (1955). *Wechsler adult intelligence scale*. New York: Psychological Corporation.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind.* Cambridge, MA: Harvard University Press.

West, T. G. (1991). I*n the mind's eye.* Buffalo, New York: Prometheus Books.