The Role of Spatial Thinking in Undergraduate Science Education

Mary Hegarty, University of California, Santa Barbara

Introduction: The Importance of Spatial Thinking in Science.

Visuospatial thinking is central to many scientific domains and professions. For example, physicists reason about the motions of objects through space, chemists justify the relationship between molecular structure and chemical reactivity, and geologists reason about the physical processes that form mountains and canyons. Scientists have also developed a variety of spatial representations such as diagrams, graphs, models, and maps to represent the objects of study and more abstract data in their respective domains.

Given the importance of spatial representations and thinking processes in science, it is perhaps not surprising that spatial ability is related to success and participation in science. In the strongest evidence for this relationship to date, longitudinal studies of intellectually talented students indicate that spatial ability accounts for a statistically significant proportion of the variance in participation in science, over and above SAT Mathematical and SAT Verbal scales (Shea, Lubinski & Benbow, 2001; Webb, Lubinski & Benbow, 2007). Moreover a recent analysis of Project Talent data, a longitudinal database including over 400,000 randomly sampled students, shows that people who received degrees in mathematics, engineering and physical sciences and those who went on to pursue scientific occupations had significantly higher spatial abilities at age 13 than those who received degrees in other fields or practiced other professions (Wai, Lubinski & Benbow, 2009). Spatial ability was more predictive of higher level achievements (e.g., receiving a Ph.D. in science as opposed to receiving a bachelors' degree in science).

In spite of the importance of spatial thinking in scientific disciplines, promoting spatial thinking is not emphasized in our educational system. A recent report (National Research

Council, 2006) claimed that spatial intelligence is “not just undersupported but underappreciated, undervalued, and therefore underinstructed” (p. 5) and called for a national commitment to the development of spatial thinking across all areas of the school curriculum. Lack of emphasis on spatial thinking not only prevents less able students from achievement in science. It also hinders us from identifying and nurturing the talents of our most spatially able students. Wai et al., (2009) point out that 70% of those who were in the top 1% of the Project Talent sample in spatial ability were not in the top 1% in either mathematics or verbal ability indicating that they were not identified as “gifted”, although they showed extraordinary potential for achievement in science.

What is Spatial Ability? There has been a long history of measurement of spatial abilities leading to the identification of several different spatial ability factors in factor analytic studies (see Hegarty & Waller 2005). In the most extensive study to date Carroll (1993) surveyed and reanalyzed more than 90 data sets that bear on the factor structure of spatial abilities. He concluded that the most commonly measured spatial ability factor is *spatial visualization* defined as “power in solving increasingly difficult problems involving spatial forms” (Carroll, 1993, p. 315) and measured by the types of tasks shown in Figure 1. This factor is sometimes dissociated from the *spatial relations* factor, which examines simpler transformations (mental rotation of 2 dimensional forms, see Figure 2a) and emphasizes speeded performance. Carroll also found support for the separability of three additional spatial factors. Two of these, *closure speed* (Figure 2b) and *flexibility of closure* (Figure 2c*)*, involve the ability to identify a stimulus or part of a stimulus that is either embedded in or obscured by visual noise. The final factor is called*perceptual speed* (Figure 2d), and described as “speed in comparing figures or symbols, scanning to find figure or symbols, or carrying out very simple tasks involving visual perception” (French, 1951).

In addition to these spatial ability factors, which have been measured and studied extensively, researchers have recently identified several aspects of spatial thinking that have not been well measured historically. One important distinction is between small-scale and large-scale spatial ability. The spatial ability factors identified in the psychometric literature primarily reflect small scale spatial ability, that is, ability to perceive and imagine transformations of manipulable objects. In contrast, individual differences in large scale tasks (also referred to as environmental spatial abilities) include facility in tasks such as learning the layout of a new environment, retracing a route, or pointing to unseen locations in a known environment. Small-scale spatial abilities and the ability to learn the layout of an environment are partially dissociated; that is, they share some variance, but someone who is good at object based tasks is not necessarily good at environmental-scale spatial tasks (Hegarty, Montello, Richardson, Ishikawa & Lovelace, 2006; Hegarty & Waller, 2005). Another important distinction is between static and dynamic spatial abilities. Dynamic spatial abilities are those that are required to reason about moving stimuli. This ability can also be dissociated from those measured by pencil and paper psychometric tests showing static stimuli (Hunt, Pellegrino, Frick, Farr & Alderton, 1988; Contreras, Colom, Hernandez & Santacreu, 2003).

Main Approaches to Studying Spatial Thinking in College Science. This paper summarizes four types of studies regarding spatial thinking in undergraduate science curricula (1) *Correlational studies* that examine the relations between measures of spatial ability and performance in science disciplines, (2) *Training studies* that attempt to train aspects of spatial thinking, (3) Focused Studies of how students understand specific *spatial representations in* science disciplines (4) studies that use *dynamic spatial representations* (models, animations and visualizations) to promote scientific understanding.

(1) Correlational Studies

Correlational studies measure the spatial abilities of students in science classes, or in the laboratory, and examine the correlations of these ability measures with various aspects of science achievement. In this section I review representative studies of this type in the domains of biology/medicine, chemistry, geology and physics/engineering.

Biology and Medicine: Studies in biology have examined correlations of spatial abilities with ability to learn the structure of anatomy and imagine cross sections of anatomy. For example, Rochford (1985) examined performance in practical exams and anatomy classes by 2nd year medical students, classifying anatomy test items as either spatial or non-spatial. While spatial ability was not predictive of non-spatial items (correlation range -.01 to .13 across different semesters) it was predictive of spatial items (range = .14 - .39) and performance in practical exams (.13 - .47). Similarly, spatial ability predicted learning of the structure of the carpal bones by university anatomy students learning from either key views of the anatomy or from multiple views (Garg, Norman, Spero & Maheshwari, 1999; Garg, Norman & Sperotable, 2001).

Ability to draw or identify the cross section of a 3-dimensional anatomy-like structure is also correlated with spatial ability. In a laboratory study with undergraduate students, Cohen & Hegarty, 2007 found that ability to draw the cross section of a novel three dimensional object was correlated with tests of mental rotation ($r = .39$, $p < .05$) and perspective taking ability (.59, $p < .01$). In more authentic studies with dentistry students spatial abilities predicted ability to choose the correct cross section of both novel objects (correlation range: .37-.52) and anatomical objects (teeth, correlation range .29-.37) (Hegarty, Keehner, Khooshabeh & Montello, 2009). Spatial abilities also predicted grades in restorative dentistry practical laboratory classes, but not in anatomy classes. Moreover, most of the correlations with spatial ability remained significant after controlling for common variance with general intelligence.

Chemistry: Chemistry is centrally concerned with how the spatial structure of molecules relates to their reactive properties, so it is not surprising that researchers have examined relationships between spatial ability measures and performance in chemistry classes. Bodner & McMillan (1986) found significant correlations (ranging from .29 to .35) between measures of spatial visualization and flexibility of closure (see Figures 1 and 2) and measures of performance in an introductory organic chemistry course. Later studies indicated small but significant correlations (in the .2-.3 range) between measures of spatial ability and performance in college courses in both general chemistry (Carter, LaRussa & Bodner, 1987) and organic chemistry (Pribyl & Bodner, 1987). Correlations were found for items that required students to mentally manipulate representations of molecules and solve problems, including problems that are not specifically spatial in nature (stochiometry problems). Spatial ability was not significantly correlated with items that measured rote knowledge or the application of simple algorithms.

The ability to relate and translate between representations has been identified as a specific difficulty for chemistry students. Keig and Rubba (1993) studied the effects of knowledge of representations, reasoning ability, and spatial ability on students’ ability to translate between formulae, electronic configuration diagrams, and ball-and-stick models in pre-college chemistry classes. They found that knowledge and reasoning ability predicted representation translation performance, but spatial ability did not add to prediction after controlling for these sources of variance. Our research group has found significant effects of spatial ability (range .32 to .38 in different studies) on ability to translate between different diagrammatic representations in organic chemistry (Stull, Hegarty, Dixon & Stieff, 2010). These representations differ in the orientation from which the molecule is viewed, explicitly display different properties of the molecule (while leaving others implicit), and use different diagrammatic conventions, e.g., for showing the 3rd dimension, so translating between them may be particularly dependent on internal spatial transformations.

Geosciences: Spatial thinking problems in geosciences include understanding spatial structures, inferring the internal structure of structures from visible outcrops or slices of block models and understanding spatial processes such as air mass circulation and plate tectonics. Orion, Ben-Chaim and Kali (1997) found sizable correlations (range .35 - .51) between measures of spatial visualization ability and performance in an introductory geology class. Liben, Kastens and Christensen (in press) also found relations between a basic measure of spatial ability, the water level task, and concepts in geology. In the water level task (classified as a test of flexibility of closure), people are shown pictures of bottles at different orientations and have to draw a line indicating orientation of level of a liquid in the bottle. Perhaps surprisingly, many adults have difficulty with this task. Liben et al classified participants as good, medium or poor on the water level task and found that these groups differed significantly in mapping an outcrop’s strike and dip, which are central concepts in geology.

Physics: Finally, researchers have documented correlations between spatial abilities and physics (mechanics) problem solving. Kozhevnikov and Thornton (2006) found correlations (range .28 to .32) between a measure of spatial visualization ability and mechanics problem solving that include relating force and motion events and interpreting graphs of force and acceleration. In a study of psychology undergraduate students, Kozhevnikov, Hegarty & Mayer (2002) found correlations between spatial visualization and performance on qualitative problems involving extrapolating motion and changing frame of reference. Measures of speeded mental rotation were not correlated with performance on this test. Finally in laboratory studies with psychology undergraduate students, Hegarty and Sims (1995) found correlations between spatial ability and ability to infer the motion of different machine components when the machine was working (mental animation), that is "run a mental model" of the machine, and Isaac and Just (1995) found that individuals' susceptibility to illusions regarding rolling motion were related to poor spatial visualization abilities

Issues in the Interpretation of Correlational Studies. While correlational studies provide good evidence for the importance of spatial skills in various science disciplines, it is important to interpret them critically. First, many of the studies conducted to date are based on small samples and specific tasks in the relevant science domains. Second most of the observed correlations of spatial ability with science achievement measures, while statistically significant are relatively small, in the .3 range. Spatial abilities are just one factor in science achievement and many other characteristics of individuals, such as general intelligence, inductive reasoning ability, mathematical/numerical ability and motivational factors are also likely to be important. Third, because non significant effects are less likely to be published (the *file drawer* effect), the published studies may over-represent the relations between spatial thinking and performance in science. Fourth, spatial ability is correlated with measures of general intelligence and many of the studies to do not control for the possibility that the correlation between spatial ability and science performance reflects common variance shared with general intelligence. However the recently published longitudinal data (Shea et al., 2007; Webb et al., 2001; Wai et al., 2009) argue against some of these concerns in that they included very large samples, controlled for verbal and mathematical ability, and examined broad and ecologically valid measures of science achievement.

To date, research on the correlation of spatial abilities with achievement in science has not focused on possible differences in the spatial demands of different sciences. Given the partial dissociation between large- and small-scale spatial abilities, a recent study questioned whether all sciences were equally demanding of spatial abilities and whether different sciences make varying demands on large-versus small-scale spatial abilities. Hegarty, Crookes, Dara-Abrams & Shipley (2010) used on-line self-report measures to collect preliminary data on the spatial abilities of scientists in different fields, as well as humanists and individuals in professional fields (total sample of over 700 professionals). Figure 3 shows the self standardized self ratings for each of the groups on large-scale spatial, small-scale spatial and verbal abilities and revealed interesting differences between the sciences. For example, geoscientists had the highest self-report ratings of both environmental and small-scale spatial abilities, whereas geographers had relatively high self ratings of environmental spatial abilities and engineers had relatively high self ratings of small-scale spatial abilities. Other scientific disciplines did not differ from the mean in self reported spatial abilities. Self ratings of verbal ability were uncorrelated with self ratings of spatial abilities and, as expected, were highest for humanities specialists. While these self-report data are intriguing, they need to be followed up with more objective measures of the spatial abilities of different scientists.

A final limitation of correlational studies is that correlations alone do not specify the direction of causality, that is, whether having good spatial abilities causes one to be good at science or whether the correlations reflect the fact that studying science enhances spatial abilities. Training studies, summarized in the next section, have begun to address this issue.

(2) Training Spatial Abilities:

Given that spatial abilities contribute to success in science, one potential way of improving performance in science disciplines is to train people's spatial skills. Here I examine whether taking science classes enhances spatial abilities, how training in spatial skills affects performance in science, and more basic research evidence for the trainability of spatial skills.

Enhancement of Spatial Abilities by Taking a Science Class. Some researchers have reported that just taking a science course can improve people's performance on spatial ability tests. Orion, Ben-Chaim & Kali (1997) found that students' spatial abilities improved significantly as a result of taking a geology course. Pallrand and Seeber (1984) tested students' spatial abilities before and after taking physics and other college classes and found that while all groups scored higher on the posttest, the gains were greater for students in physics classes than for those in liberal arts classes. However, Hegarty, Keehner, Khooshabeh & Montello (2009) found that although dental education improved student's ability to perform spatial skills specific to dentistry, it did not improve their spatial skills more generally.

Training in the Context of Science Classes. There have been several studies that aimed at training spatial skills, in the context of science classes. Some of these have focused on specific spatial skills such as inferring cross sections of three dimensional structures (Brinkman, 1996; Lord, 1985; Provo, Lamar & Newby, 2002). Lord (1985) used wooden models of geometric solids to train biology students to recognize cross-sections of primitive figures. The trained group outperformed the control group on a post-training cross section recognition task, as well as on an aggregate measure of spatial orientation and spatial visualization. Similarly, Brinkmann (1966) had some success using folding cardboard patterns and wooden geometric forms to teach spatial concepts in geometry, such as the characteristics of points, lines, angles, planes, and solids.

Other studies of this type have focused on spatial skills and knowledge in a specific domain, for example the ability to visualize molecules in chemistry. Small and Morton (1983) selected two groups of students from an organic chemistry class and gave them training outside class that focused on either manipulating 3-D molecular models and interpreting diagrams or on basic chemistry nomenclature and concepts. The model training group performed 12% higher on spatial items on the final exam. Similarly, Palrand & Seeber (1984) gave additional training in drawing to a group of students in a physics class and found that their spatial ability scores improved more than a matched group who received instruction of a similar length on the history of physics and a matched group of physics students who received no additional instruction outside class.

The most extensive spatial training program to date has been carried out by Sorby (2009; Gerson, Sorby, Sysocki & Baartmans, 2001) who identified students in engineering classes with relatively low spatial skills and developed semester-long training courses in skills such as imagining projections, sections, and rotations. These classes, which included lectures, multimedia software and workbooks, consistently resulted in significant gains in performance on measures of spatial ability compared to control groups. In tests of different versions of the training courses, a critical feature seems to be the use of workbooks in which students have to sketch the results of imagining sections, rotations etc, rather than merely observing them. Taking the training courses was also related with higher grades in engineering and science classes and with retention rates in college, especially for female students.

Psychology Laboratory Studies. There is also emerging support from psychology laboratory studies that aspects of spatial intelligence can be improved with relevant experience. Performance on tests of spatial ability and laboratory tasks such as mental rotation can be improved with practice (e.g., Kail, 1986; Wright, Thompson, Ganis, Newcombe, & Kosslyn, 2008), even if this occurs in the context of playing video games (Feng, Spence & Pratt, 2007; Terlicki, Newcombe & Little, 2008). There is also evidence that the effects of training transfers to other spatial tasks that are not practiced (Leone, Taine & Droulez, 1993;Wright et al., 2008), and that it endures for months after the training experience (Terlicki et al., 2008).

Issues in Training Studies. There is now good evidence that performance on spatial tasks improves with training, but there are important issues that need to be taken into account when evaluating the results of training studies. First, performance on spatial tests typically improves by just taking the test a second time (Uttal, Hand & Newcombe, 2009), so it is important to evaluate training interventions against control conditions. Second, there are questions of what actually improves after practice or training in spatial skills. For example, in some studies, the performance gains were very specific to the stimuli on which people were trained, suggesting that people just gained familiarity with those specific stimuli (Kail & Park, 1990; Tarr & Pinker, 1989; Sims & Mayer, 2002) whereas others suggested that people learned specific mental transformation processes (e.g., mental rotation) that could be applied to many stimuli (Leone et al., 1993; Wright et al., 2009). These results raise issues about how we should train spatial skills for science, for example, should training be in the context of specific content (e.g., facility in rotating molecules in chemistry) or can we identify fundamental spatial processes (projection, rotation, sectioning etc.) that are common to many sciences and train these processes in a way that they can be applied to any 3-D structure? Important questions also remain about the amount of training necessary for lasting gains.

(3) Focused Studies of Comprehension of Graphical Displays in the Sciences.

Scientists use a variety of graphical displays, such as diagrams, maps and graphs, to characterize the objects of study and more abstract data in their respective domains. As part of science instruction, students need to develop skills in constructing, interpreting, transforming and coordinating these domain-specific external representations. This set of skills has been referred to as *representational competence* (Ainsworth, 2006; Kozma & Russell, 1997; Kozma, Chin, Russell, & Marx, 2000; Wu, Krajcik, & Soloway, 2001).

Types of Graphical Displays and Basic Theories of Representation. Graphical displays are visual-spatial arrays that *represent,* or are symbols for, objects, events, or more abstract information. To understand the nature of a representation, we must specify the nature of the object or entity that is represented (i.e., the referent of the representation or *represented* world), the representational medium (i.e., the *representing* world), and the correspondence between these two (Palmer 1978). Graphical displays can be categorized based on the relation between the representation and its referent and how space is used to convey meaning.

There are two major categories of graphical displays. The first category consists of iconic displays. In iconic displays, space on the page represents space in the world and the properties displayed (shape, color, etc) are also visible properties of the referent (e.g., a road map, a cross section of a part of the anatomy). These types of displays are ancient (cave paintings are examples) and used by all cultures (Tversky, 2001). Note that these displays are not just pictures. They often abstract and distort information (e.g., when a subway map shows the connectivity between stations but distorts the distance), show views of objects that cannot be seen in the real world (e.g., cross sections) and represent objects that are too small or too large to be seen with the naked eye (e.g., molecules, galaxies).

The second category, relational displays are metaphorical in that they represent entities that do not have spatial extent or visible properties (e.g., when a cladogram shows the ancestral relationship between taxa, or a graph shows the velocity or acceleration of an object over time). In these displays, visual and spatial properties represent entities and properties that are not necessarily visible or distributed over space. Visual-spatial variables, such as color, shape, and location are the *representing* dimensions of the display (Bertin, 1983). The *represented* dimensions can be any category or quantity. These types of displays are a relatively recent invention. For example the invention of the graph is attributed to Playfair in the 18th Century.

Graphical displays can enhance spatial thinking in several ways. They store information externally, freeing up working memory for other aspects of thinking.They can organize information by indexing it spatially, reducing search and facilitating integration of related information. These displays can also allow the offloading of cognitive processes onto perceptual processes, for example when graphing data reveals a linear relationship between variables. When a display is interactive, people can also offload internal mental computations on external manipulations of the display itself. However, although graphical displays can enhance thinking in all of these ways, this does not mean that their use is easy or transparent. The following sections summarize examples of focused studies of how students understand specific spatial representations in biology, chemistry, geosciences, and physics, including difficulties that students have with these representations.

Examples from Biology. Cross sections (e.g., of anatomy) are a common type of external representation used in biology. Cross sections represent a 2-D slice of a 3-D object and understanding these representations is an example of what Kali and Orion (1996) refer to as “penetrative thinking” in that it involves inferring the inside structure of an object that is usually seen from the outside. As reviewed above, students often have difficulty inferring what a cross section of a 3-D object will look like, and these difficulties are often related to poor spatial abilities (Cohen & Hegarty, 2007; Hegarty et al., 2009; Provo et al., 2002; Russell-Gebbett, 1985;).

Biologists also summarize their findings in relational diagrams, for example when they show evolutionary relationships between taxa in cladograms. Novick and Catley (2007) have found that students in evolutionary biology classes often have difficulties interpreting these diagrams and that these difficulties can depend on the diagrammatic conventions used. For example, they found differences in comprehension of two types of common hierarchical diagrams (cladograms) in evolutionary biology; tree and ladder diagrams. Ladder diagrams were more difficult. This work illustrates an important theoretical points that cognitive scientists have made about representations. First, it is an example of the distinction made by Larkin and Simon (1987) between *information equivalence* and *computational equivalence* of representations, that is, representations that contain the same information but are not equally easy to comprehend and use. Novick and Catley’s (1997) work indicates that while tree and ladder formats of cladograms are informationally equivalent, they are not computationally equivalent

Examples from Chemistry: Understanding and relating external representations is particularly challenging in the study of chemistry. Chemists' inability to directly perceive atoms and molecules has led them to develop a wide range of external representations. Students must not only master representational systems unique to chemistry, but they must also relate them to more abstract representations such as equations, and they must learn the limited applicability of each representation for solving unique problems in the domain.

Researchers have identified two specific challenges of achieving representational competence in chemistry. The first is relating representations at macro and micro scales and more abstract symbolic levels (Barak & Dori, 1999; Kozma & Russell, 1997; Kozma, et al., 2000; Wu, Krajcik & Soloway, 2001; Wu & Shah, 2004). The second is relating different representations at the micro level, for example, judging whether two diagrams represent the same molecule from different perspectives (Stieff, 2007) or translating between different diagrams (with different conventions) for the same molecule (Stieff & Raje, 2010; Stull, Hegarty, Stieff & Dixon, 2010).

For example, in a current study (Stull et al., 2010) we are examining organic chemistry students' ability to translate between three types of diagrams used in organic chemistry (dash-wedge diagrams, Newman projections and Fischer projections). These show the structure of the molecules from different perspectives and use different conventions to represent the three dimensional structure of the molecule in the two dimensions of the page. One contribution that cognitive psychology can make to this problem is a specification of the cognitive processes that must be carried out to translate from one diagram format to another and the different strategies that experts and novices use to make these translations. For example, to translate from a dash-wedge diagram to a Newman diagram using a visual imagery strategy a student must first decode the conventions of the dash-wedge formulism, to infer the 3-d structure of the molecule, then rotate a representation of that structure, then project the 3-d image onto a 2-d image and finally recode the 2-D image using the conventions of a Newman diagram.

In research on diagram translations we have found that providing students with a 3-D ball-and-stick model of the molecule can improve performance on the diagram translation task, but many students do not use the models and success is strongly dependent on whether and how they manipulate the models (Stull, Hegarty, Stieff & Dixon, 2010). In other research, Stieff and colleagues have found that experts often use analytic strategies for translating between the two 2-D representations without imagining the 3-D structure of the molecule (Stieff & Raje, 2010) and while novices begin by using an imagistic strategy, many switch to using analytic strategies over the course of instruction in organic chemistry (Stieff, Hegarty & Dixon, 2010).

Examples from Geosciences: As described above, geology students often have particular difficulties imagining the internal structure of the earth's crust from external structures (Kali & Orion, 1996; Orion, et al, 1997) and understanding the concepts of strike and dip (Liben et al., in press). Kali and Orion (1996) developed a test of spatial abilities in the context of geology in which students had to infer and draw cross sections and other views of block diagrams of geological structures. They found two types of incorrect answers, "non-penetrative" answers in which students based their answers on patterns on the outside of the geological structures and "penetrative" answers, in which students attempted to visualize the internal structure but did not do so correctly. In general, students who gave non-penetrative answers had the poorest performance on the test.

In another geoscience domain, meteorology, novice students often have difficulties interpreting weather maps, which are iconic in the sense that they show the layout of a space, but also superimpose various “visualizations” of non-visual properties, such as pressure and temperature. In a series of studies comparing expert and novice comprehension of weather maps, Lowe (1994, 1996; 2004) found that novices and experts view weather maps differently. Novices tend to focus primary on a weather map’s superficial features, whereas experts focus on elements that are thematically relevant to the task at hand. Other research has shown that both novices and experts are more efficient and more accurate in making inferences from weather maps if the maps present only task relevant information, or make task relevant information more salient in the display (Hegarty, Canham & Fabrikant, 2010; Hegarty, Smallman, Stull, & Canham, 2009). However, they prefer to use more complex, detailed maps, which in fact impede performance.

More generally, Hegarty, Smallman et al (2010) found that undergraduates show a strong preference to use more realistic displays for a variety of learning and thinking tasks. They prefer animated to static displays, 3-D depth displays to 2-D displays, more realistic to less realistic displays etc. although in fact these display enhancements do not always improve task performance and can even impede it (e.g., Khooshabeh & Hegarty, 2010; Tversky, Morrison & Betrancourt, 2002). In sum, students have limits of representational metacompetence (diSessa, 2004) in that they prefer displays that simulate the real world with greater fidelity, but are often better served by simpler, more abstract displays (see also, Scaife & Rogers 1996).

Examples from Physics: In the domain of physics, researchers have identified problems that students experience in interpreting graphs of motion. A common misconception, identified in developmental studies is that children often interpret graphs literally, as if they are pictures. Kozhevnikov, et al., (2002) found the same misconception in college students with low spatial ability when they viewed graphs showing changes in velocity and acceleration over time. These students also had difficulty relating different graphs of the same data, for example judging whether a graph of velocity over time referred to the same situation as a graph showing acceleration over time.

Finally, cognitive scientists have conducted extensive studies of the cognitive processes involved in inferring motion from static diagrams of mechanical systems, often referred to as “running a mental model”. These studies have revealed naïve misconceptions that need to be overcome in achieving correct mechanical understanding, for example, many beginning physics students believe that when a ball emerges from a curved tube, it will follow a curved path (McCloskey, 1983; Kaiser, Proffitt, Whelan & Hecht, 1992. In other work, cognitive scientists have examined the mental processes involved in imagining how different components of a mechanical system (such as a spring, gear or pulley system) move when the machine is in motion. This research has revealed that people use a combination of imagistic and analytic thinking in this “mental animation” process (Clement, 2009;Hegarty, 1992; Hegarty & Sims, 1994; Schwartz & Black, 1996). For example they first decompose the machine into elementary mechanical components and infer their motion piecemeal. When the problems are novel, they use imagery processes to mentally simulate how each component will move, but they then infer rules of reasoning from their mental simulations (for example, they notice the regularity that every other gear in a gear chain turns in the same direction) and then switch from imagery-based thinking to rule-based thinking.

Issues in Focused Studies of Comprehension of Graphical Displays. In depth studies of the use of graphical representations in science necessarily look at specific representations and tasks. This raises questions of how representative the studies reviewed here are of their respective disciplines. There are also important questions of how spatial representations differ across the various science disciplines, and whether there are common challenges faced by students in these different disciplines. For example, different sciences, such as chemistry and geology, deal with very different scales of time and space, although both are centrally concerned with spatial structures. Some scientific disciplines such as geology and anatomy may depend more on penetrative thinking in that they are concerned with 3-D objects with internal structures, whereas others including geology, mechanics and meteorology are centrally concerned with motion, so they may depend more on mental animation ability. Sciences may also differ in the degree to which they depend on graphical representations versus more abstract formulas and equations. For example, a greater proportion of the space in geoscience journals is taken up by graphics, suggesting that geosciences be particularly dependent on spatial representations (Kastens, 2009).

In spite of these differences between the sciences, we can observe several common themes relating to visual-spatial representations in science. First, college students have difficulty understanding and using many visual-spatial representations in science and relating different representations of the same entity/phenomenon. Relating 2-D and 3-D representations seems to be particularly difficult for students with low-spatial ability. Second, students have particular difficulty understanding the abstract nature of representations, have a bias to view representations as realistic representations of reality, and prefer representations that resemble their referents. Third, knowing the conventions of visual-spatial representations alone is not sufficient for good comprehension and inference with these representations. Experts see patterns in representations that are not seen by novices even when they are familiar with the conventions. Novices are distracted by salient surface features that may not be task relevant. Fourth, making inferences from visual-spatial representations can involve imagistic thinking but also involves more analytic strategies such as task decomposition and rule-based reasoning.

College science is taught by domain experts who are already very familiar with the representations used in their field of study and may not realize how difficult the representations are for students to master. These observations suggest that we may need to pay more attention to introducing students to the various graphical displays that are used in scientific domain, including explicitly pointing out the relations between different displays of the same or related information, teaching students how representations with different levels of abstraction are optimized for various tasks, and giving students extensive practice on relating and translating between representations.

(4) Use of Dynamic and Interactive Visualizations in Teaching Science

In recent years, with advances in computer graphics and human-computer interaction techniques, dynamic and interactive displays have become commonplace. Technologies such as animations, interactive computer visualizations and virtual models are used by expert scientists and have enormous potential for promoting representational competence and spatial thinking in science. However, it is important to realize that these are often more complex than the static spatial representations discussed earlier, and so they can be even more demanding of spatial abilities, and representational competence.

Animations. Intuitively, it seems that animations should be helpful to students in providing an external display of information that might be difficult for students to visualize internally. However, initial results from studies comparing animated to static displays (e.g., to explain dynamic mechanical and biological processes) were disappointing and indicated that animations were not more effective than static diagrams (see Tversky, Morrison & Betrancourt, 2002 for a review). Tversky et al. (2002) point out that animations are often too fast and complex so that they present more information than can be accurately perceived and comprehended in the given time. Animations can also give an illusion of understanding, and can even cause students to “see” what they believe to be true, rather than what is actually true (e.g., Kriz & Hegarty, 2007). Hegarty, Kriz & Cate (2003) found that people learned better from an animation of how a machine works if they were first asked to predict how they thought the machine worked (perhaps, because they were first made aware of what they did not understand).

Interactive 3-D Visualizations. Given that people have difficulty understanding 2-D representations of 3-D objects, there has been much interest in the use of 3-D visualizations in education. For example, these visualizations migh be rotated at will to see different views of a structure, or sliced to see different cross sections of a structure. A recent example is a study in which Chariker, Naaz & Pani (in press) developed a computer-based visualization tool for learning neuroanatomy. Students interacted with the system to learn either whole anatomy of the brain followed by sectional anatomy or sectional anatomy alone. Learning was efficient and was retained well weeks after the learning phase (which included several learning sessions spaced over approximately 5 weeks). Furthermore learning whole anatomy transferred to learning sectional anatomy and both transferred to interpretation of biomedical images such as MRI scans.

Other studies with these types of visualizations have suggested that not all students can use them effectively. For example, Garg, Norman, Spero & Maheshwari (1999) taught medical students the anatomy of the carpal bones (of the wrist) using either an interactive 3-D model (that could be rotated interactively) or using two key views of the front and back of the hand. While high-spatial students learned well with either format, low spatial students learned better with the two key views than with multiple views. Keehner, Hegarty, Cohen, Khooshabeh & Montello (2008) examined use of an interactive 3-D visualization to perform a task that involved imagining the cross section of a three-dimensional object. There were large individual differences in how much people interacted with the models. Those who used the interactive models more often had better task performance, but surprisingly, participants who passively viewed the interactions of these model users were just as effective. This study makes it clear that just providing people with an interactive visual display does not ensure that they will use it effectively, and questions whether interactive displays are always better.

Linked Visualizations. A common format for teaching multiple representations in science is to provide students with interactive computer displays that show multiple linked representations of the same phenomenon, allowing students to manipulate one representation and observe its effects on the other. An early example of this approach was with the "thinkertools" curriculum in the domain of physics - mechanics (White, 1993) and has been used in several studies in the domain of chemistry (e.g., Kozma et al., 1996; Stieff & Wilensky, 2003; Wu, Krajcik & Soloway, 2001). For example in a visualization of this type, a student might relate a chemical formula or equation in one window, to a molecular model referring to the same substance in a second, and a video of an experiment showing a chemical reaction in another. Evaluation studies both in and out of the classroom have begun to indicate that student achievement and understanding of domain concepts improves after using these multi-representational displays (Kozma et al., 1996; Stieff & Willensky, 2003; Wu, et al., 2001). While this approach appears very promising, we need more controlled studies that compare the use of linked visualizations with more standard approaches to teaching chemical representations (i.e., "business-as-usual" control groups) and there are many questions about how to "scaffold" learning with these visualizations.

Issues in the Use of Dynamic and Interactive Visualizations. In summary, new technologies (animation, 3-D virtual models, interactive visualizations) have enormous potential for undergraduate education in science, but we are just beginning to understand how to best incorporate these technologies in instruction. While it is tempting to believe that dynamic (animated), 3-D and interactive visualizations might compensate for lack of internal visualization ability, research to date has suggested that they often depend on internal visualization ability for their use. That is, less able students often have difficulty manipulating visual-spatial representations and understanding how they can be used, just as they have more difficulty with more traditional static visual-spatial representations.

The research to date suggests several important avenues for the use of new visualization technologies in education. One avenue is to improve the visualizations so that they are more easily apprehended by students. A second avenue is to teach students explicitly how to use visualizations, given that students do not always discern how to best use interactive visualizations to accomplish the task at hand. Third, we need a better understanding of how interactive visualization technologies can and should be incorporated with more traditional instruction in the sciences, how to scaffold students learning with these visualizations and how to engage students in using these visualizations in a way that they learn the relevant concepts and are not mislead by an illusion of understanding.

Conclusions:

In summary, spatial thinking is a central component of scientific thinking and scientists have developed a large number of graphical displays to characterize the objects of study and more abstract data in their respective domains. Spatial ability is correlated with performance in college science courses, and students often struggle to understand and use spatial representations in learning. Cognitive studies are providing a better understanding of the difficulties that students have in understanding and using spatial representations and this in turn is informing studies of how to best nurture spatial thinking in science, with and without the use of new interactive technologies. It is important to continue to empirically study relations between spatial thinking and science achievement, and to evaluate the research evidence critically.

References:

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183-198.

Barak, M., & Dori, Y. J. (2004). Enhancing undergraduate student's chemistry understanding through Project-based learning in an IT environment. *Science Education 89*, 117-139.

Bertin, J. (1983). *Semiology of graphics: Diagrams networks maps* (W. Berg, Trans.). Madison: University of Wisconsin Press.

Bodner, G. M. & McMillan, T. L. B. (1986). Cognitive restructuring as an early stage in problem solving. Journal of Research in Science Teaching, 23, 727-737.

Brinkmann, E.H. (1966). Programmed instruction as a technique for improving spatial visualization. *Journal of Applied Psychology, 50*(2), 179-184.

Carter, C. S., LaRussa, M. A., & Bodner, G. M. (1987). A study of two measures of spatial ability as predictors of success in different levels of general chemistry *Journal of Research in Science Teaching, 24*(7), 645-657.

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

Chariker, J. H., Naaz, F. & Pani, John R. (in press). Computer-based learning of neuroanatomy: A longitudinal study of learning, transfer, and retention. *Journal of Educational Psychology*.

Clement, J. (1009). Creative model construction in scientists and students. The role of Imagery, analogy and Mental Simulations. Dordrecht: Springer.

Cohen, C. A., & Hegarty, M. (2007). Individual differences in use of an external visualization while performing an internal visualization task. *Applied Cognitive Psychology, 21,* 701-711.

Contreras, M. J., Colom, R. Hernandez, J. M., & Santacreu, J. (2003). Is static spatial performance distinguishable from dynamic spatila performance? A latent-variable analysis. The *Journal of General Psychology, 130*, 277-288.

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction, 22*, 293-331.

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*, 850-855.

French, J. W. (1951). Description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monographs*, 5.

Garg, A. X., Norman, G. R., Spero, L., & Maheshwari, P. (1999). Do virtual computer models hinder anatomy learning? *Academic Medicine, 74* (10), S87-S89.

Garg, A. X., Norman, G., & Sperotable, L. (2001). How medical students learn spatial anatomy. *The Lancet, 357*, 363-364

Gerson, H. B. P., Sorby, S.A., Wysocki, A & Baartmans, B. J. (2001). The development and assessment of multimedia software for improving 3-D visualization skills. *Computer Applications in Engineering Education, 9*, 105-113.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18(5)* 1084-1102.

Hegarty, M., Canham, M., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory and Cognition, 36,* 37-53.

Hegarty, M.., Crookes, R., Dara-Abrams, D. & Shipley, T. F. (2010) Do all science disciplines rely on spatial abilities? Preliminary evidence from self-report questionnaires. In C. Holscher (et al) (Eds.) *Proceedings of Spatial Cognition 2010*. Berlin: Springer.

Hegarty, M., Keehner, M., Khooshabeh, P., & Montello, D. R. (2009). How spatial ability enhances, and is enhanced by, dental education. *Learning and Individual Differences, 19,* 61-70.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition & Instruction, 21,* 325-360.

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006) Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence, 34,* 151-176.

Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory & Cognition, 22,* 411-430.

Hegarty, M. Smallman, H. S., Stull, A. T., & Canham, M. (2009). Naïve cartography: How intuitions about display configuration can hurt performance. *Cartographica, 44,* 171-187.

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.). *The Cambridge Handbook of Visuospatial Thinking*. Cambridge University Press (pp. 121–169).

Hunt, E., Pellegrino, J. W., Frick, R. W., Farr, S. A. & Alderton, D. (1988) The ability to reason about movement in the visual field. *Intelligence, 12*, 77-100.

Isaak, M. I. & Just, M. A. (1995). Constrains on the processing of rolling motion: the curtate cycloid illusion. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1391-1408.

Kail, R. (1986), The impact of practice on rate of mental rotation. *Journal of Experimental Child Psychology, 42,* 378-391.

Kail, R, & Park, Y-S. (1990). Impact of practice on speed of mental rotation. *Journal of Experimental Child Psychology, 49(*2), 227-244.

Kaiser, M. K., Proffitt, D. R., Whelan, S. M. & Hecht, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception & Performance,* 18, 669-689

Kali, Y., & Orion, N. (1996). Spatial abilities of high-school students in the perception of geological structures. *Journal of Research in Science Teaching, 33*(4), 369-391.

Kastens, K. (2009). Synthesis of research on thinking & learning in the geosciences: Developing representational competence. Presented at the annual meeting of the Geological Society of America, October 18-21.

Keehner, M. Hegarty, M., Cohen, C. A., Khooshabeh, P., & Montello, D. R. (2008). Spatial reasoning with external visualizations: What matters is what you see, not whether you interact. *Cognitive Science, 32,* 1099-1132.

Keig, P. F., & Rubba, P. A. (1993). Translation of representation of the structure of matter and its relationship to reasoning, gender, spatial reasoning, and specific prior knowledge. *Journal of Research in Science Teaching*, *30,* 883-903.

Kooshabeh, P., & Hegarty, M. (2010). Inferring cross sections: When internal visualizations are more important than properties of external visualization. *Human Computer Interaction., 25,* 119-147.

Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction, 20*, 47-78.

Kozhevnikov, M. & Thornton, R. (2006). Real-time data display, spatial visualization ability and learning force and motion concepts. *Journal of Science Education and Technology, 15, 111 – 132.*

Kozma, R., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching, 34,* 949-968.

Kozma, R. B., Russell, J., Jones, T., Marx, N., & Davis, J. (1996). The use of multiple, linked representations to facilitate science understanding. In R. G. S. Vosniadou, E. DeCorte, & H. Mandel (Eds.), International perspective on the psychological foundations of technology-based learning environments (pp. 41-60). Hillsdale, NJ: Erlbaum.

Kriz, S., & M. Hegarty (2007). Top-down and bottom-up influences on learning from animation. *International Journal of Human-Computer Studies, 65* , 911-930.

Larkin, J. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science,* 11, 65-100.

Leone, G., Taine, M. & Droulez, J. (1993). The influence of long-term practice on mental rotation of 3D objects. *Cognitive Brain Research 1:* 241-255.

Liben, L. S., Kastens, K. A., Christensen, A. E. (in press). Spatial foundations of science education: The illustrative case of introductory geological concepts. *Cognition and Instruction.*

Lord, T. (1985).Enhancing the visuo-spatial aptitude of students. *Journal of Research in Science Teaching 22,* 395-405.

Lowe, R. K. (1994). Selectivity in diagrams: Reading beyond the lines. *Educational Psychology, 14*, 467-491.

Lowe, R. K. (1996). Background knowledge and the construction of a situational representation from a diagram. *European Journal of Psychology of Education, 11,* 377-397.

Lowe, R.(2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction, 14,* 257-257.

McCloskey, M. (1983). Intuitive physics. *Scientific American, 248* (4), 122-130.

National Research Council. (2006). *Learning to think spatially*. Washington, D.C.: National Academies Press.

Novick, L. R. & Catley, K. M. (2007). Understanding phylogenies in biology: The influence of a Gestalt perceptual principle. *Journal of Experimental Psychology: Applied*, *13*, 197-223,

Orion, N., Ben-Chaim, D., & Kali, Y. (1997). Relationship between earth-science education and spatial visualization. *Journal of Geoscience Education, 45*, 129-132.

Pallrand, G. J. and Seeber, F., (1984). Spatial ability and achievement in introductory physics, *Journal of Research in Science Teaching, 21*, 507.

Palmer, S. E. (1978). Fundamental aspects of cognitive representations. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 259-303) Mahwah, NJ: Lawrence Erlbaum Associates.

Pribyl, J. R. & Bodner, G. M. (1987). Spatial ability and its role in organic chemistry: A study of four organic courses. *Journal of Research in Science Teaching,* 24, 3, 229-240.

Provo, J., Lamar, C., & Newby, T. (2002). Using a cross section to train veterinary students to visualize anatomical structures in three dimensions. *Journal of Research in Science Teaching, 39,* 10-34.

Rochford, K. (1985). Spatial learning disabilities and underachievement among university anatomy students. *Medical Education, 19*, 13-26.

Russell-Gebbett, J. (1985). Skills and strategies: Pupils' approaches to three-dimensional problems in biology. *Journal of Biological Education, 19* (4), 293-298.

Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work? *International Journal of Human-Computer Studies, 45*.

Schwartz, D. L., & Black, J. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science, 20*, 457-497.

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study*Journal of Educational Psychology, 93*(3), 604-614.

Sims, V. K., & Mayer, R. E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology, 16*, 97-115.

Small, M. Y. & Morton, M. E. (1983). Research in college science teaching: Spatial visualization training improves performance in organic chemistry.*Journal of College Science Teaching,* 13, 1, 41-43.

Sorby, S. (2009). Educational research in developing 3D spatial skills for engineering students. *International Journal of Science Education, 31* (3), 459-480.

Stieff, M. (2007). Mental rotation and diagrammatic reasoning in science. *Learning and Instruction, 17*(2), 219-234.

Stieff, M., Hegarty, M., & Dixon B. L. (2010). Alternative strategies for spatial reasoning with diagrams. In A. Goel, M. Jamnik & N.H. Narayanan (Eds.) *Diagrammatic Representation and Inference (Proceedings of Diagrams 2010).* Berlin: Springer-Verlag

Stieff, M., & Raje, S. (2010). Expert algorithmic and imagistic problem solving strategies in advanced chemistry. *Spatial Cognition and Computation, 10* (1), 53-81.

Stieff, M., & Wilensky, U. (2003). Connected Chemistry - Incorporating interactive simulations into the chemistry classroom. Journal of Science Education and Technology, 12(3), 285-302.

Stull, A. T., Hegarty, M. Stieff, M. & Dixon, B. (2010). Concrete models as aids to representation translation of molecular diagrams? Presented at the Annual Meeting of the Cognitive Science Society, Portland OR, August.

Tarr, M.J., & Pinker, S.J. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*(2), 233-282.

Terlecki, M. S., Newcombe, N. S., & Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Applied Cognitive Psychology, 22*(7), 996-1013.

Tversky, B. (2001). Spatial schemas in depictions. In M. Gattis (Ed.). *Spatial schemas and abstract thought.* Cambridge, MA: MIT Press.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate? International Journal of Human-Computer Studies, 57, 247-262.

Uttal, D. H., Hand, L. L. & Newcombe, N. S. (2009). Malleability of spatial cognition: Results of a meta-analysis. Presented at the biennial meetings of the Society for Research in Child Development, Denver.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817-835.

Wallace, B, & Hofelich, B. (1992). Process generalization and the prediction of performance on mental imagery tasks. *Memory & Cognition, 20*(6), 695-704.

Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. *Journal of Educational Psychology, 99*(2), 397-420.

White, B. Y. (1993) ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction, 10,* 1-100.

Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., & Kosslyn, S. M. (2008). Training generalized spatial skills. *Psychonomic Bulletin & Review, 15*(4), 763-771.

Wu, H.-K., Krajcik, J. S. & Soloway, E. (2001). Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching,* 38, 7, 821-842.

Wu, H., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning *Science Education, 88* (3), 465-492.

Zhang, J. & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18,* 87-122.

Figure 1: Examples of tests of Spatial Visualization Ability.

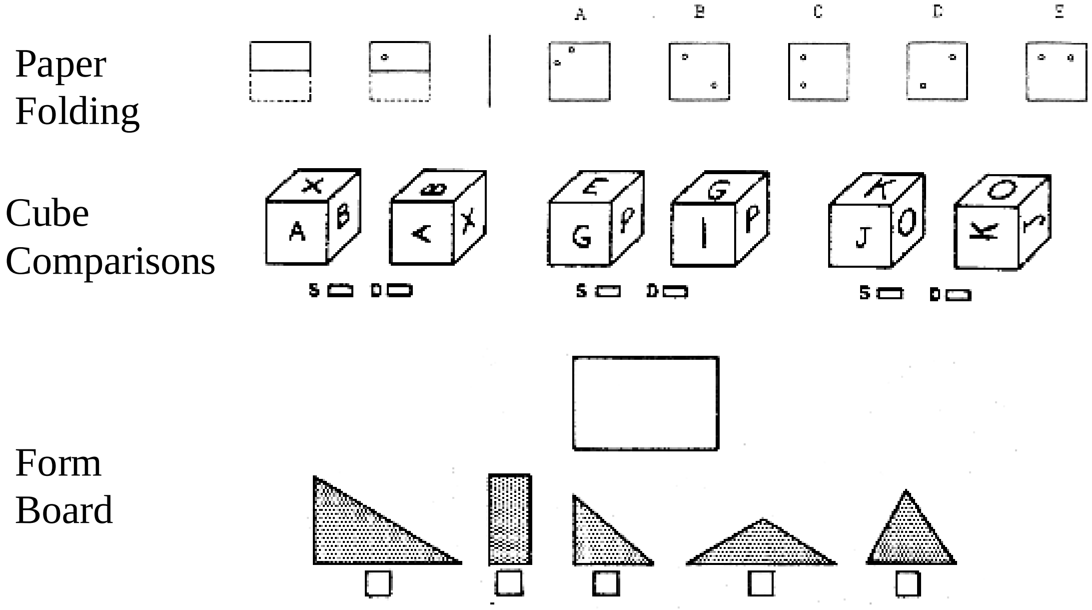
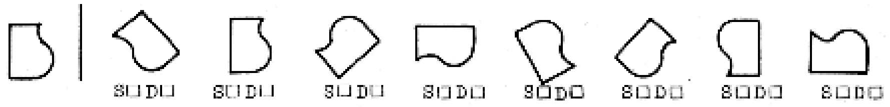
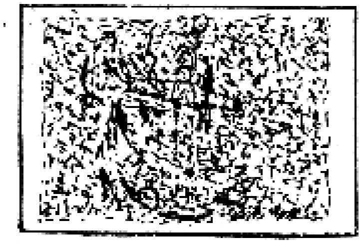


Figure 2: Examples of test items from tests of spatial ability factors identified by Carroll (1993)

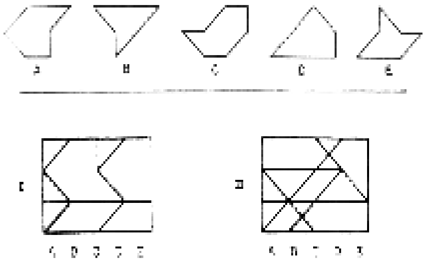
Spatial Relations:



Closure Speed:



Flexibilty of Closure:



Perceptual Speed:

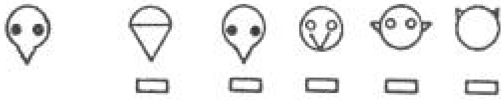


Figure 3: Standardized scores showing self ratings of scientists in different disciplines in large-scale spatial ability (blue bars), small scale spatial ability (green bars) and verbal ability (red bars).