Renata Gorska, Cracow University of Technology
Renata A. Gorska is on the faculty in the school of Architecture at the Cracow University of Technology in Cracow, Poland. She spent the 1997-98 academic year on sabbatical to Michigan Technological University. Her research interests are in assessing and training for spatial skills.

Sheryl Sorby, Michigan Technological University
Sheryl A. Sorby is a Professor of Civil and Environmental Engineering and is the former Associate Dean for Academic Programs in the College of Engineering at Michigan Technological University. She is currently serving as a Program Director in the Division of Undergraduate Education at the National Science Foundation. Sorby is active in the American Society for Engineering Education and the American Society of Civil Engineers. She is a past chair of the Engineering Design Graphics Division of ASEE. She was a recipient of the Dow Outstanding New Faculty Award and the Distinguished Teaching Award, both from the North Midwest Section of ASEE. Her research interests include spatial visualization and computer aided design.
TESTING INSTRUMENTS FOR THE ASSESSMENT OF 3-D SPATIAL SKILLS

Abstract — Assessment and improvement of spatial skills is a significant topic of research in the engineering graphics community. In order to conduct this type of research, various instruments are available for assessing spatial skill levels. This paper describes several instruments used by the authors with various audiences in conducting research in spatial skills development. Typical scores obtained in the research are also presented.

INTRODUCTION: THE THEORY BEHIND THE PRACTICE

Spatial visualisation abilities play a role important for success in engineering studies as many researchers have found (Leopold 2005, Sorby 2001, 2005, Juščáková 2003, 2004, Suzuki 2004, Tsutsumi 2004). A significant number of students enter technical studies without well-developed spatial skills. Thus one of the main objectives in teaching subjects such as descriptive geometry, engineering graphics, computer graphics and/or computer aided design, is to provide students with the kind of teaching materials and to introduce new methods for enhancing “seeing” and “thinking” in a three-dimensional world. The introduction of this paper covers various aspects of the theories of perception and visualisation. Following that, testing instruments used in research on spatial abilities will be described and the results of the experiments using these testing instruments will be presented.

SPATIAL ABILITY

According to L.L. Thurstone (1938) spatial ability is an important component of intellectual ability. Thurstone (1950) listed seven factors, three of these having to do with visual orientation in space, which he labelled S₁, S₂ and S₃. S₁ was interpreted as the “ability to recognize the identity of an object when it is seen from different angles”, or as “the ability to visualise a rigid configuration when it is moved into different positions” as in the Flags test (Figure 1). S₂ was interpreted as “the ability to imagine the movement or internal displacement among the parts in configuration”. This second factor S₂ was involved in tests of mechanical movement and surface development. The third spatial factor S₃ was said to represent “the ability to think about those spatial relations, in which the body orientation of the observer is an essential part of the problem”. In 1951 Thurstone carried out further studies on Mechanical Ability and reported ten factors extracted from the correlations of 32 spatial group tests given to 350 boys.

J.W. French (1951) carried out a number of factorial investigations prior to 1951. He called the three spatial factors: the space factor, spatial orientation and spatial visualisation respectively. The space factor represented “the ability to perceive spatial patterns accurately and to compare them with each other”. This factor related to perception of three-dimensional as well as two-dimensional space objects. The spatial orientation factor had not been clarified according to French, but seemed “to involve a person’s ability to remain unconfused by the varying orientations in which a spatial pattern may be presented” (Smith, p.86). Spatial visualisation factor was interpreted as “the ability to comprehend imaginary movement in three-dimensional space, or the ability to manipulate objects in imagination”.

In 1957, Michael, Guilford, Fruchter and Zimmerman attempted to synthesize the findings of research on spatial abilities. They formed three groupings of factors from which the psychologists distinguish between three categories of spatial abilities, namely spatial relations and orientation (SR-O), visualisation (Vz), and kinesthetic imagery (K). Michael et al. suggested that their SR-O factor was “more or less a composite of Thurstone’s S₁ and S₂ factors”, while their Vz factor was identical with Thurstone’s S₁ factor. Their K-factor might be considered identical to Thurstone’s tentatively identified K-factor (Smith, p.90).

Recent, as well as historical factor analytic studies, provide strong and consistent support for the existence of at least two distinct spatial abilities: spatial visualisation and spatial orientation (McGee, 1979). However, American psychologists are finding difficulties in clarifying the distinctions between the spatial factors Vz and SR-O. Spatial Visualization is an ability to mentally manipulate, rotate, twist, and pictorially invert presented visual stimuli. The underlying ability seems to involve a process of recognition, retention, and recall of a configuration in which there is movement among the internal parts of the configuration, or of an object manipulated in three-dimensional space, or the folding or unfolding of flat patterns. Spatial orientation involves comprehension of the arrangement of elements within a visual stimulus pattern, the aptitude to remain unconfused by the changing orientations in which a configuration may be presented, and the ability to determine spatial relations in which the body of the observer is an essential part of the problem.

Activities as disparate as perception of horizontality, mental rotation of objects, and location of simple figures within complex figures have all been referred to as measures of spatial ability (Linn and Petersen, 1985). Linn and Petersen
categorize spatial ability into three categories: (a) **spatial perception**, which can be done efficiently using a gravitational or kinesthetic process; (b) **mental rotation**, which can be done efficiently using a Gestalt-like mental rotation process analogous to physical rotation of the stimuli; (c) **spatial visualization**, which can be done efficiently using an analytic process. In **spatial perception** tasks, subjects are required to determine spatial relationships with respect to the orientation of their own bodies, in spite of distracting information. The example of such test is a Rod and Frame Test (RFT), in which subjects must place a rod vertically while viewing a frame oriented at 22° (Linn & Petersen, 1985). Another is water level, a task that requires subjects to draw or identify a horizontal line in a tilted bottle (Linn & Petersen, 1985). A widely acknowledged test used to measure student ability to rotate a two or three dimensional figure rapidly and accurately is the test developed by Shepard and Metzler (1971). Originally the test measured the speed of response to different amounts of rotation. Subsequently Vanderberg & Kuse (1978) modified the Shepard-Metzler Mental Rotation Test for group administration. The other test measuring this potential is the Flags and Cards from the French kit (French et al., 1963) as shown in Figure 1.

**Spatial visualization** is commonly associated with those spatial ability tasks that involve complicated, multi-step manipulations of spatially presented information. These tasks may involve the processes required for spatial perception and mental rotations but are distinguished by the possibility of multiple solution strategies. Spatial visualization tasks include EFT (Embedded Figure Test), Hidden Figures, Paper Folding, Paper Form Board, Surface Development, Differential Aptitude Test, Block Design, and Guilford-Zimmerman spatial visualization (Linn & Petersen, 1985).

**FIGURE 1**
The Flags Test: Spatial Relations test; the subject must indicate whether the two flags are the same: (S) if one can be slid around so that it is identical to the other or (D) for different flags

**VARIous TESTING Instruments USEd FOR S PATIAL ABILITY EVALUATION**

In conducting research in spatial skills and in the development of spatial skills, a standardized instrument is typically chosen to measure gains. There are several instruments from which to choose and the authors have varying degrees of experience in using these instruments to conduct educational research. The remainder of this paper discusses some of the common instruments as well as describing some that are in development. Typical results from the use of these instruments are also presented.

**Mental Rotation Test (MRT)**

The Mental Rotations Test (MRT) was developed by Vandenberg and Kuse (1978) and consists of 20 items. Each problem contains a criterion figure, which undergoes rotation. Among four alternatives, two are depicting the three-dimensional criterion figure after rotation, two choices depict a different object. Students are asked to identify which two of the alternatives are rotated images of the criterion figure. Solution time provided for the test is 4 minutes for the first set of 10 items, and after a short time interval another 4 minutes is provided to solve the other 10 problems. In the MRT, the components of spatial visualisation and mental rotation are tested. Figure 2 shows an example problem from the MRT.

**Figure 2. Example problem from the MRT**

Figure 3 shows the results of testing using the MRT with students in the Civil Engineering programs at the Cracow University of Technology over several years and Table 1 includes data on the gains measured on the MRT after...
participation in a semester-long descriptive geometry course. The data presented in Figure 3 and Table 1 shows that the average pre-test score for these students on the MRT is around 65% with a gain score of around 15% after participation in a descriptive geometry course. The gender differences in pre-test scores are evident from Figure 3; these gender differences are statistically significant. The data from Table 1 shows that in general, the female students experience higher gains on the MRT than do the male students, likely because they start with a lower mean pre-test. In testing done with first-year engineering students at Michigan Tech, pre-test scores on the MRT were generally at about the same level (mean score ~71%) as those obtained with the CUT students. Similar gains on the MRT (~10%) were observed after participation in an engineering graphics course at Michigan Tech.

![Figure 3. Mean MRT Pre-test Scores at CUT](image)

<table>
<thead>
<tr>
<th>Year</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>94/95</td>
<td>9.81 (n=82)</td>
<td>10.68 (n=17)</td>
</tr>
<tr>
<td>99/00</td>
<td>23.89 (n=133)</td>
<td>29.39 (n=41)</td>
</tr>
<tr>
<td>02/03</td>
<td>14.81 (n=77)</td>
<td>15.44 (n=39)</td>
</tr>
<tr>
<td>03/04</td>
<td>17.02 (n=68)</td>
<td>14.82 (n=27)</td>
</tr>
</tbody>
</table>

**Table 1. Mean Gain Scores on MRT for Students at CUT**

**Mental Cutting Test (MCT)**

The Mental Cutting Test (MCT) was first developed as part of a university entrance examination in the U.S. (CEEB, 1939) and has subsequently been applied in spatial visualization research in Japan, Australia, Germany, Poland and the U.S. The test consists of 25 items and the time provided for solution is 20 minutes. In each problem, a three-dimensional criterion figure is presented on the left-hand side of the problem with an imaginary cutting plane indicated. The students have to choose the correct resulting cross-section from among five alternatives. With the MCT, the components of spatial visualization and spatial relations are tested. A sample problem from the MCT is shown in Figure 4.

![Figure 4. Sample Problem from MCT](image)

Figure 5 shows the results of testing using the MCT with students in the Civil Engineering programs at the Cracow University of Technology over several years and Table 2 includes data on the gains measured on the MCT after participation in a semester-long descriptive geometry course in Poland. The data presented in Figure 5 and Table 2 shows that the average pre-test score for these students on the MCT is around 60% with a gain score of around 10% after participation in a descriptive geometry course. The gender differences in pre-test scores are evident from Figure 5; these gender differences are statistically significant. The data from Table 2 shows that in general, the female students experience higher gains on the MCT than do the male students, likely because they start with a lower mean pre-test. Once again, the pre-test scores for Michigan Tech students were similar (~60%) to those observed at CUT and the gains (~8%) after an engineering graphics course were approximately the same also.
In the fall of 2005, a subset of items from the MCT was administered to students at the middle and high school levels in the US. The items were selected to be of varying difficulty. Average scores on this subset of problems were 26.3% for middle school students and 36.3% for high school students. Gender differences in test scores, favoring males, were observed at both grade levels. However, on further examination of the results from this testing it was found that this subtest had a very low reliability (Cronbach’s alpha ~0.3) and item analysis seemed to indicate that guessing predominated. During the fall of 2006, a revised subset of items was administered to pre-college students. In the revised test, the ten easiest items were selected from the original MCT (the determination of the easiest items was made by co-author Sorby). For the revised instrument, the average for middle school students 36.5% and for high school students it was 42.6%; however, the test reliability was greatly improved. For the revised test, the Cronbach’s alpha was 0.67 which is within the acceptable range for test reliability.

**Differential Aptitude Test: Space Relations (DAT:SR)**

The Differential Aptitude Test: Space Relations (DAT:SR) consists of 50 items. The task is to choose the correct three-dimensional object from four alternatives that would result from folding the given two-dimensional pattern. In the DAT:SR, the components of spatial visualisation and spatial relations are tested. In testing done with first-year engineering students at Michigan Tech, average pre-test scores on the DAT:SR are generally higher than those observed for other testing instruments (~85%) and gain scores are generally lower (~5%). Gains are likely lower because with a high pre-test score, there is less “room” for students to improve their scores. Interestingly, gender differences on the DAT:SR are rarely observed; sometimes there are gender differences favouring female students on this test. An example problem from the DAT:SR is shown in Figure 6.

![Figure 6. Sample Problem from DAT:SR](image)

In the fall of 2005 and 2006, a subset of items from the DAT:SR was administered to students at the middle and high school levels in the US. The items were selected to be of varying difficulty. Average scores on this subset of problems was 50.4% for middle school students and 60.2% for high school students. No gender differences were observed for either grade level for this test and Cronbach’s alpha was ~0.68, indicating a reasonable test reliability for this audience.
The Purdue Spatial Visualization test: Rotations (PSVT:R) was developed by Guay (1977) to assess a person’s ability to visualize rotated solids. The test has 30 items and is scored as the percent correct. The top line of the problem presents an exemplary rotation of a three-dimensional object. In the second line, a different object is shown and the student selects from the five choices what this second object would look like if it were rotated by the same amount as the first object. In the PSVT:R, the components of spatial visualisation and mental rotation are tested. A sample problem from the PSVT:R is shown in Figure 7.

The PSVT:R has been used extensively at Michigan Tech over the past 15+ years for assessment of spatial visualization skills. In general, pre-test scores for first-year engineering students on the PSVT:R are ~78% at Michigan Tech with gain scores of ~7% observed after participation in an engineering graphics course. The PSVT:R has also been used at Michigan Tech to assess the effectiveness of a special spatial skills course. For the spatial skills course, average pre-test scores on the PSVT:R are generally ~50% with gains of ~28% after participation in the course. Thus, the post-test scores for the students in the spatial skills course (~78%) are generally equivalent to the pre-test scores for the general engineering population at Michigan Tech.

In the fall of 2005, a subset of items from the PSVT:R was also administered to students at the middle and high school levels in the US. The items were selected to be of varying difficulty. Average scores on this subset of problems were 47.9% for middle school students and 60.0% for high school students. Gender differences in test scores, favoring males, were observed at both grade levels. The Cronbach’s alpha for this subtest was 0.64 which is within the acceptable range for test reliability.

**Test of Spatial Imagination (TPS)**

The “Test of Space Imagination” was developed by Zuzana Juščáková from the Technical University of Košice (Slovakia) as part of a VEGA project (No 1/1407/04), granted by the Slovak Ministry of Education in 2004. TPS was developed based on an earlier TPP test (a testing instrument that was used as a measure of spatial ability of engineering students at a number of Slovak, Polish, and Czech Schools). The test consists of three distinct parts: Subtest 1, Subtest 2 and Subtest 3. Each of the parts of the test consists of 10 items. Solution times are 13 minutes for Subtest 1 and Subtest 2, and 8 minutes for Subtest 3. In the TPS test, the component of spatial relations is tested.

**Subtest 1: Parallel or Intersecting?**

The subtest 1 tasks deal with regular polyhedra. Specifically, either a tetrahedron, a cube or an octahedron is used for each item. A straight line KL and a plane ABC are defined by the characteristic points lying either on the edges or on the faces of each solid. The relationship of mutual position in three-dimensional space between the line KL and the plane ABC is under consideration, i.e., the relationship between the line KL and plane ABC is either parallel or intersecting. The task is to choose one item out of four given in which the relationship between line KL and plane ABC is different from relationships existing in the three other drawings. Figure 8 shows an example of a problem subtest 1 of the TPS. For this example, it is evident that the first item on the left side shows an intersection between line KL and plane ABC, while in the other drawings parallelism between the line KL and the plane ABC is observed.
The TPS was administered to first-year students at both the Cracow University of Technology and in various schools in Slovak and the Czech Republic. Mean scores on this subtest were 44.95% at CUT and 40.1% at the Slovak and Czech schools. Gender differences, favoring males were observed at all locations.

**Subtest 2: A Snake in a Cube**

The second subtest of the TPS instrument was titled “Snake in a cube.” For this test, subjects are told that the snake rolls and bends within the cube without overlapping parts or knots and it passes along the sides of the cube or inside its walls. It can bend in canters of the walls, it can turn at midpoints of edges of the cube or finally it can bend in the shape of a circular arc. Three views in the third angle projection method (European Method) are given. The task is to draw an axonometric view of a snake in a cube based on the given three views. Figure 9 shows an example problem from this subtest.

The second subtest of the TPS was administered to students at CUT and at various schools in Slovak and the Czech Republic. The average scores students at CUT were 57.3% and for the Slovak and Czech schools the average was 44.5%. Gender differences on this subtest were not observed for CUT but were observed for the students in the Slovak and Czech schools.

**Subtest 3: Complementary Parts of a Cube**

For this subtest of the TPS, students are to choose one of the given $a$, $b$, $c$, or $d$ parts of the cube, which when joined with the criterion part $K$ form a full cube, i.e., $K$ and the choice must be mating parts. Figure 10 shows an example problem for the third subtest of the TPS.

Subtest 3 of the TPS was administered to students at CUT and at various Slovak and Czech schools. The mean score at CUT were 72.7% compared to 62.6% at the other schools. Significant gender differences on this subtest were observed in both Poland and in the Slovak and Czech schools.

Table 3 summarizes the testing results from the three subtests of the TPS for the Polish, Slovak, and Czech schools.

<table>
<thead>
<tr>
<th></th>
<th>Cracow University of Technology (CUT)</th>
<th>Slovak and Czech Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subtest 1</strong></td>
<td>Group result</td>
<td>Subtest 2</td>
</tr>
<tr>
<td>Men</td>
<td>47.91 s.d.=24.53 (n=67)</td>
<td>58.75 s.d.=28.65 (n=64)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>76.15 s.d.=17.29 (n=65)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44.7 s.d.=24.4 (n=582)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54.2 s.d.=31.4 (n=582)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>69.7 s.d.=21.3 (n=582)</td>
</tr>
<tr>
<td><strong>Subtest 2</strong></td>
<td>57.28 s.d.=28.87 (n=92)</td>
<td>44.1 s.d.=24.2 (n=905)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44.5 s.d.=32.4 (n=905)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>62.6 s.d.=23.6 (n=905)</td>
</tr>
<tr>
<td><strong>Subtest 3</strong></td>
<td>72.69 s.d.=18.25 (n=93)</td>
<td>62.6 s.d.=23.6 (n=905)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54.2 s.d.=31.4 (n=582)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>69.7 s.d.=21.3 (n=582)</td>
</tr>
</tbody>
</table>
LAPPAN TEST

Lappan (1981) developed a test for assessment of spatial skills aimed at the middle school grades. The items assess student understanding of basics in isometric sketching and orthographic projection. However, upon careful examination of the test items it was determined that most problems did not conform to engineering standards of graphical representation. The test items were modified to reflect standard conventions. Figure 11 shows an original item from the Lappan test and Figure 12 shows the item after modification. In this case, students are presented with an isometric view of an object and are asked to identify a specific view (in this case the Back View) from five choices given.

Figure 11. Original Isometric Item from Lappan Test

Figure 12. Modified Isometric Item from Lappan Test

Figure 13. shows a second original item from the Lappan Test that assesses a student’s understanding of orthographic views and their relationship to coded plans. Figure 14 shows the same item after it has been modified to conform to engineering graphics conventions. For this type of problem, students are presented with three views of an object and must identify which partial coded plan could be used to define the object.

Figure 13. Original Orthographic Item from Lappan Test

Figure 14. Modified Orthographic Item from Lappan Test
Ten items from the modified Lappan test were administered to students in the spatial skills course at Michigan Tech during the fall 2007 semester. The average pre-test score for college freshmen on this test was 60%. The modified Lappan test was also administered to middle school and high school students during the fall of 2005 and 2006. At the middle school level, the average pre-test score was 34.0%; at the high school level, the average score was 47.8%. Significant gender differences were observed at both pre-college levels. The Cronbach’s alpha for this test was ~0.63 indicating a reasonable degree of test reliability.

CONCLUSIONS

From the data presented in this paper it appears that spatial skills can improve through instruction in various graphics courses, including descriptive geometry, engineering graphics, or a special course aimed at developing spatial skills. The pre-test scores for US and Polish students are roughly equivalent as measured by a variety of instruments. Significant gender differences are typically observed at both the pre-college or university levels; however, gender differences may or may not be apparent depending on the instrument you choose. For example, gender differences on the DAT:SR are typically not found.

As presented in this paper, various testing instruments have been used for evaluation of spatial skills of engineering students with various audiences in Europe and in the USA. In selecting an instrument for educational research in spatial skills development, it is important to know what you want to test for from the outset. Currently it appears that the most frequently and popular tests used to measure spatial skills among engineering students all over the world are the MRT and the MCT tests. However, it should be noted that there are newly developed instruments like the Lappan test or the TPS test, which give a good evaluation of students’ spatial skills. The TPS test was standardized with a representative group of subjects (n>1065) and gave good correlation results with typical intelligence tests. Because both the Lappan test and the TPS test provide evidence of a student’s level of understanding of the relationships between isometric and orthographic views it likely can be an adequate instrument for measuring spatial abilities among engineering students.

Students enter a technical university with varying levels of spatial abilities. If the average pre-test score is too high, one may not be able to observe gains in spatial skills development. This phenomenon is known as the “ceiling affect.” If the test is too difficult for the subjects, there will likely be a great deal of guessing on the instrument, resulting in unreliable data. Thus if you are considering conducting research in spatial skills development, you should choose your instrument carefully to achieve the results you desire.

REFERENCES

[1] CEEB Special Aptitude Test in Spatial Relations (MCT). Developed by the College Entrance Examination Board, USA, 1939.


