Journal for Geometry and Graphics
Volume 7 (2003), No. 2, 237–246.

A Pilot Study of a New Testing Method for Spatial Abilities Evaluation

Renata A. Górska[1], **Zuzana Juščáková**[2]

[1]*Faculty of Architecture, Div. of Descriptive Geometry and Engineering Graphics A-9, Cracow University of Technology, Warszawska st. 24, PL 31-155 Kraków, Poland*
email: rgorska@pk.edu.pl

[2]*Faculty of Civil Engineering, Department of Descriptive Geometry, Technical University of Košice, Vysokoškolska 4, SK 040 01 Košice, Slovakia*
email: Zuzana.Juscakova@tuke.sk

Abstract. A new study on spatial visualization abilities based on the *Space Imagination Test (TPP)* has been conducted in international cooperation between Poland and Slovakia. The goal of this study was not only to provide and analyze the levels of spatial skills of Polish and Slovak engineering students at technical universities but also to provide a validation of the test itself. The TPP-test has been developed in frame of the project VEGA[1], which was granted by the Slovak Ministry of Education in 2000/2002. The test results have been statistically evaluated and conclusions formulated.

Keywords: Spatial Visualization Ability, Graphics Education

MSC 2000: 51N05

1. Introduction

According to Thurstone [17] *spatial ability* is one of the factors important for general intelligence. Recent as well as historical factor analytic studies provide strong and consistent support for the existence of at least two distinct spatial abilities: *spatial visualization* and *spatial orientation*.

Mark McGee [10] defines spatial visualization as *"the ability to mentally manipulate, rotate, twist, or invert pictorially presented stimuli"*. McGee identifies five components of spatial skills:

spatial perception, spatial visualization, mental rotations,
mental relations, and *spatial orientation.*

[1]VEGA Grant Project No 1/7319/20

ISSN 1433-8157/$ 2.50

However, there is still no real consensus about what is meant by the term *"spatial visualization skills"* or *"spatial abilities"* [14]. Spatial ability can be defined as the "innate" ability to visualize spatial objects and their transformations in one's imagination.

It is not possible to measure a person's spatial ability directly. However, it can be measured indirectly by examination of the skill-levels of a person performing specific tasks. Therefore, in order to ascertain spatial abilities we use standardized tests to measure spatial skills and, from the tests' results we infere conclusions regarding a person's spatial abilities. Tartre [16], studied the earlier work of McGee [10], and proposed a classification for spatial skills based on the mental processes that are expected to be used in performing a given task. Two distinct categories of 3-D spatial skills — *spatial visualization* and *spatial orientation* have been distinguished. Spatial visualization ability is understood as to be able to perform mental rotations and mental transformations.

In the past years there were several international research studies about spatial visualization abilities from different viewpoints [3, 4, 6, 7, 12, 13, 14, 15]. Standardized tests, such as Mental Rotations Test [18] or Mental Cutting Test [1] have been used to evaluate the spatial visualization skill-levels of engineering students from technical universities. The overall interest in this field arises largely from the estimation of spatial visualization abilities as necessary and until now not well enough developed conditions for visual communications. In our daily lifes visual communication gets more and more important by the emergence of computer graphics and multimedia applications. Especially for those persons forming and designing the three-dimensional environment, working in the field of engineering, spatial visualization abilities are critical important.

In this paper we analyze and compare the results of a new testing instrument, which is a *"Spatial imagination test"* (TPP) developed under suspicion of the Slovak Ministry of Education. This new testing instrument consists of four parts. Specifically, its purpose is to examine the ability to operate with the figural matter in the three-dimensional space. The test had been administered at the Slovak University of Technology in Bratislava (SUT), Technical University of Košice (SvF TU) and Cracow University of Technology (CUT) respectively at the Faculties of Civil Engineering in the winter semester of 2001. The test results have been statistically analyzed.

This new test is still under development, and international comparison of the test results will give a basis for future validation of the test itself. The point of future research will be the evaluation of this new test and providing a comparison of the test results with the results of the tests that have been commonly used for spatial abilities evaluation (MCT, MRT, or others). The authors deeply believe that the test presentation will provide a new light into the research of spatial abilities.

2. Test of Space Imagination (TPP)

The "Spatial Imagination Test" has been developed in frame of a VEGA project, which was granted by the Slovak Ministry of Education in 2000/2002. Originally, the research on spatial skills with the use of TPP had been conducted primarily at Slovak technical universities. The test itself was verified by means of the use of so called content validity ratio (CVR) factor [8]. A group of experts was tested to evaluate CVR. For particular parts of the test CVR index was respectively equal to

0.05 for Subtest I, 0.71 for Subtest II, 0.90 for Subtest IV,

which is a good measure of a test quality. The content validity ratio for Subtest II and

Subtest IV shows a high level. Thus, having earlier experience in testing spatial abilities with the use of MRT, MCT and the other tests, the authors decided to perform a pilot study with the use of TPP test in Poland.

In consequence, the TPP test was also used as a measure of spatial abilities of engineering students at the Faculty of Civil Engineering in the winter semester of 2001. The test itself consists of four different Subtests I–IV. In the following the tests' problems are described.

2.1. Subtest I: Parallel or intersecting?

The test example of one item is shown in Fig. 1. In Subtest I the subjects are asked to determine whether the line p is parallel to or whether it intersects a specific plane ABC. All the solids used as testing models are regular, namely a tetrahedron, a cube or an octahedron. Line p has been determined by two various points, which can be chosen among a center of gravity of a solid's face, an orthocenter of a solid and a midpoint of an edge. Plane ABC can be determined by a solid face, a diagonal plane or optionally determined by three non-collinear points on the edges. The subtest consists of 12 items and solution time is 5 minutes.

This part of the test has been designed to measure the *knowledge of figural relations.*

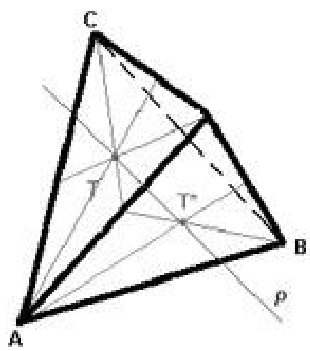


Figure 1: Parallel or intersecting?

2.2. Subtest II: The wire in the cube

This part of the test (Fig. 2) consists of problems related to testing spatial abilities based on a multiview projection. One piece of a wire has been bent inside the cube (along its faces or its diagonals) in such a way that it has neither knots nor "crossings". The test basically consists of two tasks. In II-1-a and II-1-b the subjects are asked to choose one out of five 3D models corresponding to the given three views (2 items). In the task II-2-a a 3D pictorial of a wire inside a cube has been provided as an isometric pictorial and the subjects are supposed to draw a front, a top and a side view related to the pictorial view (4 various items). In the task II-2-b based on the three views a subject is supposed to draw a pictorial view of a wire in a cube (4 items).

Especially in this problem mental rotation and visualization ability were required as the points of observation. They were assumed to be assigned in a special way, but not as it is traditionally determined in a multiview projection method.

Subtest II has been designed to check the practical ability to *transfer between 3D and 2D* or vice versa. This ability is substantial for engineering practice, especially for those who are involved in design practice. The time provided for solution is 20 minutes.

Figure 2: An example of a problem “The wire in the cube”

2.3. Subtest III: Solid’s complement

In this part of the test a cube has been cut into two parts, one of the parts is called a *cut*, the other its *complement* (Fig. 3). There have been 4 cuts denoted with letters A, B, C, and D and 16 complements shown as various isometric pictorials in various positions. The problem to solve was

a) to identify the shapes and volumes of the cuts and complements, and

b) to match corresponding cuts and complements in this way that it makes a cube.

In order to solve this task, mental rotation has been required. Thus the *mental manipulation and shape recognition* was subject to testing. Solving time for these 16 items is 5 minutes.

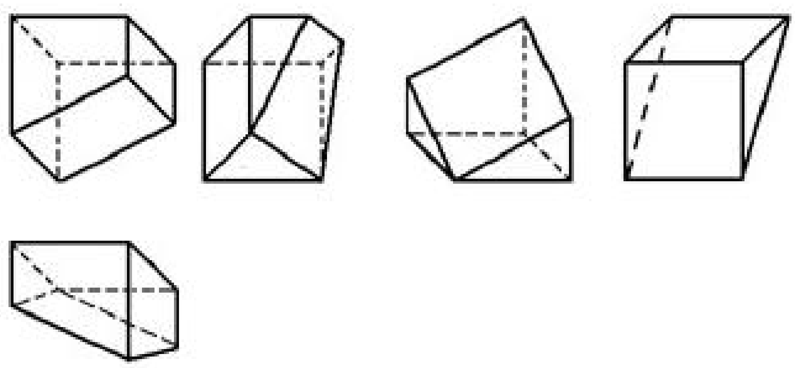


Figure 3: An example of a cut-out part: one of 16 items

2.4. Subtest IV: Cube’s complement

A cube consists of 2 parts (see Fig. 4), one of them has been drawn with a blue pencil in the problem pattern. The subject has to draw with a red pencil a complementary part of the cube. This subtest consists of 12 items. This problem required from the subjects not only to mentally visualize a missing part but also to present ability to draw correctly and to denote hidden lines. It has to be pointed out that missing parts of the cube had to be complemented in pictorials with changed viewpoint.

Thus the task required *flexibility and space orientation* and certainly a *mental and manual manipulation* from the subjects. Time provided for solving this part — 20 minutes. Within

Guilford's [5] classification of spatial abilities this part of the test falls into a group of a convergent product of figural systems.

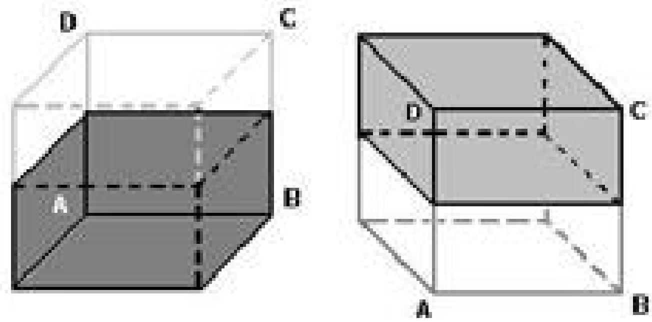


Figure 4: Cube's complement — problem example

3. International collaboration

Several previous research studies on spatial abilities [3, 4, 12, 14] have been conducted to measure a person's skill levels with regards to spatial visualization by using various testing instruments, such as Mental Rotations Test [18], Mental Cutting Test [1], or PSVT:R (Purdue Spatial Visualization Test: Rotations) [14]. The testing went in frame of international co-operation and the test results have been statistically analyzed and compared.

As Sorby [14] states there is still disagreement about exactly what constitutes spatial visualization skills. It is generally recognized that these skills are important for success in engineering and other technical careers.

A new testing instrument for assessing spatial skill levels has been developed by Z. Juščakova [6] and used as a testing instrument. In winter semester of 2001/02 a pilot study using the same testing material had been administered at the Faculties of Civil Engineering of the Cracow University of Technology (CUT), Slovak University of Technology in Bratislava (SUT) and at the Technical University of Košice (SvF TU). The data collected were split and elaborated in two groups: one set of data received in Poland (CUT), and the other set of data received in Slovakia and called STU (Slovak Technical Universities), which mean all the data collected at SUT and SvF TU together. The frequency distribution of the scores for Subtests I, II, III, and Subtest IV has been shown respectively in Fig. 5 for CUT and in Fig. 6 for STU. The X-axis in each chart shows ranges of scores to gain in particular sub-tests, Y-axis shows dependent value meaning the number of occurrence in percentage values. The numbers of subjects and mean scores for the tests together with the standard deviation values have been shown in Table 1.

4. Tests results analysis

Both at CUT and at STU the number of subjects who were tested with TPP was balanced and equal to 134, while 96 were male and 38 were female at CUT and respectively 95 men and 39 women at STU (Table 1). Mean scores for individual students were statistically analyzed using a two-tailed statistical T-test to determine the significance of means between the two universities for each part of Subtest I, II, III, IV, and for the TPP test as a whole. T-test

Table 1: Average scores (scores in percentages) of the Subtests I – IV in a group as a whole and by gender

	Cracow University of Technology (CUT)			
	Subtest I	*Subtest II*	*Subtest III*	*Subtest IV*
Group result	71.83 $s.d. = 15.18$ $(n = 134)$	50.04 $s.d. = 24.92$	93.80 $s.d. = 10.44$	65.05 $s.d. = 18.91$
Men	74.48 $s.d. = 14.65$ $(n = 96)$	54.91 $s.d. = 24.69$	94.40 $s.d. = 11.01$	67.22 $s.d. = 19.03$
Women	65.13 $s.d. = 14.61$ $(n = 38)$	37.72 $s.d. = 21.23$	92.27 $s.d. = 8.77$	59.58 $s.d. = 17.69$
	Slovak Technological Universities (STU)			
	Subtest I	*Subtest II*	*Subtest III*	*Subtest IV*
Group result	73.32 $s.d. = 17.81$ $(n = 134)$	60.22 $s.d. = 22.88$	93.47 $s.d. = 17.08$	45.29 $s.d. = 27.66$
Men	74.22 $s.d. = 17.73$ $(n = 95)$	63.25 $s.d. = 21.49$	94.92 $s.d. = 14.12$	45.52 $s.d. = 28.92$
Women	71.05 $s.d. = 18.04$ $(n = 39)$	52.58 $s.d. = 24.72$	89.80 $s.d. = 22.76$	44.74 $s.d. = 24.56$

Table 2: Significance of differences in mean test scores at CUT and STU (T-test results) and correlation between the universities

	Subtest I	*Subtest II*				*Subtest III*	*Subtest IV*	*TPP*
		II-1-ab	*II-2-a*	*II-2-b*	*Together*			
Pearson's coeff. $r =$	0.9667	0.9510	0.86.58	0.9348	0.9348	0.9026	0.9629	0.9711
F-test	0.0669	0.9912	0.5748	0.6854	0.0265	2.7E-08	1.5E-05	1.6E-07
T-test	0.4611	0.5052	0.0143	4.6E-07	2.4E-07	0.8504	7.4E-11	0.7231

Table 3: Correlation between the subtests for CUT and STU

Subtests	*I/II*	*I/III*	*I/IV*	*II/III*	*II/IV*	*III/IV*
CUT	0.9586	0.8138	0.9754	0.7232	0.9612	0.8269
STU	0.9786	0.7294	0.9504	0.7210	0.9350	0.5607

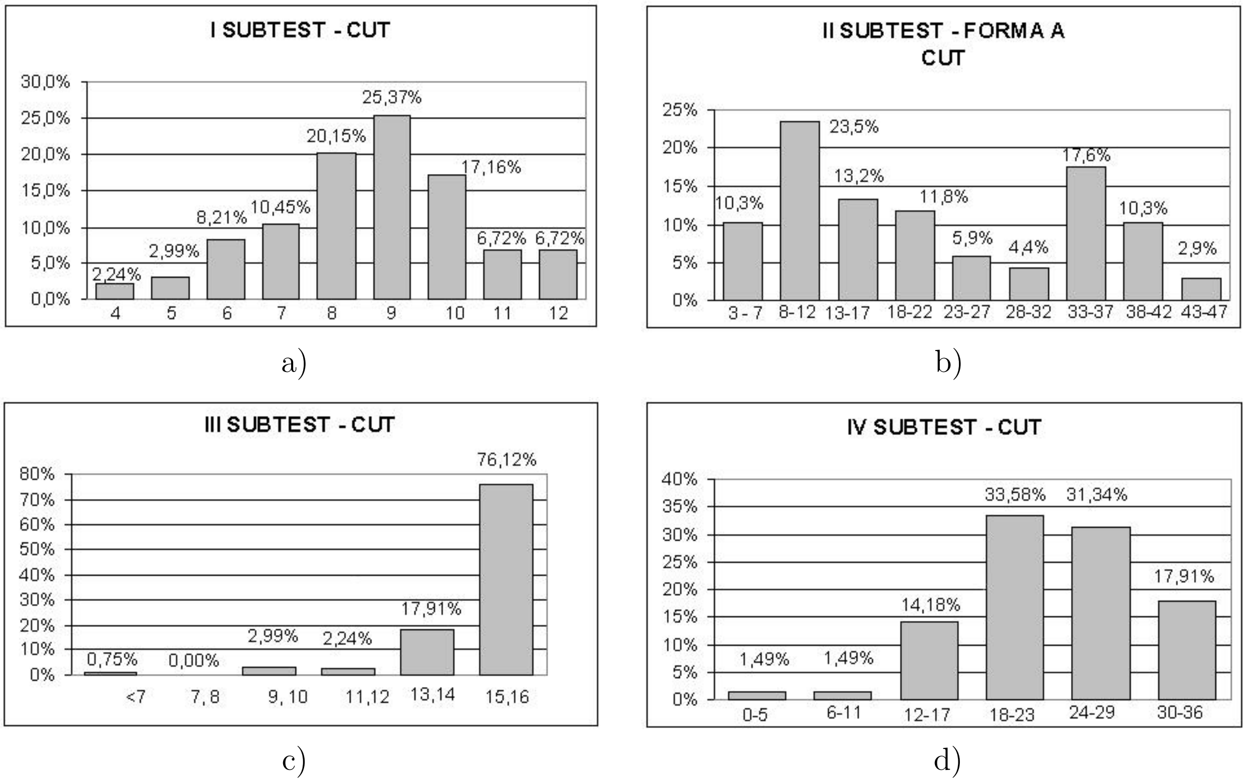


Figure 5: Frequency distribution in percentage values related to the score ranges at CUT: a) Subtest I, b) Subtest – Form A, c) Subtest III, d) Subtest IV

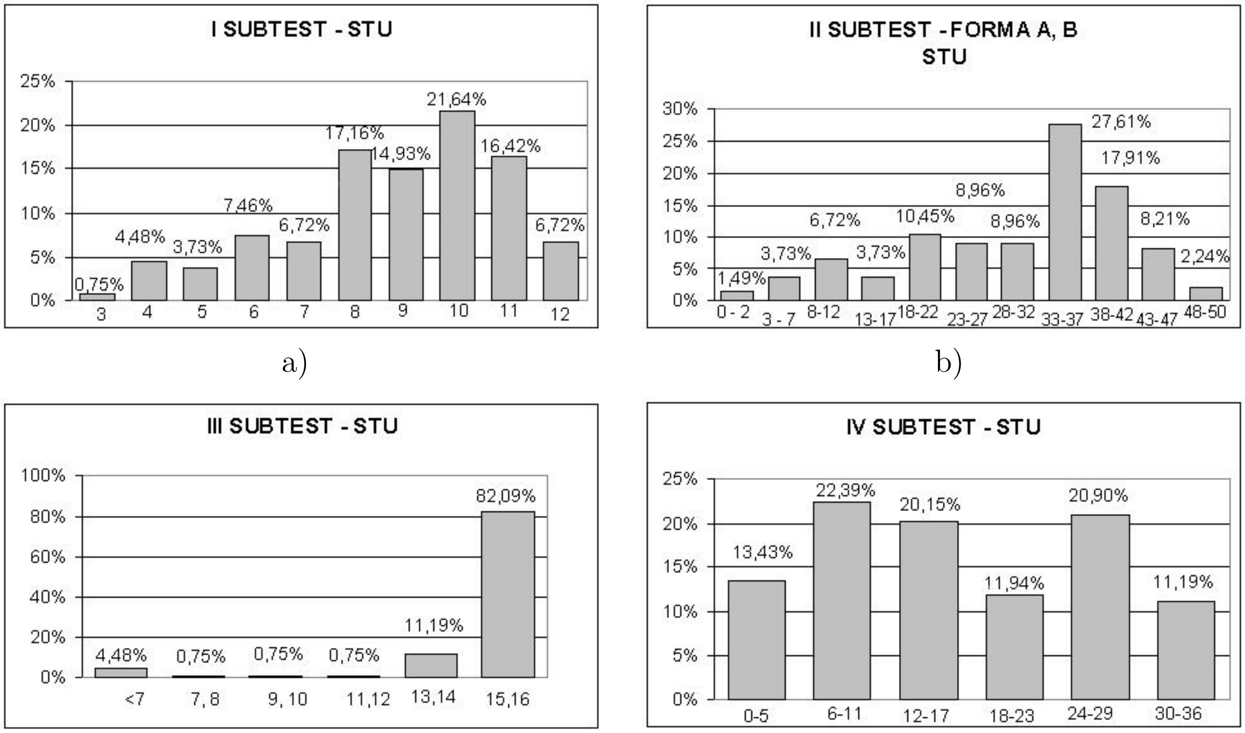
values of a hypothesis concerning no difference in a mean score on each subtest at CUT and STU have been included in Table 2.

The mean score for the group as a whole on Subtest I and Subtest III was on the same level at both universities and there was no significant difference between the results. In student's T-test we obtained respectively $p = 0.4611$ for Subtest I and $p = 0.8504$ for Subtest III. However, the mean score of Subtest II at CUT was obviously lower than that of STU (Table 1: $50.04 < 60.22$, Table 2: $p = 2.4E - 07$).

Similar, but the other way round observation can be made regarding Subtest IV, where the result at CUT for the group as a whole was evidently higher than at STU (Table 1: $65.05 > 45.29$, Table 2: $p = 7.4E - 11$). The last result may be explained by the fact that at STU the test was administered at the beginning of the winter semester 2001/02, while at CUT administration took place at the end of the same semester. On the other hand Subtest II proved to show some deficiencies and needs to be re-modeled.

The test TPP as a whole did not show statistically significant difference of mean results between STU and CUT ($p = 0.7231$). Regarding particular parts of the Subtest II we used a statistical F-test to investigate if there is no difference between the variances of mean scores between the universities. Only for one part of Subtest II, namely for a problem II-1-ab value $p > 0.05$ while for the other parts of this subtest $p < 0.05$, and for the subtest as a whole is $p = 2.4E - 07$ (Table 2). Thus, we may conclude that there is significant difference in variances of mean scores.

T-test has been also used to compare mean scores in a group by gender for each university. At CUT significant gender differences showed up Subtest I ($p = 0.0014$), Subtest II ($p =$



a) b) c) d)

Figure 6: Frequency distribution in percentage values related to the score ranges at STU: a) Subtest I, b) Subtest – Form A, c) Subtest III, d) Subtest IV

0.0067) and Subtest IV ($p = 0.0306$). Subtest III ($p = 0.2431$) showed no significant gender differences. At STU no significant gender differences showed up Subtest I ($p = 0.3250$) and on Subtest IV ($p = 0.7610$), while on Subtest II ($p = 0.0156$) and Subtest III ($p = 0.0201$) significant gender difference occurred.

The TPP test has been still under reconstruction and evaluation. However, when we compare adjacent figures for CUT (Fig. 5) and STU (Fig. 6) similarity of frequency distribution is evident for Subtest I and Subtest III. The most representative results were obtained in Subtest II and Subtest IV. The test was also analyzed on its reliability. In Table 2 a Pearson product moment coefficient of correlation r between CUT and STU has been inserted separately for each subtest and for the group mean score. It has been shown that apparent linear correlation of group mean scores in each subtest exists between the two universities (Table 2).

The data presented has indicated that there is a ‘ceiling effect’ for Subtest III and this subtest will not have much discriminatory power. The measure of linear correlation has been used to investigate if the subtests are independent. The results are presented in Table 3 and they prove strong linear relationship between the results. There is high correlation between the two of each subtests at each university. However, the correlation coefficient is evidently lower in all cases where Subtest III is compared to the other tests.

5. Conclusions

Two aspects of testing were taken into consideration while administering the TPP test. Regarding the test elaboration and validation it has been shown that not all subtests have a normal (Gauss) frequency of distribution and that by mean scores there is no significant difference in both testing groups at CUT and STU for the test as a whole TPP. We came to the same conclusion when analyzing Subtest I and Subtest III and part II-1-ab of Subtest II. Normal distribution have Subtest I for both groups reliable, Subtest II for STU and Subtest IV for CUT approximately. As it has been investigated and graphically interpreted in Fig. 5 and Fig. 6, the Subtest-III's results together with statistical evaluation indicated that there was a 'ceiling effect' for subtest III and will not have much discriminatory power. It has been suggested that this part of the test should be re-modeled or dropped. Continued research will provide analysis of the test validation. Up until now, the CVR factor has been the only measure used for the test evaluation.

The results of TPP indicated significant difference between the female students and the male students. This conclusion has been identified in many previous research [3, 4, 7, 12, 13, 14]. A new tool of TPP has been used in research on spatial visualization abilities at Polish and Slovak universities in a pilot study. However, there still remain the questions, which are open for the future research. The most important question is if the test that was developed could be given to the general population. This problem will undergo further research.

References

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

[2] J. Ferjenčik: *Uvod do Metodologie Psychologického Výskumu*. Portál, Praha 2000, 207–210.

[3] R. Górska, S.A. Sorby, C. Leopold: *Gender Differences in Visualization Skills – An International Perspective*. The Engineering Design Graphics Journal **62**, no. 3, 9–18 (1998).

[4] R. Górska, S.A. Sorby, C. Leopold: *Cross-cultural Comparisons of Gender Differences in the Spatial Skills of Engineering Students*. Journal of Women and Minorities in Science and Engineering **5**, no. 3, 279–291 (1999).

[5] J.P. Guilford: *Psychometric Methods*. McGraw-Hill, New York 1954.

[6] Z. Juščáková: *Mentálne reprezentácie priestorovej predstavivosti*. Proc. 21st KoGaPG, Lednice v Moravě 2001, 48–51.

[7] Z. Juščáková, Z. Zvariková: *Inter-sexuálne rozdiely v priestorovej predstavivosti*. Proc. 20th KoGaPG, Mechanical Engineering Faculty, ČVUT Praha 2000, 105–110.

[8] Z. Juščáková: *Obsahová a konštruktová validita TPP*. Proc. Symposium on Computational Geometry SCG'2001, Faculty of Mechanical Engineering, Civil Engineering Faculty, STU in Bratislava, Kočovce 2001, 59–63.

[9] P.H. Maier: *Räumliches Vorstellungsvermögen*. Lang, Frankfurt a.M., Berlin, Bern, New York, Paris, Wien 1994.

[10] M.G. McGee: *Human Spatial Abilities: Psychometric Studies and Environmental, Genetic, Hormonal, and Neurological influences*. Psychological Bulletin **86**, 899–918 (19??).

[11] M. Smith: *Spatial Ability – Its Educational and Social Significance*. University of London, London 1964.

[12] S.A. Sorby, R. Górska: *The Effect of Various Courses and Teaching Methods on the Improvement of Spatial Ability*. Proc. 8th ICECGDG, Austin 1998, 252–256.

[13] S.A. Sorby: *Developing 3-D Spatial Visualization Skills*. Engineering Design Graphics Journal, Spring 1999, 21–31.

[14] S.A. Sorby: *Assessing and Improving Spatial Visualization Skills of Engineering Students: International Collaborations and Studies*. Proc. IVth International Conference on Graphics Engineering for Arts and Design, Sao Paolo, Brasil, Nov. 2001, 1285–1313.

[15] K. Shiina, K. Suzuki, E. Tsutsumi: *Evaluation of Students Spatial Abilities by a Mental Rotations Test*. Proc. 5th ICECGDG, Melbourne 1992, 286–290.

[16] L.A. Tartre: *Spatial Skills, Gender, and Mathematics*. In E.H. Fennema & G.C. Leder (eds.): *Mathematics and Gender*, Teachers College Press, New York 27–59.

[17] L.L. Thurstone: *Primary Mental Abilities*. Univ. of Chicago Press, Chicago 1938.

[18] S.G. Vanderberg, A.R. Kuse: *Mental Rotations, a Group Test of Three Dimensional Spatial Visualization*. Perceptual and Motor Skills **47**, 599–604 (1978).

Received August 1, 2002; final form December 16, 2003