Journal for Geometry and Graphics
Volume 8 (2004), No. 1, 107–115.

# Training Visualization Ability by Technical Drawing

**Gerardo Prieto Adanez**[1], **Angela Dias Velasco**[2]

[1] *Psychology Faculty, Salamanca University*
*Avda. de la Merced, 109–131, 37005 Salamanca, Spain*
*email: gprieto@usal.es*

[2] *Mechanics Dept., Engineering Faculty, Paulista State University*
*Av. Ariberto Pereira da Cunha, 333, Guaratinquetá – SP – Brasil 12516-410*
*email: avelasco@feg.unesp.br*

**Abstract.** To analyze if learning Technical Drawing improves the spatial visualization aptitude, a visualization test was applied, at the beginning and at the end of a Technical Drawing course, on a sample of first year engineering students. At the end of the Technical Drawing course, it was observed that more than one third of the students increased their performance on the visualization test. This improvement was statically significant and equal both for men and women.
These results support the conclusion that spatial visualization is an aptitude that could be improved with training and, although teachers do not explicit this objective, Technical Drawing courses are an efficient way of doing this. It can be suggested that the change in spatial visualization aptitude may be considered as a efficiency indicator of the teaching-learning process.

*Key Words:* Spatial aptitude, Technical Drawing, aptitude training

*MSC 2000:* 51N05

## 1. Introduction

*Spatial ability* may be defined as the ability to generate, retain, retrieve, and transform well-structured visual images [6]. There are several spatial abilities. The core of *visualization*, the most complex of them, is the process of transforming visual patterns. The ability to transform the visual images is an important skill for architects, engineers, and those in the drawing field. Hsi et al. [4] interviewed a group of engineering instructors and engineers working in industry. They agree that spatial reasoning skills were important and could help engineering students. Moreover, it has been empirically demonstrated that visualization is a moderately good predictor of success in courses such as Engineering Drawing [4, 6, 10, 11, 14].

ISSN 1433-8157/$ 2.50 

Because of this influence on academic and professional success, the training of spatial abilities has elicited a great interest.

Baeninger and Newcombe [1] published several meta-analysis of spatial ability training studies. For the studies included in the review, the duration of instruction ranges from a very brief session to a year-long program. The content of instruction ranges from very specific training such as repeated exposure to a spatial test to an indirect training such as an entire college engineering course. They concluded that training of a variety of types and duration increases the spatial test performance, and that training does not benefit the sexes differentially. In all cases, the magnitude of the improvement between the first and the last administration of a spatial test was significantly different from zero, and at least moderate in size. Thus, levels of spatial ability appear as environmentally malleable. According to Newcombe et al. [8] this conclusion could be generalized for a variety of spatial measures such as mental rotation tests, dynamic spatial tests, and horizontality-verticality tasks. However, because of the scarce number of studies, the effect of the indirect and long training is not sufficiently known, mainly its influence on the visualization ability.

The purpose of the study presented here is to examine if activities involved in an engineering course of Technical Drawing could improve the performance on a visualization ability measure.

From our point of view, this objective is valuable for the following arguments:

1. The improvement of visualization ability of engineering students would be an important benefit because of its influence on job and academic success.
2. The effect on visualization ability of an indirect, sizable, and long training will be more permanent and general than that of a specific-task training. Technical Drawing could be one of the more efficient training methods.
3. To contest the traditional thinking that this aptitude is a gift, making the teachers of Technical Drawing free of responsibilities for its development.

## 2. Method

### Participants

In this study 159 first year Brazilian engineering students participated. 57 were from the Paulista State University — Guaratingueta Campus (UNESP), 24 from the Polytechnic School of São Paulo University (EPUSP) and 78 from the Chemistry Engineering Faculty of Lorena (FAENQUIL). The mean age was 18 years and 9 months and the standard deviation was 2 years and 11 months. 31,45% of the sample were women.

### Instrument

We applied the instrument TVZ2002 form B, a paper and pencil test with 20 items, selected from an Item Visualization Bank constructed and founded on Cognitive Theory [9, 10]. The time limit was 20 minutes. The task consists of a cube that has all faces identified with letters. To the right, the cube is shown unfolded with one of its faces identified and another face marked with an interrogation *'?'*. The student has to identify the letter of the face with the interrogation and its relative position. The student has to choose the right answer among 9 given options (Fig. 1). Studies with similar versions of this test show that it is an excellent indicator of the visualization construct [9, 10].

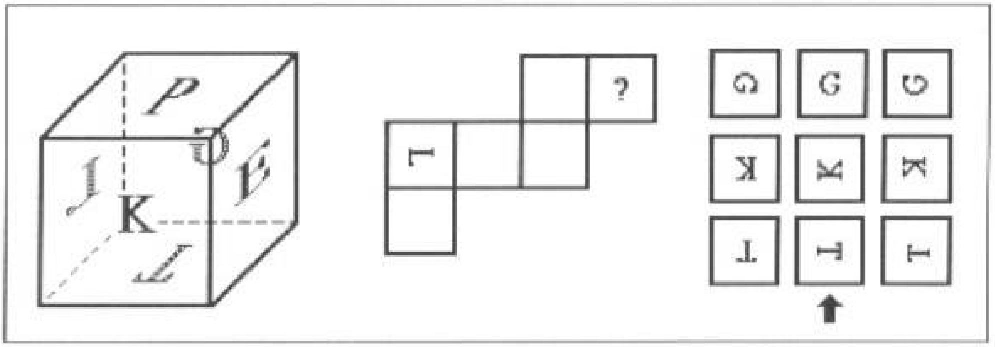

Figure 1: One task of TVZ2002. The right answer is the one marked with the arrow

## Procedure

The test was applied to each participant at the beginning and at the end of the Technical Drawing course. At the first application of the test in March 2002, there were 159 students. 101 of them had participated in the second application in November 2002.

The Technical Drawing course in the Engineering Faculties in this work has items characteristic for the graphic area. This include making and interpreting orthographic and pictorial views, imagining sections and details of objects, knowing the norms of technical graphics expression, dimensioning and being able to efficiently manipulate traditional and computerized instruments of drawing. Although there are some differences among these faculties, synthetically this was the basic content of the course. The activities in these courses are predominantly individual and based on practice, as well as the evaluations.

## Measurement Model

The item response matrix was analyzed with the RASCH *Model (RM)* [12]. As we have recently summarized in this journal [11], RM provides measurement equivalent to that available in the physical sciences, because it synthesizes the key features of requirements for fundamental measurement (invariance, unidimensionality and additivity) [16]. This model is known as a one parameter logistic model because the probability of a correct answer, $P(X_{ij} = 1)$ depends on the difference between the ability of the examinee ($\theta_j$) and the parameter of the difficulty of the item ($\beta_i$): the greater the difference $\theta_j - \beta_i$, the greater the probability that the examinee answers the item correctly. The equation (1) describes the relation between both values.

$$P(X_{ij} = 1) = \exp(\theta_j - \beta_i)/\left(1 + \exp(\theta_j - \beta_i)\right) \quad (1)$$

The meter to score jointly persons and items (usually the *logit* scale) has interval properties. The location of point zero on the scale is arbitrary. In the RASCH tradition, it is situated in the mean of the difficulty of items. In this case, the interpretation of the persons' parameters is simple (values of ($\theta_s$) greater than 0 mean that these persons have a probability of success greater than 0,50 on items that have a mean difficulty). Although the logit scale can adopt values between $\pm\infty$, in the majority of the cases it is situated between $\pm 5$. The interval property has a great importance because it guarantees the invariance of differences of scores all through the continuous scale (a necessary requirement in the analysis of the changes due to development or training).

## 3. Results

### Fit of the data to the Rasch Model

First, we show the results of the fit model analysis for items and examinees. The fit is crucial. If it does not exist, the values does not have a theoretical meaning and the RM advantages, such as interval property, disappear. The most employed statistic is called *infit*, which is an information-weighted sum. The statistical information in a RASCH observation is its variance. To calculate infit, each squared standardized residual value is weighted by its variance and then summed. Infit statistic is reported as mean squares divided by their degrees of freedom, so that they have a ratio scale form with an expected value of 1 and a range from 0 to positive infinity. In this form, the mean squared fit statistics are used to monitor the compatibility of the data with the model. Traditionally it is considered that values greater than 1,3 show maladjustment [2]. The descriptive statistics of infit values are shown in Table 1.

Table 1: Item and person fit. Descriptive statistics of infit values

| *Object* | *Mean* | *St. D.* | *Maximum infit* | *Percentage with infit* $> 1.3$ |
|---|---|---|---|---|
| *Pretest Items* | 0.98 | 0.15 | 1.27 | 0 |
| *Pretest Persons* | 0.97 | 0.37 | 2.10 | 18.92 |
| *Posttest Items* | 0.97 | 0.12 | 1.20 | 0 |
| *Posttest Persons* | 0.98 | 0.34 | 2.22 | 16.30 |

The statistics show a good fit. On the one hand, the means and standard deviations of values are those that are expected when there are not substantial divergences between the model predictions and the raw data. On the other hand, non of the items present values greater than 1.3 . Only a moderate subsample of the students shows this result.

### Establishing stable item calibrations to measure changes

Second, we analyze whether the items are stable across the two measurement occasions. The measurement of a meaningful progress and development along a learning process requires a stable definition of the measurement variable [7]. Comparison requires a stable frame of reference. If the item calibrations are invariant, then differences between person measures at the two occasions are valid indicators of changes in persons over time [18, 19, 20].

The stability of item parameters that are obtained on two occasions ($\beta_1 - \beta_2$) is evaluated by inspecting the graphical relationship of the item parameters for pretest and posttest, and by examining the standardized difference (2) between the two estimates [21].

$$z = \beta_1 - \beta_2 / [SE^2_{(\beta_1)} + SE^2_{(\beta_2)}]^{1/2}. \tag{2}$$

$SE_{(\beta_1)}$ and $SE_{(\beta_2)}$ are standard errors of the estimates. The standardized differences for an item pool that conform to stability have an expected value of 0.00 and an expected standard

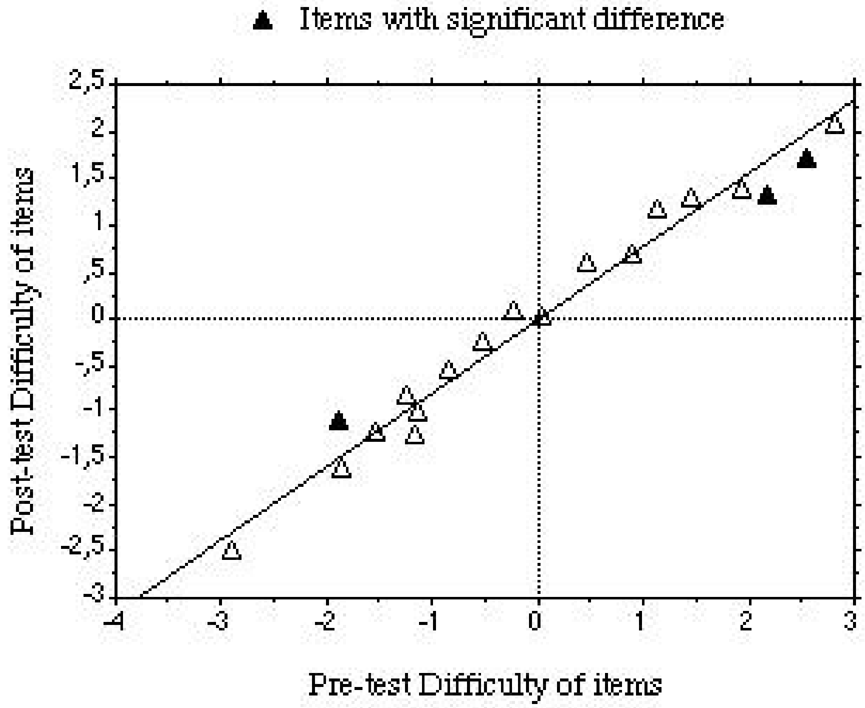


Figure 2: Scatter plot of item difficulty in pretest and posttest

deviation of 1. Large deviations from these expectations in the observed data indicate change over time [18].

A scatter plot of the item parameters for pretest and posttest is shown in Fig. 2.

The analysis shows that there are minor differences between pretest and posttest calibrations. Only three items (3, 16 and 19) exhibit significant standardized differences ($z \geq 1.96$; $p \leq 0.05$), but the size of these differences is low and without theoretical and practical consequences. Since such differences between the items may confound the inferences about changes in the person measures for Time 1 and Time 2, we estimate the abilities of persons in the posttest using pretest parameters of stable items across time as anchor [7].

### Measuring Pretest-Posttest Change

The traditional method of assessing change using a $t$-test demonstrated statistically significant gains for males and females from pretest to posttest (Table 2). The effect sizes, explained as Cohen's $d$, range between 0.77 (females) and 0.88 (males), which are large effect sizes according to Cohen's taxonomy [13]. The difference between change means of males and of females was not statistically significant. However, males outperform females both in pretest ($t = 3.74$, $p = 0.0003$, $d = 0.60$) and posttest measures ($t = 3.52$, $p = 0.0007$, $d = 0.71$).

According with Smith et al. [15], rather than concentrating on group differences, it would be of greater value to see which individuals demonstrated statistically significant gains or losses. Rasch measurement also produces a distinct advantage over classical test theory: standard errors for each individual measure [11]. Standard errors allow for the statistical comparisons of pre-post scores at the individual rather than group level. Using this information one is able to target individuals that demonstrated statistically significant gains in visualization.

The change of person parameters that are obtained on the two occasions ($\theta_2 - \theta_1$) is evaluated by examining the standardized difference (3) between the two estimates [18].

$$z = \theta_2 - \theta_1 / [SE^2_{(\theta_2)} + SE^2_{(\theta_1)}]^{1/2} \tag{3}$$

The standardized differences have an expected value of 0.00 and an expected standard deviation of 1. Large deviations in observed data from these expectations indicate a change over time [18]. In order to test the nullity hypothesis, we perform a two tail test ($\alpha = 0,05$).

The results of this procedure are summarized in Table 3. The data shows that the program benefited a large number of persons, that have similar size in males and females.

Table 2: Results of post-pre analysis at the group level

| TVZ2002 | *Mean* | *SD* | *Mean post-pre* | *SD post-pre* | *t* | *df* | *p* | *Effect size(d)* |
|---|---|---|---|---|---|---|---|---|
| *Pretests (males = 109)* | -.68 | 2.05 | | | | | | |
| | | | 1.32 | 1.49 | 7.07 | 63 | .0001 | .88 |
| *Posttests (males = 64)* | -.46 | 2.33 | | | | | | |
| *Pretests (females = 50)* | -1.93 | 1.74 | | | | | | |
| | | | 1.18 | 1.54 | 4.65 | 36 | .0001 | .77 |
| *Posttests (females = 37)* | -1.17 | 2.08 | | | | | | |

Table 3: Number of persons in categories of standardized change scores

| | *Reduction* ($z \leq -1.96$) | *Stability* ($-1.96 < z < 1.96$) | *gain* ($z \geq 1.96$) | *Totals* |
|---|---|---|---|---|
| *Female* | 0 (0 %) | 23 (62.16 %) | 14 (37.84 %) | 37 (100 %) |
| *Male* | 0 (0 %) | 40 (62.50 %) | 24 (37.50 %) | 64 (100 %) |

The gain was not correlated with the visualization level. On one hand, the correlation between change scores and pretest in the visualization test was not statistically significant ($r = -0.14$, $p = 0.1522$). On the other hand, the percentage of persons with significant gain was similar between persons who tested above the pretest-median and those who tested below the pretest-median.

## 4. Discussion and conclusion

The aim of this work was to analyze if the learning of Technical Drawing improves the spatial visualization aptitude. A visualization test was applied at the beginning and at the end of the Technical Drawing course on a sample of first year students of engineering. As a quantitative indicator of change we used the difference of the scores obtained at the beginning and at the end of the course. To acquire the same meaning of the change scores on the range of the variable, it is necessary to use a procedure with the following characteristics:

1. Additivity: measure with an interval property.

2. Invariance of the measure frame: scaling based on the invariant items on both applications (pretest and posttest).

In order to estimate the examinees scores on the spatial visualization aptitude with the additivity property, the Rasch model [12] was used. On both pretest and posttest, the data showed a good fit to the model. Therefore, it can be concluded that the scores had the required property.

To analyze the invariance of the measure frame on both pretest and posttest, the standardized differences between the items parameters on the first and second application was estimated. Only 3 items among the 20 of the TVZ2002-b showed significant differences. However the size of these differences was low and theoretically irrelevant. To situate the examinees' scores of pre- and posttests on a common scale, the posttest scores were estimated using as anchor the parameters of the pretest of the invariant items. Finally, the change scores for each person were calculated.

The mean on the TVZ2002-b was greater for males than for females, on both pretest ($t = 3,74$; $df = 157$; $d = 0,60$) and posttest ($t = 3,52$; $df = 99$; $d = 0,71$). This difference, favouring males, has a medium effect size, according to Cohen's classification criteria [3]. This effect size is greater than those obtained in other studies with visualization tests [17]. Probably the magnitude of the difference favouring males is due to the presence of strong demands of mental rotation in items of TVZ2002-b. Generally, males outperform females in mental rotation tasks. This is one of the most consistent data of psychometrical research about sex differences in spatial cognition [17].

At the end of the Technical Drawing course, more than a third part of the students has a statistically significant improvement on their performance in the visualization test. The improvement was high, both on the male group ($d = 0,88$), and on the female group ($d = 0,77$). Due to the long time lapse between the pretest and posttest, the improvement can not be interpreted as a practice effect. According to Kaufman [5], long intervals allow forgetting of the test's content, and therefore reduce the magnitude of the practice effects.

Converging with others studies, the improvement was similar on both males and females [8].

The results support the conclusions that spatial visualization is an aptitude that could be improved with training and that Technical Drawing courses are an efficient way of achieving this. Although the duration and generality of this kind of training are topics to future investigations, it could be suggested that they will be greater than those obtained by specific procedures related with spatial tests.

It is important for teachers of the graphic area to be conscious of the role of spatial visualization aptitude on their didactical attitudes. As we have recently shown in this journal [11], the spatial visualization aptitude is not only determinant to the students professional success but is directly related to the performance on Technical Drawing courses. To care about the spatial visualization is to care about the improvement of the teaching-learning process. From another point of view, because of the importance of visualization aptitude on the performance in Technical Drawing, changing this aptitude could be considered as an indicator of the efficiency of the teaching-learning process.

## Acknowledgements

- To the Engineering Faculty of UNESP, Guaratingueta Campus, the Polytechnic School of the University of São Paulo and the Chemistry Engineering Faculty of Lorena (FAENQUIL).
- To the students who did the tests.

## References

[1] M. Baenninger, N. Newcombe: *The Role of Experience in Spatial Test Performance: A Meta-Analysis*. Sex Roles **20**, 327–344 (1989).

[2] T.G. Bond, C.M. Fox: *Applying the Rasch Model: Fundamental measurement in the Human Sciences*. LEA, Mahwah/NJ 2001.

[3] J. Cohen: *Statistical power Analysis for the Behavioral Sciences*. Academic Press, New York 1977.

[4] S. Hsi, M.C. Linn, J.E. Bell: *The Role of Spatial Reasoning in Engineering and Design of Spatial Instruction*. Journal of Engineering Education **86**, 151–158 (1997).

[5] A.S. Kaufman: *Practice Effects*. In R. Sternberg:*Encyclopedia of Human Intelligence*, McMillan, New York 1994, pp. 828–833.

[6] D. F. Lohman: *Spatial Ability*. In R. Sternberg:*Encyclopedia of Human Intelligence*, McMillan, New York 1994, pp. 1000–1007.

[7] L. H. Ludlow, P. L. Andres, S. M. Haley: *Measuring Change: Functional Activity Item Calibrations at Admission vs. Discharge*. American Journal of Physical Medicine and Rehabilitation, in press.

[8] N.S. Newcombe, L. Mathason, M. Terlecki: *Maximization of Spatial Competence: More Important then Finding the Cause of Sex Differences*. In A. Mcgillicuddy-de Lisi, R. de Lisi: *Biology, Society, and Behavior: The Development of sex differences in cognition*. Ablex, Wesport/CT 2002, pp. 183–206.

[9] G. Prieto, A. R. Delgado: *Diseño Cognitivo de un banco de items de Visualización Espacial*. Metodología de las Ciencias del Comportamiento, vol. especial, 452–455 (2002).

[10] G. Prieto, A.D. Velasco: *Construção de um teste de Visualização a partir da Psicologia Cognitiva*. Avaliação Psicológica **1**, Porto Alegre / Brasil 2002, 39–47.

[11] G. Prieto, A. D. Velasco: *Predicting Academic Success of Engineering Students in Technical Drawing from Visualization Test Scores*. J. Geometry Graphics **6**, 99–109 (2002).

[12] G. Rasch: *Probabilistic models for Some Intelligence and Attainment Tests*. Danish Institute for Educational Research, Copenhagen 1960.

[13] R. Rosenthal, R.L. Rosnow: *Essentials of Behavioral Research: Methods and Data Analysis*. McGraw-Hill, New York 1984.

[14] T. Saito, K. Suzuki, T. Jingu: *Relations between Spatial Ability evaluated by a mental Cutting Test and Engineering Graphics Education*. Proc. 8th ICECGDG Austin/Texas 1998, pp. 231–235.

[15] E.V. Smith Jr., K.A. Lawless, L. Curda, S. Curda: *Measuring Change in Efficacy*. Popular Measurement **2**, 31–33 (1999).

[16] I. Styles: *The study of intelligence-the interplay between theory and measurement*. In M. Anderson: *The Development of Intelligence*, Psychology Press, Hove 1999, pp. 19–42.

[17] D. Voyer, S. Voyer, P. Bryden: *Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables*. Psychological Bulletin **117**, 250–270 (1995).

[18] E.W. Wolfe, C.W.T. Chiu: *Measuring Pretest-Posttest Change with a Rasch Rating Scale Model*. Journal of Outcome Measurement **3**, 134–161 (1995).

[19] B. Wright: *Comparisons Require stability*. Rasch Measurement Transactions **10**, 506 (1996).

[20] B. Wright: *Time 1 to Time 2 Comparison*. Rasch Measurement Transactions **10**, 478–479 (1996).

[21] B. Wright, G.N. Masters: *Rating scale analysis*. MESA Press, Chicago 1982.

Received June 9, 2003; final form June 7, 2004