Training of Spatial Visualization Using Computer Exercises

Gerardo Prieto[1], **Angela D. Velasco**[2], **Rosario Arias-Barahona**[3],
Mercedes Anido[4], **Ana-María Núñez**[3], **Patricia Có**[4]

[1]*Departamento de Psicología Básica, Psicobiología y Metodología, Universidad de Salamanca*
Avda. de la Merced 109–131, 37005 Salamanca, Spain
emails: gprieto@usal.es

[2]*Universidade Estadual Paulista (Brazil)*

[3]*Universidad de San Martin de Porres (Peru)*

[4]*Universidad Nacional de Rosario (Argentine)*

Abstract. This work describes the construction, application and results obtained with a battery of computerized exercises for training Spatial Visualization in engineering students. The battery contains four exercises based on very common tasks in the basic teaching of Technical Drawing. Each exercise is composed of 18 items with four options to answer from, only one of which is correct. After answering each item, the student receives immediate feedback of whether the answer is right or wrong. The response format of the exercises is called "Answer-Until-Correct", because if the answer the student gives is incorrect, he or she must keep on seeking the right answer. When the test is over, the student receives information on the score obtained. To assess the effectiveness of the training in Spatial Visualization, tests in this ability, scaled in the same metrics, were administered at the beginning and end of a Technical Drawing course. The figures of the exercises and tests were constructed with AutoCad and programming was carried out with Revolution Studio 2. Two models were used to obtain the measures: the Partial Credit Model [17] and the Rasch Model [24]. It was observed that, on average, the students' improvement in Spatial Visualization was moderate.

Key Words: Visualization ability, Technical Drawing, feedback, training, Answer-Until-Correct, Rasch model

MSC 2000: 51N05

1. Introduction

All the most important factor analyses with psychometrical figural tests [4, 13, 18, 20] concluded that *Visualization*, defined as the ability to manipulate, rotate, twist or invert images of objects, is the most important dimension in *Spatial Ability*. The susceptibility of Spatial Ability to training is one of the topics included in research programs. Training for Spatial Ability is of interest owing to its involvement in academic and work success: improving abilities would benefit learning and performance in the technical professions, such as architecture, engineering, and graphic and mechanical design [21]. In this sense, BAENNINGER and NEW-COMBE [3] published a meta-analysis of studies on the trainability of spatial abilities. The studies analyzed were classified into several categories as a function of the training contents (specific, general or indirect) and duration (long, medium, or short). The term *specific* refers to the preparation of a single spatial test and the term *general* refers to training in a variety of spatial tests. *Indirect* training includes activities associated with spatial ability, but is not concerned with preparing for specific spatial tests. As far as the duration is concerned, less than three weeks training is considered *short*, more than three weeks and less than a semester is considered *medium*, and a semester or more is considered *long*. Long-term training usually coincides with formal instruction in curricular subjects such as engineering or architecture.

Certain important conclusions can be drawn from the studies cited above. In the first place, it seems clear that training increases the level of execution in spatial tests. In second place, it is evident that most of the benefits were obtained from instruction in tasks very similar to those of the spatial tests used to measure the change. In our judgment, in order to avoid merely artificial effects and reach more generalizable conclusions, indirect training in tasks different from those of the measuring instruments is of greater interest. Finally, there is a scarcity of studies about long-term training and its effects on certain important spatial abilities, such as Visualization, a recognized mediator in academic and work performance in the technical professions. Thus, the objective of this work is to assess the effect on Visualization of long term indirect training consisting of a semester course in Technical Drawing aided by computerized exercises.

We describe the characteristics of a battery of computerized exercises constructed to indirectly train Visualization through tasks that are common in a Technical Drawing course, and that were introduced during a semester course in this subject addressed to first year engineering students. We then analyze the improvement in Spatial Visualization after the students finished the course in Technical Drawing.

2. The exercises

The battery was constructed with the program *Revolution Studio 2*, from *Runtime Revolution Limited* (https://secure.runrev.com) and the three-dimensional figures with AutoCad. At present, it is composed of four exercises based on common tasks in the basic teaching of Technical Drawing. All the representations of the objects in perspective are in isometrics. The four exercises are introduced with a brief summary of the contents of the task, illustrated with short texts and animated graphics. Each exercise is composed of 18 items with four possible answers, only one of which is correct. After answering each item, the student receives immediate feedback of whether the answer is right or wrong. He or she also receives an animated demonstration of what is right or wrong about it. Feedback is also given in the case of the right answer so that the student can see the solution in case he or she had simply
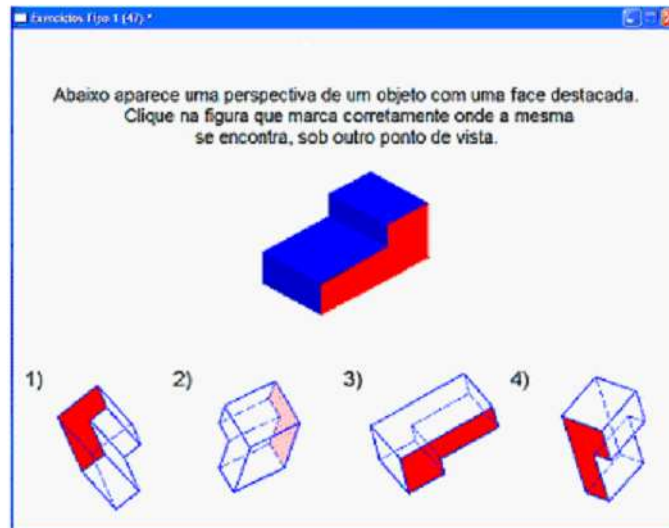
Figure 1: Example of an item in Exercise 1. The item reads:
Below you can find an object with a highlighted face. Click on the option
corresponding to the figure above from another point of view.

guessed it. This format is called *"Answer-Until-Correct"*, since if the answer is incorrect the student must look for another solution [26]. The program stores the answers to the 18 items and calculates a score that reflects the student's competence in the exercise. When the test is over, the student receives information on the score obtained.

The first exercise is based on a mental rotation task of three-dimensional figures. The student has to recognize the position of an edge or a face of an object rotated in space (see Fig. 1).

For the items of the second exercise, the student is asked to choose the option in which the main orthogonal views of an object appear correctly (see Fig. 2).

The items of the main orthogonal views of an object are presented in the third exercise.
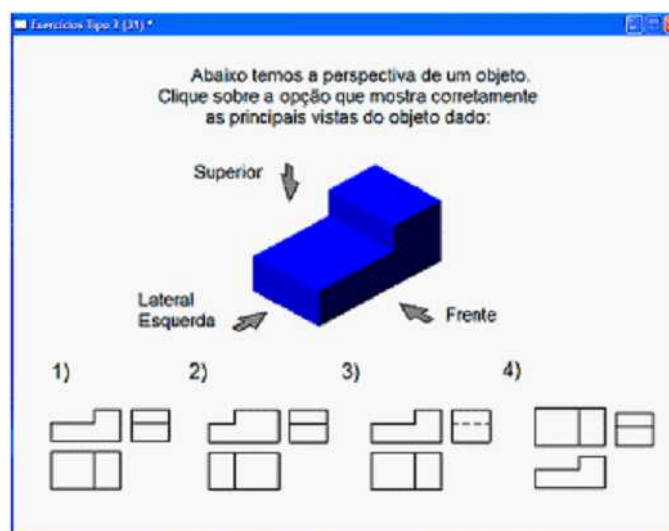

Figure 2: Example of an item in Exercise 2. The student is asked
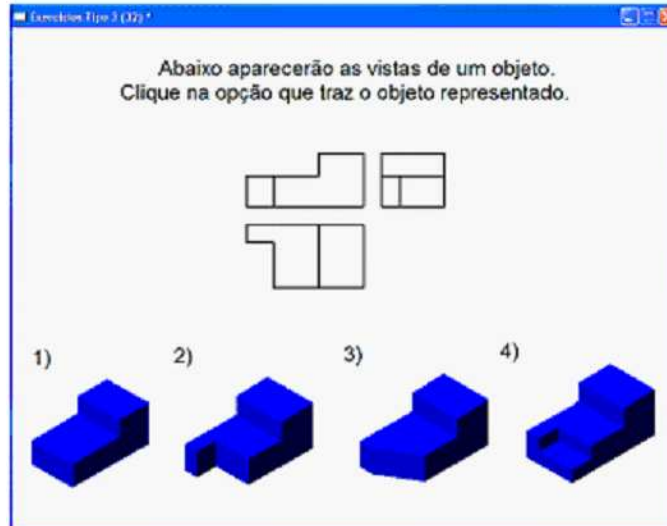to choose the option showing the principal views of the figure above.

Figure 3: Example of an item in Exercise 3. The student is asked to choose
the option corresponding to the views that appear above.

The student must identify the object to which the views correspond (see Fig. 3)

The fourth exercise is based on a task of recognizing sections (comprehension of the internal parts of an object).

To devise the distractors of exercises 2, 3, and 4, we took into account the students' most frequent errors when doing these types of exercises with paper and pencil in previous years (for example, errors in the proportions of the figures, in the positions of the views, in the types of line, absence of edges, etc.)

As mentioned earlier, the items are multiple choice questions and have four answer options, only one of which is correct. After choosing each answer, the students learn whether it is right or wrong and receive a feedback demonstration so that they can verify and improve their learning. If the answer is incorrect, they are obliged to look for the solution again. That is, they can not go on to the next item until their answer is correct (see Fig. 5).
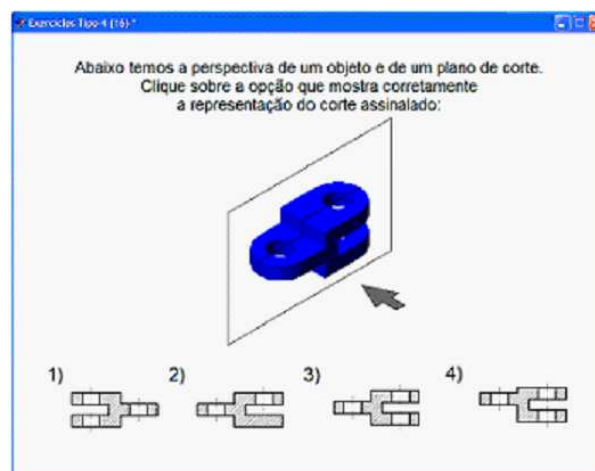

Figure 4: Example of an item in Exercise 4. The student is asked to choose
the option that corresponds to the section in the plane of the figure above.
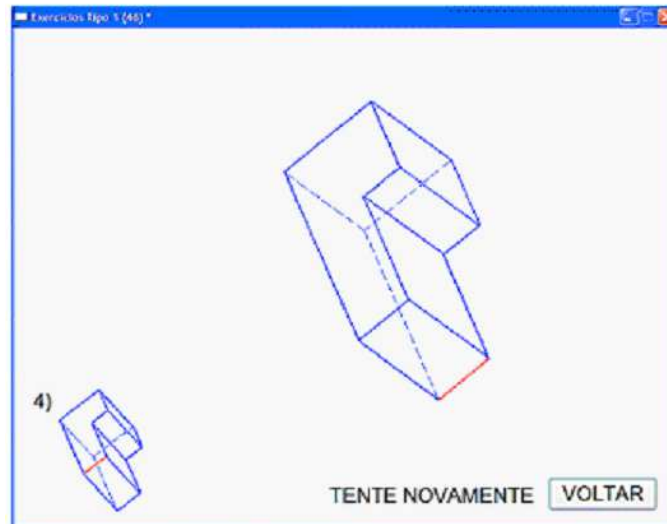
Figure 5: After seeing that his or her answer is wrong (smaller figure),
the student is asked to choose another option.

3. Method

Subjects

The participants were 239 university students from Brazil ($n = 47$), Argentina ($n = 59$), and Peru ($n = 133$), in their first year of engineering studies, with a mean age of 18.74 (SD = 1.83). Sixty-nine point five of them were men (69.5%).

Instruments

1. Two computerized Visualization tests (TVZ2006-A and TVZ2006-B) based on a Surface Development task were administered (Fig. 6). Each test was composed of 20 items with 9 options, only one of which was correct [23]. The students had thirty minutes to answer the items on each test. Both tests were calibrated in the same metrics by means of three anchor items. Test TVZ2006-A was easier than test TVZ2006-B. In order to minimize the effect of item memorization and the ceiling effect in the post-test, two tests of different levels of difficulty were included.

2. The first three training exercises described previously.

Versions of the exercises and the tests with instructions in Spanish and Portuguese were constructed.

Procedure

The TVZ2006-A test was administered at the beginning of the Technical Drawing course and the TVZ2006-B test at the end. There was a five month lapse between the two administrations. Technical Drawing is a theoretical-practical course about the graphic representation of objects. The purpose of the course is to give the students a basic knowledge of the representation of objects and to have them learn a software program for drawing and for plans. Course contents: graphic software for Windows; coordinate systems; geometric constructions; applied geometry; layers, dimensioning, isometric drawing, visualization of solids and multiple or dihedral projections. The contents of the course are very similar in the schools of the three
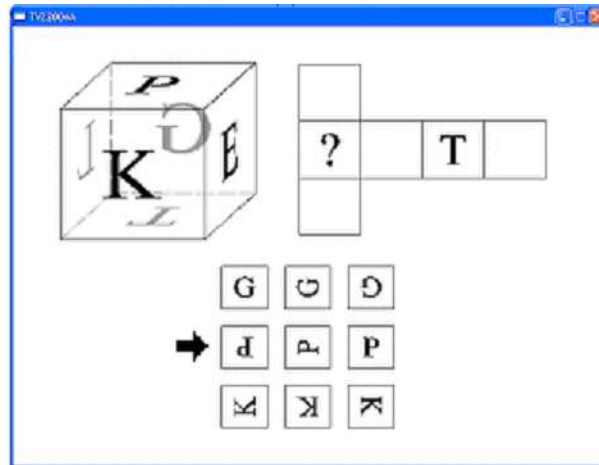
Figure 6: Example of a test item. The correct answer is indicated with an arrow.

countries involved. The three computerized exercises were included during the course with a separation of approximately one month between them.

Measurement models

The Rasch model [24] was used to obtain the pre- and post-measurements of Visualization tests and to measure the change, because it has the following advantages [7, 27]: measurement at interval level, estimation of specific measurement errors for each changed score, equating of test scores from tests having different levels of difficulty using anchor item design.

The Rasch Model for dichotomous responses (right/wrong) specifies the probability, P, that person n of ability B_n succeeds on item i of difficulty D_i, as follows:

$$\log_e(P_{ni}/(1 - P_{ni})) = B_n - D_i \quad [24].$$

It is a one-parameter logistic model. A person's ability in logits is the natural log odds for succeeding on items of the kind chosen to define the scale zero (which is conventionally set to the difficulty mean). An item's difficulty in logits is the natural log odds for failure on that item by participants with abilities at the scale zero. The Rasch parameter estimators are sufficient, consistent, efficient, and unbiased [1, 2, 25]. LUCE and TUKEY's [16] ordinal conditions for conjoint measurement are satisfied by this model and thus, given enough fit, the probability of a response can be expressed as an additive function of a participant parameter and an item parameter, which is consistent with the quantitative assumption [15].

Correct answers to the items in the exercises were weighted differently according to the order in which they occurred: on the first attempt (4), on the second (2), on the third (1) or on the fourth (0). This has been shown to be the most effective weighting in previous studies with the "Answer-Until-Correct" format [6, 10, 19]. That is, solving an item correctly on the first attempt indicates a much higher level of competence than solving it on later attempts.

As a result, to analyze the data of the exercises we used a measurement model for polytomous items called the *Partial Credit Model* [17]. The Partial Credit Model (PCM) is an extension of the Rasch model for polytomous items, i.e., it specifies the probability, P_{nij}, that person n of ability measure B_n is observed in category j of a rating scale specific to item i of difficulty measure D_i as opposed to the probability $P_{ni(j-1)}$ of being observed in category $(j-1)$, as follows: $\log_e(P_{nij}/P_{ni(j-1)}) = B_n - D_i - F_{ij}$. In this extension, F_{ij}, the step calibration, is the point on the latent variable, specific to item i, where the probability of being

observed in category j equals that of being observed in category $j - 1$. The PCM, being a Rasch-type model, has optimal metric properties.

To analyze the data, the WINSTEPS 3.64 program was used [12].

4. Results

Table 1 and Figure 7 show the descriptive statistics of the performance (in logits) on the three exercises and on the pretest and posttest measures of Visualization. Figures 7 and 8 show Wright maps, one of the utilities of the Rasch type models most employed [28]. The map allows us to visualize the joint distribution of the students' values and the items on the variable measured.

Table 1: Descriptive statistics of exercises and tests

Index	E1	E2	E3	VZ-Pretest	VZ-Posttest
Mean B	1.06	1.32	1.51	-.29	.21
SD	.63	.62	.67	1.51	1.70
Alpha	.73	.73	.63	.87	.89
R Br	.34*	.40*	.52*	.33*	.28*
R P	.34*	.47*	.32*	.46*	.48*

Mean B = Mean of person parameters

SD = Standard deviation

Alpha = Cronbach's alpha

R Br = Correlation with the grade in Technical Drawing (Brazil)

R P = Correlation with the grade in Technical Drawing (Peru)

Note: Grades for Argentina were not available

* $p < .05$

In Table 1 we can see that the performance of the students in the three exercises is high, given that the means are much higher than 0 (by convention, the mean difficulty of the items). The high performance in the tasks may be due to the low difficulty of the items (see their location and their low variability and the large difference on the scale between the mean of the students and that of the items in Fig. 7). Nevertheless, the variability in the students' performance is moderately high. The reliability or internal consistency (Cronbach's alpha) of the scores in the exercises is acceptable.

Table 1 and Figure 8 show the statistics and the distributions of the scores on the Visualization tests used in the pretest and posttest. The students showed high variability in the aptitude tests. It should be pointed out that the reliability of the scores is very high, and higher than that observed in the training exercises. The better reliability of the tests is due to the fact that the difficulty of the items is more suitable for the student's level of ability (Fig. 8).

Since the Visualization pretest and posttest are on the same scale, it can be seen (Table 1 and Fig. 8) that the students improved after taking the Technical Drawing course aided by computerized exercises. The difference between means of the posttest and the pretest (.50) is statistically significant (Student's $t = 3.6$; $p = .00$). To assess the magnitude of the improvement, we quantified the effect size using Cohen's d statistic [5], which indicates
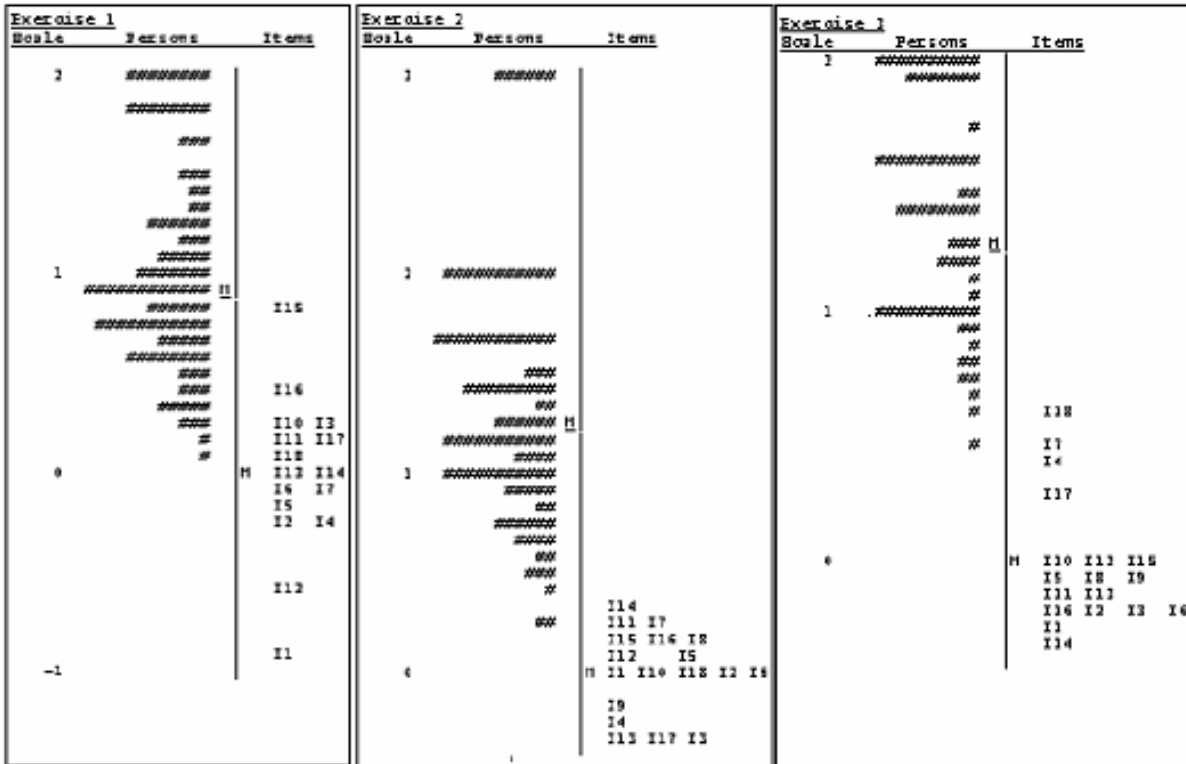
Figure 7: Wright Maps: Joint Person and Item Representation
along the Achievement Variable for Exercises.
M̲ = Person mean location; M = Item mean location

how many standard deviations the two measures differ by. In this case, the effect size is moderate ($\mathbf{d = .55}$). No significant differences in improvement were observed between the male students (mean = .54; SD = .94) and the female students (mean = .42; SD = .89): Student's $t = .98$; $p = .33$.

Table 1 shows the correlations between the scores on the aptitude test and the exercises and the final grades received in the Technical Drawing course. All the correlations are moderate and statistically significant. These data show that performance on the exercises is a predictor of success in the final exam. We have also replicated the results of studies that concluded that Visualization ability is associated with learning Technical Drawing [8, 9, 11, 14, 22].

5. Conclusions

Two fundamental conclusions can be drawn from this study. First of all, we have presented a battery of computerized exercises to assist the learning of engineering students in the subject of Technical Drawing. The tasks in the battery help to facilitate the mental representation of the internal structure or the transformations of three-dimensional objects. The exercises were conceived as a means for training. Feedback is therefore given after the student answers, offering an animated demonstration so that the students can understand why their answers were right or wrong. If they gave the wrong answer, they must continue to seek the solution. This answer format, called answer-until-correct, is considered very useful in computer assisted learning. Informal interviews given to a small group of students showed that they found
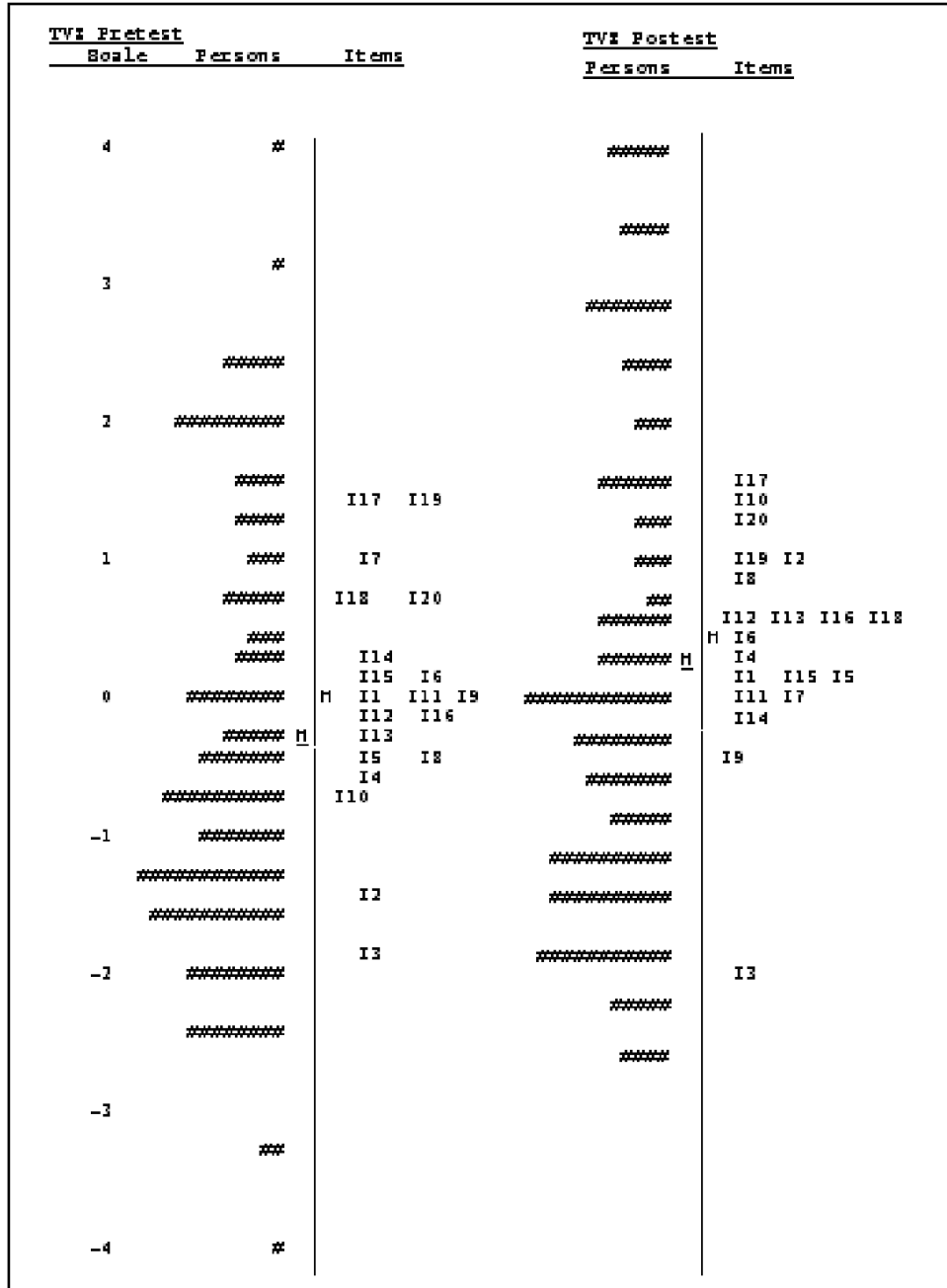
Figure 8: Wright Maps: Joint Person and Item Representation
along the Visualization Variable for Pretest and Postest.
$\underline{\text{M}}$ = Person mean location; M = Item mean location

both the contents of the exercises and the answer format very attractive and motivating. Furthermore, the exercises give a score for each student that assesses their competence in the exercise. The scores were obtained using the Partial Credit model. These scores show a moderate association with the final grades in the Technical Drawing course.

Secondly, we have obtained evidence that Visualization ability, a classic mediator of success in professions such as engineering, architecture and graphic design can be modified through long term indirect training. In this study, the improvement that occurred in stu-

dents with a high level of ability, engineering students, was moderate. We consider that the size of the improvement could be increased by introducing more exercises with greater difficulty than the ones used in this study. However, as an anonimous reviewer has indicated, our experimental design did not count with a control group and so the unique contribution of the computer exercises to the TVZ score improvement cannot be estimated. If possible, future studies should add a control group. Finally, despite the influence of the range restriction on the magnitude of the correlations, we replicate the association of Visualization ability with performance in the learning of Technical Drawing. For this reason, the use of a Visualization test at the beginning of a course will allow students with possible learning problems to be identified, so that they can receive special attention during the course.

Acknowledgements

References

[1] E.B. ANDERSEN: *Asymptotic properties of conditional maximum likelihood estimators.* Journal of the Royal Statistical Society **32**, 283–301 (1970).

[2] E.B. ANDERSEN: *A goodness of fit test for the Rasch model.* Psychometrika **38**, 123–140 (1973).

[3] M. BAENNINGER, N. NEWCOMBE: *The role of experience in spatial test performance: A meta-analysis.* Sex Roles **20**, 327–343 (1989).

[4] J. CARROLL: *Human cognitive ability: A survey of factor-analytical studies.* Cambridge University Press, New York 1993.

[5] J. COHEN: *Statistical power analysis for the behavioral sciences.* NJ: Erlbaum, Hillsdale 1988.

[6] B.G. DODD, W.R. KOCH: *Effects of variations in item stepvalues on item and test information in the partial credit model.* Applied Psychological Measurement **19**, 5–22 (1987).

[7] S.E. EMBRETSON: *Improving the Measurement of Spatial Aptitude by Dynamic Testing.* Intelligence **11**, 333–358 (1987).

[8] D.F. HALPERN, M.L. COLLAER: In P. Shah, A. Miyake (eds.): *The Cambridge Handbook of Visuospatial Thinking*, pp. 170–212, Cambridge University Press, Cambridge 2005.

[9] M. HEGARTY, V.K. SIMMS: *Individual differences in mental animation during mechanical reasoning.* Memory and Cognition **22**, 411–430 (1994).

[10] I. HIROSE: *Answer-Until-Correct Item Response Model with Restricted Number of Responses.* Japan Journal of Educational Technology **24**, 53–62 (2000).

[11] S. HSI, M.C. LINN, J.E. BELL: *The role of spatial reasoning in engineering and design of spatial instructions.* Journal of Engineering Education **86**, 151–158 (1997).

[12] J.M. LINACRE: *A User's Guide to WINSTEPS & MINISTEPS: Rasch Model Computers Programs.* IL: Winsteps.com, Chicago 2007.

[13] D.F. Lohman: *Spatial abilities as traits, processes, and knowledge.* In R.J. Sternberg (ed.): Advances in the psychology of human intelligence, pp. 181–248, NJ: Lawrence Erlbaum, Hillsdale 1988.

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

[15] J.D. Long, D. Feng, N. Cliff: *Ordinal analysis of behavioral data.* In J.A. Schinka, W.F. Velicer (Vol. eds.): *Handbook of Psychology, vol. 2: Research Methods in Psychology*, pp. 635–661, NJ: John Wiley & Sons, Hoboken 2003.

[16] R.D. Luce, J.W. Tukey: *Simultaneous conjoint measurement: a new type of fundamental measurement.* Journal of Mathematical Psychology **1**, 1–27 (1964).

[17] G.N. Masters: *A Rasch model for partial credit scoring.* Psychometrika **47**, 149–174 (1982).

[18] M.G. McGee: *Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences.* Psychological Bulletin **86**, 889–918 (1979).

[19] L.F. Menéndez: *Modelos psicométricos para el análisis de ítems aplicados con el formato responder hasta acertar.* Tesis doctoral, Universidad de Oviedo, Spain, 2008.

[20] W.B. Michael, J.P. Guilford, B. Fruchter, W.S. Zimmerman: *The description of spatial-visualization abilities.* Educational and Psychological Measurement **17**, 185–199 (1957).

[21] N.S. Newcombe, L. Mathason, M. Terlecki: *Maximization of Spatial Competence: More Important than Finding the Cause of Sex Differences.* In A. McGillicudy-De Lisi, R. De Lisi (eds.): *Biology, Society, and Behavior: The development of sex differences in cognition*, pp. 183–206, CT: Ablex, Wesport 2002.

[22] 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).

[23] G. Prieto, A.D. Velasco: *Training Visualization Ability by Technical Drawing.* J. Geometry Graphics **8**, 107–115 (2004).

[24] G. Rasch: *Probabilistic models for some intelligence and attainment tests.* University of Chicago Press, Chicago 1960.

[25] G. Rasch: *A mathematical theory of objectivity and its consequences for model construction.* Report from European Meeting on Statistics, Econometrics and Management Sciences, Amsterdam 1968.

[26] R. Wilcox: *Solving Measurement Problems with an answer-Until-Correct Scoring Procedure.* Applied Psychological Measurement **7**, 67–74 (1981).

[27] E.W. Wolfe, C.W.T. Chiu: *Measuring Pretest-Posttest Change with a Rasch Rating Scale Model.* Journal of Outcome Measurement **3**, 134–161 (1999).

[28] B.D. Wright, M.H. Stone: *Best test design, Rasch measurement.* MESA Press, Chicago 1979.