Journal for Geometry and Graphics
Volume 3 (1999), No. 2, 211–218

Design of Modified Mental Rotations Test and its Error Analysis

Kumiko Shiina[1] and Kenjiro Suzuki[2]

[1] *Research Division, The National Center for University Entrance Examinations
Tokyo 153-8501, Japan
email: shiina@rd.dnc.ac.jp*

[2] *Department of Computer and Graphic Science, College of Arts and Sciences,
The University of Tokyo, Tokyo 153-8902, Japan
email: ksuzuki@idea.c.u-tokyo.ac.jp*

Abstract. A new version of the MRT (M-MRT) was designed by modifying the original MRT (V-MRT) in three ways. (1) The structural questions that may evoke strategies other than mental rotation were excluded. (2) Each part was composed of the same three-dimensional objects. (3) In order to make the difficulty of each part equal, the order of the questions was changed on the basis of the simulation of the data of the V-MRT. The M-MRT was administered to 250 subjects and the data was analyzed. There were significant differences of the mean scores in the M-MRT between sexes and majors. The mean score in the M-MRT was significantly larger than that in the V-MRT as for the low score groups. The larger mean score in the V-MRT in the low score groups might be due to the increase of the solving speed caused by the changes in the design of the M-MRT. The mean points of the complete-subjects were almost the same between questions which belong to different parts but are composed of the same objects. This suggests that the difficulties of each part are almost same. The increase of the mean score is considered to be caused by the higher solving speed for Part 2.

Key Words: spatial ability, mental rotations, error analysis

MSC 1994: 51N05

1. Introduction

Recently, many researchers in graphics education have taken interest in evaluating students' spatial abilities. The Mental Rotations Test developed by Vandenberg and Kuse [1] (hereafter V-MRT) has been widely used for this purpose. The analysis of the problem solving process of the V-MRT using eye fixation data and protocol data revealed that scores in the V-MRT reflect the following aspects of spatial abilities: the speed of the mental rotation of

ISSN 1433-8157/$ 2.50

three-dimensional objects and the ability to unify strategies for mental rotation (Shiina et al. [2, 3]). The V-MRT is constructed from two kinds of items: mirror-image questions and structural ones. The latter evokes strategies other than mental rotation, such as that of detecting structural features. Such items may prevent low-ability subjects' taking a mental rotation strategy and this makes it difficult to directly evaluate mental rotation abilities. Previous research also suggested that the difficulties in Part 1 and Part 2 might be different. In order to eliminate such problems in the V-MRT and to directly evaluate mental rotation abilities a new version of the MRT was designed by modifying the original MRT. The modified MRT (hereafter the M-MRT) was administered to 250 subjects and an error analysis was made.

2. Design of modified MRT

2.1. Unification of strategies

The V-MRT contains 20 questions. The perfect score of the V-MRT is 40. Each question is composed of a criterion figure, two correct alternatives, and two incorrect ones or "distractors". The subjects are required to find the two correct alternatives. For half the questions, the distractors are rotated mirror-images of the criterion, while in the other 10 questions, the distractors are rotated images of one or two other structures. Hereafter, these questions shall be called, respectively, "mirror-image questions" and "structural questions".

Previous research reported that some subjects with low V-MRT scores solved the questions by using strategies other than mental rotation, such as detecting structural features or matching encoded descriptions of objects (Shiina et al. [2, 3]). As for the structural questions, subjects were able to reach the correct answer not by using mental rotations but by detecting structural features; for example, whether the upper arm is parallel to the lower arm or not. Structural questions may evoke strategies other than mental rotation and prevent subjects from taking a mental rotation strategy. In addition, in some structural questions errors occur easily. Shiina et al. [4] have pointed out three such structural questions: Q. 3, in which an invisible cube occurs by mental rotation and Q. 15 and 19, in which attention to the number of cubes is required in order to get the correct answers. To lead subjects to a mental rotation strategy, the structural questions were taken away from the V-MRT. So the M-MRT is composed of 10 mirror-image questions, which will be divided into two parts.

In addition to eliminate the structural questions, instructions were modified to guide subjects in mental rotation strategy. Vertical axes, which illustrate rotation in depth direction, were added to the example figures in the instructions, and the word "rotation" was used explicitly in them.

2.2. Equality between parts

In the V-MRT, five objects and their mirror-image objects compose the criterion figures and alternatives (Vandenberg et al. [6]). So, there are five kinds of mirror-image questions. Five objects (Ap, Bp, Cp, Dp, Ep) are shown in Fig. 1. There are five other objects (An, Bn, Cn, Dn, En) which are the mirror-images of the objects shown in Fig. 1.

Table 1 shows the name of the objects which compose each mirror-image question in the V-MRT, together with the completion-rates and mean points of the complete-subjects. The term "completion-rate" for each question is defined as the rate at which subjects chose two alternatives. The term "completion-subjects" for each question is defined as the subjects who chose two alternatives, or completed the question. As shown in Table 1, the completion-rates

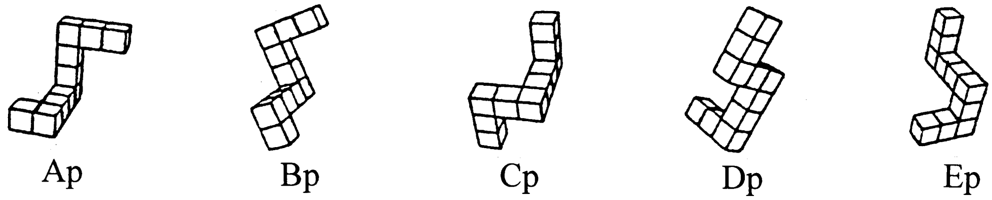

Figure 1: Objects used in the V-MRT

Table 1: Objects, completion-rates, mean points of complete-subjects in mirror-image questions in the V-MRT

Part	Part 1						Part 2			
Q. No.	1	2	5	6	9	10	13	14	17	18
Objects	Ap, An		Bp, Bn		Cp, Cn		Dp, Dn		Ep, En	
Completion-rates (%)	99.5	99.0	86.6	70.9	31.1	20.6	98.3	96.5	73.6	60.2
Mean points of complete-subjects	1.74	1.68	1.55	1.74	1.41	1.40	1.49	1.53	1.50	1.64

decrease in the latter questions in each part. These results indicated that a small number of subjects could reach the latter questions. Completion-rates are an index of solving speed. Mean points of the complete-subjects are an index of the difficulty of each question. The completion-rates and mean points of the complete-subjects were calculated using the data of 2058 subjects (Male: $n = 1407$, Female: $n = 651$) (Shiina et al. [5]). As shown in Table 1, objects which compose the questions in Part 1 are different from those in Part 2. Such asymmetry might influence the difficulty of each part. In the M-MRT, each part will be composed of the same objects, although these objects' presented positions are different between parts.

It is possible to make the 32 combinations of the mirror-image questions by choosing the questions as follows: (Q. 1 or Q. 2) and (Q. 5 or Q. 6) and (Q. 9 or Q. 10) and (Q. 13 or Q. 14) and (Q. 17 or Q. 18). Each part of these combinations contains five kinds of mirror-image questions, in which objects are used, respectively, (Ap and An), (Bp and Bn), (Cp and Cn), (Dp and Dn), or (Ep and En). To make subjects solve the same question as the first question, Q. 1 in the V-MRT will be also placed as Q. 1 in the M-MRT. The above placement conditions are satisfied by $32/2 = 16$ combinations.

To make the difficulty of each part equal, the sums of the mean points of the complete-subjects for each part were calculated on the basis of the 2058 subjects (Shiina et al. [5]) for each of the 16 combinations. A combination, whose difference of the sums for each part are minimum and negative, was chosen from 16 combinations. Negative difference will make it easier to detect the practice effect, which might occur between Part 1 and Part 2.

On the basis of the above consideration, the M-MRT was designed as the following combination of the mirror-image questions used in the V-MRT; Part 1: Q. 1, 5, 10, 14, and 18 (the calculated sum of the mean points of the complete-subjects was 7.85) and Part 2: Q. 2, 6, 9, 13, and 17 (the calculated sum of the mean points of the complete-subjects was 7.81). The solving time for each part of the M-MRT is 90 seconds. Full score in the M-MRT is 20.

3. Results and discussions

The M-MRT was administered to 250 freshmen and sophomores of three universities in March '97. The number of subjects whose major was Science and/or Engineering or Medical Science (hereafter labeled as "S-major") was 123 (Male: 102, Female: 21). The number of subjects whose major was Literature or Law or Science of Economics (hereafter labeled as "L-major") was 127 (Male: 86, Female: 41).

3.1. Mean scores

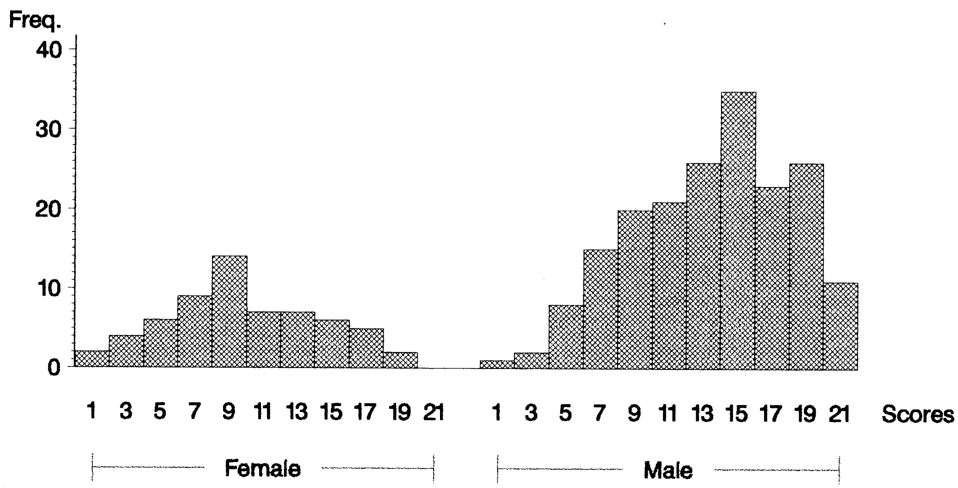

Figure 2: Histogram of scores in M-MRT

Fig. 2 shows the histogram of the scores in the M-MRT for each sex. The scores of the male subjects were distributed over higher ranges than those of the female subjects. Within the same sex and major, there was rarely significant difference between universities. So, data from three universities were dealt all together in the following analysis. The mean score in the M-MRT for all subjects ($n = 250$) was 11.9. Table 2 shows the mean scores in the M-MRT of each major and sex. There were significant differences between sexes ($p < 0.01$) and majors ($p < 0.05$). The highest mean score was that of the male group with "S-major" and the lowest mean score was that of the female group with "L-major".

Table 2: Mean scores and standard deviations in M-MRT of each major and sex

	S-major	L-major	All
Male	13.3 (4.3)	12.0 (4.6)	12.7 (4.5)
Female	10.2 (3.8)	8.8 (4.7)	9.3 (4.5)
All	12.8 (4.4)	11.0 (4.9)	11.9 (4.7)

3.2. Comparison of V-MRT and M-MRT

In order to examine differences between the V-MRT and the M-MRT, subjects with the similar properties were picked from the subjects who were given the V-MRT and the M-MRT, respectively. Male subjects with “S-major” of the same university were picked from the subjects who were given the V-MRT and the M-MRT (respectively, 592 and 37 subjects). We didn’t have enough data for female subjects with the same major of the same university. So, female subjects with the similar majors (respectively, Home Economics and “L-major”) were picked from the subjects who were given the V-MRT and the M-MRT (respectively, 318 and 41 subjects). The male groups are typical of high score subjects and the female groups are typical of low score subjects.

Table 3: Mean scores in percentages and standard deviations

Groups	V-MRT	M-MRT	Difference
High (Male)	72.1 (17.3)	69.9 (24.1)	−2.2
Low (Female)	37.6 (16.7)	44.0 (23.6)	6.4*

*) $p < 0.05$

Table 3 shows the mean scores in percentages in the V-MRT and the M-MRT and the differences between them for each group. As for the high score groups, there was no significant difference between the mean score in the V-MRT and that in the M-MRT. While, as for the low score groups, the mean score in the M-MRT was significantly higher than that in the V-MRT ($p < 0.05$).

Table 4: Mean percentages of the numbers of completed-questions

Groups	V-MRT	M-MRT	Difference
High (Male)	81.8 %	79.5 %	−2.3
Low (Female)	55.5 %	72.4 %	16.9

Table 4 shows the mean percentages of the numbers of the completed-questions in the V-MRT and the M-MRT. As for the high score groups, both of the mean percentage of the number of completed-questions in the V-MRT and that in the M-MRT were high and they were almost the same. As for the low score groups, the mean percentage in the M-MRT was, on the other hand, much higher than that in the V-MRT. This suggests the larger solving speed in the M-MRT. The subjects might have been encouraged to use mental rotation strategy by the elimination of the structural questions and the modification of the instructions. The increase in solving speed for the low score subjects might be caused by the unification of problem solving strategies to mental rotation.

Table 5 shows the mean points of the complete-subjects and the standard deviations for each group in the V-MRT and the M-MRT, respectively. As for the high score groups, there was no significant difference. As for the low score groups, there was a significant difference between the mean point of the complete-subjects in the V-MRT and that in the M-MRT. This suggests that the M-MRT was more difficult than the V-MRT for the low score subjects.

Table 5: Mean points of complete-subjects and standard deviations for all questions

Groups	V-MRT	M-MRT	Difference
High (Male)	1.74 (0.67)	1.71 (0.70)	−0.03
Low (Female)	1.32 (0.95)	1.16 (0.99)	−0.16**

**) $p < 0.01$

As discussed previously, the changes in the design of the M-MRT might encourage the low score subjects to unify their solving strategies to mental rotation. The unification might, then, prevent them from using strategies other than mental rotation such as detecting structural features and matching encoded descriptions of objects, by which they sometimes reached correct answers in the V-MRT (Shiina et al. [2, 3]). In addition, the structural questions had been eliminated in the M-MRT. One of the possible reasons of the higher difficulty in the M-MRT might be the changes in the design of the M-MRT.

As discussed above, the solving speed for the low score group was much larger in the M-MRT than in the V-MRT, though the problems were more difficult in the M-MRT. The larger mean scores for the low score group in the M-MRT shown in Table 3 might be due to the larger solving speed caused by the change of the design of the M-MRT. It should be, however, recalled here that the majors and the universities of the subjects were similar but slightly different between the two low score groups discussed here. The differences between subjects might cause differences in the mean scores, solving speed and difficulties of the tests. Further research regarding the source of the difference between the V-MRT and the M-MRT is required.

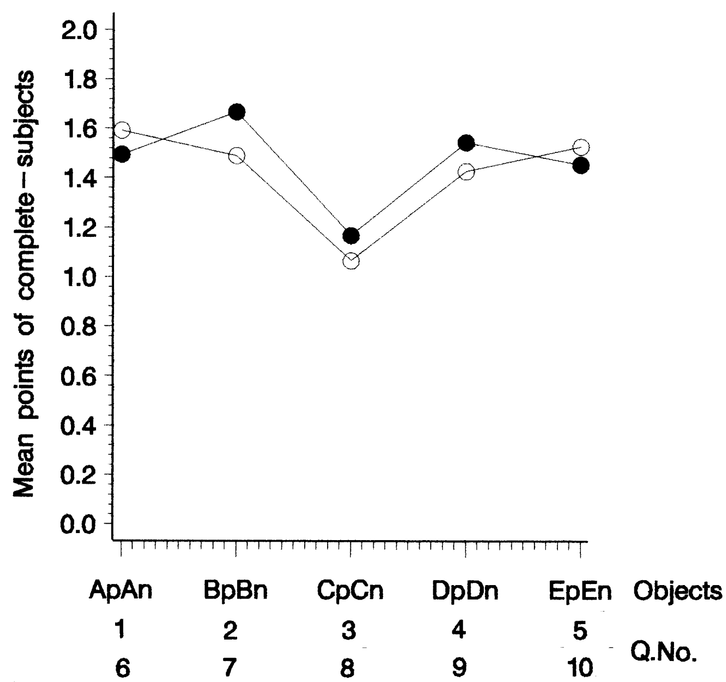

Figure 3: Mean points of complete-subjects for questions composed of the same objects
(○ : Part 1 (Q. 1–5), ● : Part 2 (Q. 6–1))

3.3. Difficulty of each part in M-MRT

In Fig. 3, the mean points of the complete-subjects, which were calculated using data of all subjects, are compared for questions which belong to different parts but are composed of the same objects. Among 5 pairs of such questions in the M-MRT, significant difference was not detected with the exception of a pair of the questions which are composed of objects "Bp" and "Bn". It suggests that the difficulties of each part are almost same, although objects are exposed to subjects for a second time in Part 2.

Fig. 3 indicates that the mean points of the complete-subjects for questions composed of "Cp" and "Cn" are lower than those of the questions composed of other objects. It can be said that the questions composed of "Cp" and "Cn" are more difficult than the questions composed of other objects.

(a) "Cp" (b) "Cn"

Figure 4: "Cp" and "Cn" as alternatives

Fig. 4 indicates "Cp" and "Cn" as the alternatives in the questions. As shown in Fig. 4, parts of the objects hide other parts. It may be difficult to recognize the whole shapes in the projected figures. The difficulties may be caused by the shapes of the projected-figures of the the objects "Cp" and "Cn" as the alternatives.

Table 6 shows the mean scores and the mean numbers of completed-questions of each part. The mean score of Part 2 was significantly higher than that of Part 1. The mean number of completed-questions of Part 2 was significantly larger than that of Part 1. This suggests that the solving speed of Part 2 was higher than that of Part 1.

Table 6: Mean scores and mean numbers of completed-questions for each part

	Part 1	Part 2	Difference
Mean scores	5.40	6.46	1.07**
Mean numbers of completed-questions	3.37	3.81	0.44**

**) $p < 0.01$

As already shown in Fig. 3, the difficulties of each part are almost same. So, the higher mean score in Part 2 can be considered to be caused by the increase in solving speed.

4. Summary and conclusion

A new version of the MRT (M-MRT) was designed by modifying the original MRT (V-MRT) in three ways.

(1) The structural questions that may evoke strategies other than mental rotation were excluded.

(2) Each part was composed of the same three-dimensional objects.

(3) In order to make the difficulty of each part equal, the order of the questions was changed on the basis of the simulation of the data of the V-MRT.

The M-MRT was administered to 250 subjects and the data was analyzed. There were significant differences of the mean scores in the M-MRT between sexes and majors. In order to examine differences between the V-MRT and the M-MRT, subjects with similar properties were picked from the subjects who were given the V-MRT and the M-MRT, respectively. The mean score in the M-MRT was significantly larger than that in the V-MRT as for the low score groups, while the M-MRT was more difficult than the V-MRT for low score groups. The larger mean score in the M-MRT in the low score groups might be caused by the increase of the solving speed. The mean points of the complete-subjects were almost the same between questions which belong to different parts but are composed of the same objects. This suggests that the difficulties of each part are almost same. However, the mean score of Part 2 was significantly higher than that of Part 1. The increase of the mean score is considered to be caused by the higher solving speed for Part 2.

Acknowledgements

This research was conducted as a part of Joint Research (April '96 – March '99) at the National Center for University Entrance Examinations. The authors express sincere appreciation to all the principals and researchers who assisted with the data collection for this study by taking the time to complete the survey.

References

[1] S.G. Vandenberg, A.R. Kuse: *Mental rotations a group test of three-dimensional spatial visualization.* Perceptual and Motor Skills **47**, 599–604 (1978).

[2] K. Shiina, T. Saito, K. Suzuki: *Analysis of problem solving process of a Mental Rotations Test.* Proc. 6th ICECGDG, Tokyo 1994, pp. 810–814.

[3] K. Shiina, T. Saito, K. Suzuki: *Analysis of problem solving process of a Mental Rotations Test.* J. Geometry Graphics **1**, 185–193 (1997).

[4] K. Shiina, T. Saito, K. Suzuki, T. Jingu, E. Tsutsumi: *Evaluation of students' spatial abilities by a Mental Rotations Test.* Proc. 5th ICECGDG, Melbourne 1992, pp. 286–290.

[5] K. Shiina, K. Suzuki: *Consideration of problem solving process of Mental Rotations Test by error analysis* [in Japanese]. Journal of Graphics Science of Japan **78**, 3–10 (1997).

[6] S.G. Vandenberg, A.R. Kuse: *Answer Key for Mental Rotations Test.*

Received August 14, 1998; final form November 2, 1999