Individual Differences in Imagery Ability and Mental Size Comparison

著者	HATAKEYAMA　TAKAO
journal or publication title	Tohoku psychologica folia
volume	43
page range	112-126
year	1985-03-30
URL	http://hdl.handle.net/10097/62652

Tohoku Psychologica Folia
1984, **43** (1-4), 112-126

INDIVIDUAL DIFFERENCES IN IMAGERY ABILITY AND MENTAL SIZE COMPARISON

By

TAKAO HATAKEYAMA (畠山孝男)[1]

(*Yamagata University*)

The present research examined how several imagery tests could predict the objective performance and the subjective estimates in the two mental size comparison tasks, i.e., object comparison and clock comparison. In both tasks, a clear "symbolic distance effect" (Moyer & Bayer, 1976) was revealed. Linear regression tests for the reaction time of individual subjects, however, was not related to the degree of the subjective estimates on imagery use, though that degree was related to the performance. The Betts QMI was predictive of the response speed in both tasks, namely, the subjects with high vivid imagery responded more quickly, especially at small size differences. The Space Relations predicted the faithfulness to the imagery instruction and the confidence of correctness. The Flags and the Gordon TVIC, in the clock task, was related to error responses, and the latter was in turn also to the estimates of imagery use frequency. When the QMI was combined with the Flags, the Space Relations and the TVIC, respectively, the high vivid-low control subjects responded quickly and the low vivid-high control subjects did slowly in the clock task, especially at small size differences. The MPFB, the VVIQ and the VVQ did not have any prominent predictive efficiency. Consideration was made on how individual subjects actually solved the tasks by means of their imagery ability.

Key words: imagery ability, mental size comparison, individual differences, symbolic distance effect, Betts' Questionnair upon Mental Imagery.

Though researches concerning imagery have been accumulated, few researches have been made on the individual differences in imagery process on cognitive tasks and in imagery ability, and very little is known about them. The author revealed some interesting findings in the study which examined how three spatial and three questionnaire tests, which are all widely used as imagery tests, could predict the performance on two sorts of mental rotation tasks, i.e., post-stimulus and pre-stimulus (preparatory) rotation tasks (Hatakeyama, 1981). It was found in this research that to investigate how imagery tests are related to well-investigated cognitive tasks, such as mental rotation, would be a good means to clarify not only what aspects of and how each test would predict the tasks, but how imagery would actually function and how differently each individual would use imagery. The present research examined how several imagery tests could predict the performance in mental size comparison tasks.

Moyer (1973) showed the fact that the time needed for comparing the size of animal in memory was inversely related to the estimated difference in animal size.

1. Section of Psychology, Faculty of Education, Yamagata University, Kojirakawa-machi, Yamagata, 990 Japan.

This "symbolic distance effect" (Moyer & Bayer, 1976) han been replicated not only for the size dimension (Holyoak, 1977; Holyoak, Dumais & Moyer, 1979; Jamieson & Petrusic, 1975; Kosslyn, Murphy, Bemesderfer & Feinstein, 1977; Paivio, 1975, 1978 a) but for abstract dimensions such as intelligence (Banks & Flora, 1977), pleasantness and monetary value (Paivio, 1978 b).

About the relation with imagery ability, Paivio (1978 a, b, c) already examined in the mental comparison tasks, i.e., comparison of mental clocks, shape comparisons, and comparison on pleasantness and on value, and revealed that the subjects with high imagery ability responded faster than low imagery ones. In Paivio, however, imagery ability was solely defined by the sum of standard scores of three spatial tests. As Kosslyn (1980) pointed out, the test approach to individual differences in imagery use would be useful if one had some priori reason for believing that given tests do in fact require imagery (p. 404). Each of the imagery tests, including self-report techniques, must be examined by the tasks which have been proved to reflect imagery processes. Particularly, we regret that Paivio abandoned the two questionnaire tests which he had administered as part of a battery, bacause we got interesting findings in relation to the mental rotation tasks by combining self-report vividness tests with self-report and objective controllability tests (Hatakeyama, 1981), and J. Richardson (1979) revealed by analyzing subjects' introspective reports that mental imagery was used in making mental comparisons among concrete objects in terms of both physical properties and abstract properties.

The present research was made on two mental size comparison experiments; Experiment 1 on objects and Experiment 2 on clocks. In the former, imagery instruction was not given, as in Moyer (1973) and Paivio (1975, Experiment 1). In the latter, imagery instruction was given, as in Paivio (1978 a). By analyzing the relation of the experiments with imagery tests, we might clarify how each of the tests would predict the performance of the mental size comparison tasks and how individual subjects would actually solve the tasks by means of their imagery ability. Moreover, if any relation with imagery tests would appear, we might prove that imagery would function spontaneously when subjects compare object sizes and that it would be used voluntarily when they compare clocks. We got some subjective estimates made by the subjects themselves centering around on their imagery use. Of course we should not give an explanation to the phenomenon solely on the ground of their introspections, but we would use them supplementarily to understand subjects' solving processes. The estimation data on the imagery use frequency would be especially important if imaginary process would be indispensable in the tasks. The estimates on the correctness of responses would be indices of subjects' confidence.

This research was somewhat tentative and did not lead directly to some plain hypotheses. Some predictions would be shown below.

(a) If imagery is not a mere accompanying phenomencn but has an substantial function, the vividness tests will have something to do with the tasks but the controllability tests have nothing, since the subjects should be required to visualize images but not be required to transform them in the tasks.

(b) The vividness tests will be related not only to objective indices of reaction time and errors but to subjective estimates on imagery use frequency, etc.

(c) The combinations of vividness and controllability tests will not have substantial predictive efficiency.

(d) Individual differences in imagery ability will play a more essential part in the clock task (Experiment 2) than in the object task (Experiment 1), since the former should be more difficult to solve without visualizing, unless the subjects will use a mathematical strategy.

Experiment I

The question was examined how the imagery tests predicted objective performance (reaction time and error) and subjective estimates (imagery use frequency and response correctness) in the mental size comparison task of objects.

Method

Subjects: One hundred and seven subjects (45 men and 62 women), all students at Yamagata University, were selected for the analysis. They participated in all the imagery tests and Experiment 1 and 2, and did not need extremely long reaction time and/or make too many error responses.

Materials: 129 items, names of 88 common objects and 41 animals, were selected out of 158 items in Holyoak, Dumais & Moyer's (1979) Table 3 and out of 176 items in Paivio's (1975) Table 1, considering the familiarity to Japanese students. For the size scales of the former items, the mean ratings of Holyoak *et al.*'s were wholly used, because those norms were established more strictly than Paivio's, and for the latter items, the author, positioning them arbitrarily in the ordered list of Holyoak *et al.*, gave expedient norms to them. Using these items, the author selected two lists of 40 pairs so that each of the five size differences, namely, 0, 1, 3, 5 and 7, had eight pairs of items. The size differences were calculated by subtracting the size norm of one member from the size norm of the other. Any given item occurred only once in any list, and these two lists had 118 items in all. The items were expressed in daily Japanese, namely, in *kanji* (Chinese character), *hiragana* (the cursive *kana* character) or *katakana* (the square form of *kana*), and each word was expressed by using one to eight characters. The left-right positions of larger members of the pairs were counterbalanced in each list, under the restriction that the same side would not be used for larger members more than three times, except the pairs of "0" size difference. The stimulus pairs were photocomposed horizontally in Gothic types on 20×20 cm white field, with 2 cm separating each item in the center. The height of the stimulus words was 1.6 cm and the width was 1.3 to 8.7 cm, which approximately corresponded to the number of characters.

Procedure: The stimuli were presented by means of a Takei DP-type three-field tachistoscope, and reaction times were recorded on a Takei Digitimer in msec. The timer started when the stimulus was illuminated and stopped when the subject pressed either left- or right-hand key. In the same instance, one of the two little red lamps lighted on, which told the experimenter which side key the subject pressed. Subjects were instructed to press the key on the side the member on which is thought larger one in real life for each pair of named objects. Each trial went on as follows: first there was a "ready" signal, and a pre-exposure field with a little black cross centered was presented for 2,000 msec, which was followed by the exposure of the stimulus pair. After the subject pressing the key, the experimenter reset the tachistoscope. The procedure was illustrated by showing an example of a pair of printed names. 6 practice trials were given prior to the 40 experimental trials. After the experimental task, subjects were asked to give an introspection about their strategies they had used in the task, and asked to make two estimates in percent concerning how often they had used imagery and how correctly they had responded.

Imagery Tests: Imagery ability was measured by four questionnaire and three spatial-manipulation tests. The former included Sheehan's modification of Betts' Questionnaire upon Mental Imagery (QMI) (A. Richardson, 1969), Gordon Test of Visual Imagery Control (TVIC) (A. Richardson, 1969), Marks' (1973) Vividness of Visual Imagery Questionnaire (VVIQ) and A. Richardson's (1977) Verbalizer-Visualizer Questionnaire (VVQ)[2]. The latter included the Flags (Thurstone & Jeffrey, 1959), the Space Relations (a subtest of the Differential Aptitude Tests, Bennett, Seashore & Wesman, 1974) and the Revised Minnesota Paper Form Board Test (MPFB) (Likert & Quasha, 1970)[3].

Results and Discussion

For each subject, median and quartile deviation of reaction times were used for each of the five size differences as the measures of subject's reaction time (RT) and its dispersion (Q). The measure of error (Er) was the frequency of error responses for each size difference.

(a) RTs, Qs and Ers, averaged over subjects, all were decreasing according to the size differences (Fig. 1). This demonstrates a typical "symbolic distance effect" (Moyer & Bayer, 1976), also with respect to the Q, i.e., the dispersion of reaction times within each subject. Fig. 1 shows that the symbolic distance effect diminishes markedly when the size norm differences are beyond "3" (e.g., *penguin-ear*) for both RT (including its Q) and Er, and Er disappears substantially beyond "5" (e.g.,

2. We used these questionnaire tests in Japanese versions, which were translated by the author.
3. The time limits were shortened for the Space Relations and the MPFB

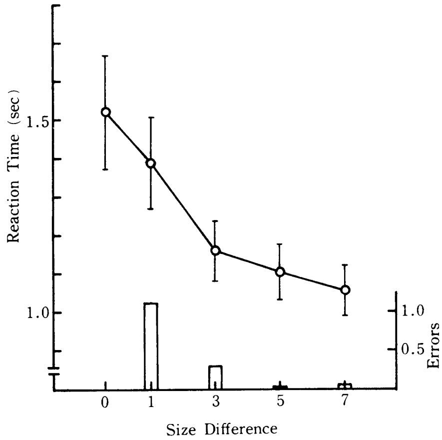

Fig. 1. Mean RT, Q and Er as a function of scaled size difference in Experiment 1.

key-cow).

Subjects' individual estimates in percent on the frequency of imagery use ranged from 0 to 100% and their mean estimate was 68.7%. According to their introspection, many subjects reported that they could judge directly, i.e., without imagining, when the words were the names of extreme large or small objects (e.g., *whale*, *flea*), and that they visualized objects when the words were the names of similar sized objects. Therefore, we can say subjects believed that they used imagery in this mental comparison task.

Did their subjective recognition concerning imagery use have substantial function in performing the task ? To examine this, we divided the estimates into three parts and classified the highest as high frequency (H), i.e., upper than 80%, and the lowest as low frequency groups (L), i.e., lower than 50%. A two-way analysis of variance, with repeated measure on the size difference variable, was made on RTs, Qs and Ers (Kirk's
06 (1968) unweighted-mean solution for Type SPF-$p.q$ Design). The H group showed a tendency to make less errors at small size differences [main effects, F (1,62)=3.31, $.05<p<.10$; interaction, F (3,186)=2.25, $.05<p<.10$]. A t-test revealed that the H group made higher estimates on the correctness than the L group [t (62)=2.67, $p<.01$]. It said that the subjects who reported more frequent imagery use had more confidence in the performance and made less errors than the subjects who reported less frequency. This might suggest that the subjective recognition concerning imagery use has a substantial relation to the performance in the mental size comparison task of objects.

Moreover, to examine if the symbolic distance effect reflected subjective imagery use, we performed linear regression tests for reaction times of individual subjects and observed how the subjects with significant linear regression slopes at $p<.05$ were included in the groups mentioned above, but in this case, including the middle frequency group. Table 1 revealed no differences among the groups. Therefore, it can be said that the symbolic distance effect which appears in the mental object task does not reflect directly the subjective imagery use.

(b) Sex difference was analyzed by performing a two-way analysis of variance on RTs, Qs and Ers, and it was analyzed by a t-test on imagery use and correct response estimates. Sex difference was found only on the estimates of correct responses, namely, men estimated more highly than women, means of 80.0 and 74.7%, respectively [$t(105)=2.26$, $p<.05$]. Thus, women gave estimations modestly for some unknown reason.

Table 1. Number of subjects with linear regression slopes for each frequency level of imagery use in Experiment 1

Frequency Level		Regression	Non-Regression
High	(N=54)	49 (90.7%)	5 (9.3%)
Middle	(N=21)	21 (100.0)	0 (0.0)
Low	(N=32)	28 (87.5)	4 (12.5)

Table 2. Correlations between imagery tests

	Flags	Space	MPFB	TVIC	QMI	VVQ
Space	.24*					
MPFB	.00	.22*				
TVIC	.23*	.14	.02			
QMI	.07	.09	−.03	−.06		
VVIQ	.02	−.05	−.03	−.26**	.32**	
VVQ	.03	−.02	−.02	.27**	−.09	−.16

*$p<.05$ **$p<.01$

(c) The correlation coefficients between each of the tests were in Table 2. They showed low or no correlations. Therefore, these tests are all called and used as imagery tests, but they would measure different aspects of imagery, if they would do.

The relation between imagery ability and performance was analyzed. First, the scores on each imagery test were divided into three parts, whose upper one third and lower one third were called high imagery (H) and low imagery groups (L), respectively. For the Betts QMI and the VVIQ, the upper was classified as L and the lower as H. Since the sex difference was found in the scores of the Flags and the MPFB, namely, men were superior than women in the former, and the reverse in the latter, classifications in terms of these tests were made within each sex.

The RTs, Qs and Ers were analyzed by a two-way analysis of variance. For the QMI alone, the H group revealed a tendency to respond faster than the L group (Fig. 2) 〔$F(1, 71)=3.63$, $.05<p<.10$〕. Though interaction did not reach significance, there was a large discrepancy between the mean RTs of both groups at "0" size difference, as shown in Fig. 2. The dispersion of the reaction times was also less in the H group 〔$F(1,71)=5.78$, $p<$

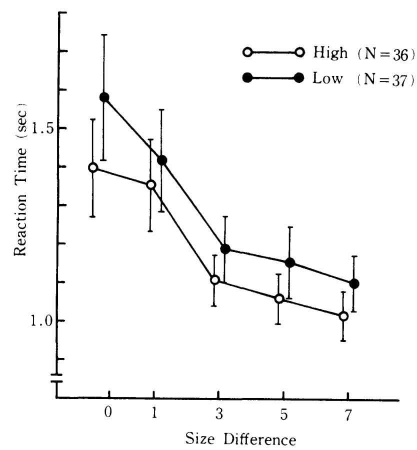

Fig. 2. Mean RT and Q as a function of scaled size difference for the two imagery groups on the Betts QMI in Experiment 1.

.05].

Estimates of the imagery use and of the correct responses were analyzed by a t-test. Here again, for the Betts QMI, the H group estimated more highly than the L group on imagery use frequency, means of 73.6 and 61.1%, respectively [t (71)=2.28, $p<.05$]. The QMI test is supposed to measure imagery vividness. Therefore, these two results say that the subjects who thought themselves to possess vivid imagery reported more imagery use and responded faster than the subjects who thought themselves to possess weak imagery. This finding suggests that imagery is actually used in the process of mental size comparison of objects, even if any imagery strategy instruction is not given. Here, we can make a claim that the vividness of imagery is a matter of entirely subjective phenomenon but has an important function in cognitive process.

The VVIQ should be expected to have been related to the task in the same manner as the QMI, but it was not the case.

Another interesting result was found for the Space Relations on the estimates of correct responses, namely, the H group gave a higher estimation than the L group [$t(68)=3.45$, $p<.005$]. Though there was not any significant difference between both groups in the actual performance, the H group seemed to have possessed higher confidence in performance than the L group. This test can measure, according to the manual, both "the ability to visualize a constructed object from a picture of a pattern" and "the ability to imagine how an object would appear if rotated in various ways." The higher confidence of the H group in the present task will be discussed later.

(d) The relation between vividness and controllability of imagery was analyzed. The scores on each test were divided at the median into high vivid or high control (H) and low vivid or low control groups (L). Each vividness test, i.e., the Betts QMI and the VVIQ, was combined with each controllability test, i.e., the Flags, the Space Relations, the MPFB and the Gordon TVIC, and for each combination we had high vivid-high control (H-H), high vivid-low control (H-L), low vivid-high control (L-H) and low vivid-low control groups (L-L). There was not any prominent difference among the groups in each combination. Considering this together with the fact that any controllability test was independently unrelated to the performance, it seems that the present mental size comparison task does not require any imagery controllability.

Experiment II

The question was examined how the imagery tests predicted objective performance (reaction time and error) and subjective estimates (faithfulness to the imagery instruction, clearness of images, imagery use frequency and response correctness) in the mental size comparison task of clocks. The two estimations were added in the present task in accordance with the imagery strategy instruction.

Method

Subjects : The same one hundred and seven subjects as in Experiment 1 were also used for the analysis in the present experiment.

Materials : In accordance with Paivio's (1978a) Experiment 1, two 40-pair lists of numerical times were formed so that each of the five angular size differences, namely, 30°, 60°, 90°, 120° and 150°, had eight pairs of numerical times. The angular differences were calculated by subtracting the angular separation between the hour and minute hands of one time from the angular separation of the other. Discrepancies from Paivio (1978a) were found in the following two points. First, when selecting pairs of times, their mutual positions in the clock face were not considered in our experiment. Second, none of the times were repeated in each of our lists.

The left-right positions of larger members of the pair were counterbalanced in each list, under the restriction that the same side would not be used in succession for larger members more than three times. The stimulus pairs of times were printed horizontally in Gothic face capitals on 20×20 cm white field, with 3.5 cm separating each item in the center. The height of the stimulus was 1.3 cm and the width of each stimulus time was 3.5 to 4.2 cm.

Procedure : After Experiment 1, subjects participated in the present experiment. The general procedure was same as in Experiment 1. Subjects were instructed to press the key on the side of the larger member in a given pair of times, by imagining a pair of clock faces and comparing the angles formed by the clocks' hands. The procedure was illustrated by showing an example of a pair of printed times with the corresponding clock faces and hands drawn under. They were then given 6 practice trials prior to the 40 experimental pairs. After the experimental task, subjects were asked, in the first place, to make two ratings on 7-point scale concerning two questions, "Did you follow the experimenter's instruction when you compared the angles ?" and "Did you imagine clock faces clearly ?" The number 1 and 7 of the scales corresponded to the descriptions of "never" and "very well," respectively. Then they were asked to make another two estimates in percent concerning how often they had used imagery and how correctly they had responded. Lastly, they were asked concerning another specific devices, if any.

Imagery Tests : The same scores of the same tests as those in Experiment 1 were used.

Results and Discussion

The same analyses were made as in Experiment 1.

(a) Mean RTs, Qs and Ers all decreased with the increase of angular size difference (Fig. 3), which revealed a clear symbolic distance effect, as in Experiment 1

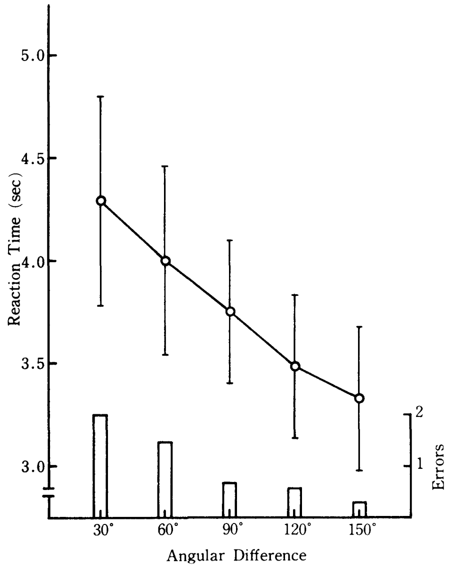

Fig. 3. Mean RT, Q and Er as a function of angular size difference in Experiment 2.

and as Paivio (1978a). In the present task, an imagery instruction was given to the subjects in advance. Subjects' estimates on their faithfulness to the instruction ranged from 2 to 7 in 7-point scale and the mean was 4.7. Their estimates on the frequency of imagery use ranged from 10 to 100% and the mean was 65.3%. For the clearness of their images, they gave a mean estimate of 3.7. Therefore, on the average, they performed the task more or less faithfully to the instruction and imagined clock faces in about two third trials, but the clock faces were subjectively not so very clear. According to their introspection, they could distinguish the difference of the size without imagining clock faces when numerical time was known to have extremely large or small angle or when it was accustomed for the subjects for some reasons. Mathematical strategies were reported by 17 subjects (15.7%), as far as we did not take the degree of the frequency into consideration. On the whole, subjective estimates and reports tell us that the imagery use is occurred in performing the mental clock task.

A two-way analysis of variance was made for the estimation on imagery use frequency. The high frequency subjects (H), i.e., upper than 80%, made responses more slowly and made less error responses than the low frequency subjects (L), i.e., lower than 50% [$F(1,72)=3.22$, $.05<p<.10$; $F(1,72)=4.87$, $p<.05$].

A t-test revealed that the H group made higher estimates than the L group on the faithfulness to the instruction [$t(72)=7.29$, $p<.001$], on the clearness of imagined clock faces [$t(72)=4.75$, $p<.001$] and on the correctness [$t(72)=4.17$, $p<.001$]. It might be that the subjects who reported more frequent use of imagery performed the task more faithfully to the imagery instruction, imagined more clearly and had more confidence in the performance. And they, interestingly, took longer time to respond, and made less errors. It may be said that the subjects who reported less frequent imagery use responded after an insufficient comparison and made more errors. The subjective recognition concerning imagery use has a substantial relation to the performance in the mental size comparison task of clocks.

We observed how the subjects with significant linear regression slopes at $p<.05$ were included in the three groups, which were classified by the degree of imagery frequency estimation. Table 3 revealed no differences among the groups. Here again, we can propose that the symbolic distance effect which appeares in the mental clock task does not necessarily reflect the subjective use of imagery.

Table 3. Number of subjects with linear regression slopes for each frequency level of imagery use in Experiment 2

Frequency Level		Regression	Non-Regression
High	(N=38)	25 (65.8%)	13 (34.2%)
Middle	(N=33)	25 (75.8)	8 (24.2)
Low	(N=36)	27 (75.0)	9 (25.0)

(b) Sex difference was found in the subjective estimates on correctness and on clearness of imagined clocks, namely, men were higher than women [$t(105)=4.97$, $p<.001$; $t(105)=2.16$, $p<.05$]. In the actual performance, significant interaction was found on Ers, namely, men showed more error responses at smaller angular differences and less at larger angular differences than women [$F(4,420)=2.16$, $.05<p<.10$]. The difference of the estimates of correctness between both sexes became larger in this task when compared with Experiment 1. This might be attributed to the fact that women imagined clock faces which were less clear subjectively and they made more errors at the small angular differences, though their estimates concerning the performance had a tendency to be moderate.

(c) For the Betts QMI, the H group responded faster than the L group, especially at small angular differences (Fig. 4) [main effects, $F(1,71)=3.29$, $.05<p<.10$; interaction, $F(4,284)=5.16$, $p<.01$]. The dispersion of the reaction time was less in the H group as well [$F(1,71)=3.10$, $.05<p<.10$]. However, very interestingly, there were no differences both on the estimates of imagery use and of the clearness of imagined clocks. No difference on the former might be caused by the fact that an imagery strategy instruction was explicitly given in advance. About the latter, we could expect higher estimates for the H group, because they are supposed to possess more vivid imagery. But that was not the case. It might be that the task did not need high degree but some degree of clearness of imagery clock faces that made it possible for the subject to judge the difference. Considering the results altogether, it might be said that one of the factors underlying individual differences in imagery vividness is the difference in the speed with which the image can be produced in some degree of clearness that makes it possible to

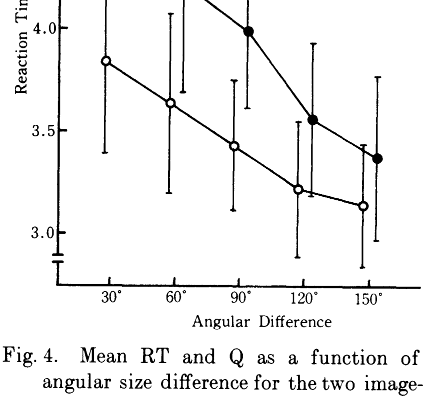

Fig. 4. Mean RT and Q as a function of angular size difference for the two imagery groups on the Betts QMI in Experiment 2.

be used.

The VVIQ did not relate to the present task, too.

For the Space Relations, a significant effect on the estimation of correctness was revealed, as was in Experiment 1 [$t(68)=2.05$, $p<.05$]. Also on the faithfulness and the imagery clock clearness, the H group made higher estimates than the L group [$t(68)=2.08$, $p<.05$; $t(68)=1.78$, $.05<p<.10$]. However, again, there were no significant effects in the performance. This test was already proved to predict the preparatory mental rotation, i.e., the subjects who got high scores on the test performed the rotation in a manner *faithful* to the instruction and made a template match well (Hatakeyama, 1981). The facts are very interesting that the H group had higher confidence on performance in both of the object and the clock comparison tasks, and had higher faithfulness to the imagery strategy instructions in the clock task and in the preparatory mental rotation task (the latter required subjects to rotate an image in the abscence of stimulus). In sum, the H group classified by the Space Relations obeyed the imagery instruction well, imagined more clearly, and had more confidence in the performance. The reason why this test was not related to the actual performance would be that the size comparison tasks in this article did not require subjects to transform images.

For the Flags, the H group made less error responses than the L group [$F(1,69)=5.42$, $p<.05$]. Moreover, the H group gave somewhat higher estimates on the clearness of imagery clocks than the L group [$t(69)=1.96$, $.05<p<.10$]. This test is, according to the test manual, a measure of "the ability to visualize a rigid configuration when it is moved into different positions." Through the examination of the test with reference to mental rotation tasks (Hatakeyama, 1981), this test was confirmed that the measurement of an orientation ability was made, i.e., an ability to identify fast and accurately the top of the rotatedly-presented stimulus. In the mental clock task, it is probably required to orientate accurately the directions or positions of the clock hands in the mind. It might be considered that the H group classified by the Flags was able to perform it well.

For the Gordon TVIC, the H group made less errors than the L group [$F(1,66)=2.80$, $.05<p<.10$], made higher estimates on the faithfulness to the instruction [$t(66)=1.68$, $.05<p<.10$], and on the imagery frequency [$t(66)=2.41$, $p<.05$]. It is generally supposed that this questionnaire test measures the controllability of imagery, i.e., the ability to imagine a series of varied scenes in turn according to the descriptions. In the clock task, if they perform faithfully to the instruction, subjects have to visualize two clock faces with hour and minute hands in turn, which might have some common requirement with the Gordon test. It must be an important finding that the H group classified by this subjective test could, from their point of view, obey more faithfully to the imagery instruction and imagined clock faces more frequently, and made actually less error responses in performance.

(d) For all combinations of the Betts QMI with four controllability tests, significant interactions were revealed on RTs 〔QMI-Flags, $F(12,408)=1.87$, $p<.05$; QMI-Space Relations, $F(12,404)=2.64$, $p<.01$; QMI-MPFB, $F(12,408)=2.46$, $p<.01$; QMI-TVIC, $F(12,408)=2.34$, $p<.01$〕. They all reflected the interaction which appeared between the H and the L groups for the QMI, that is, the vivid imagery group responded more quickly than the weak imagery group, especially at smaller size differences. The H-L groups in common responded most quickly and the L-H groups most slowly for the all combinations except the QMI-MPFB, but the L-L group responded as slowly as the L-H group for the QMI-Flags (Fig. 5).

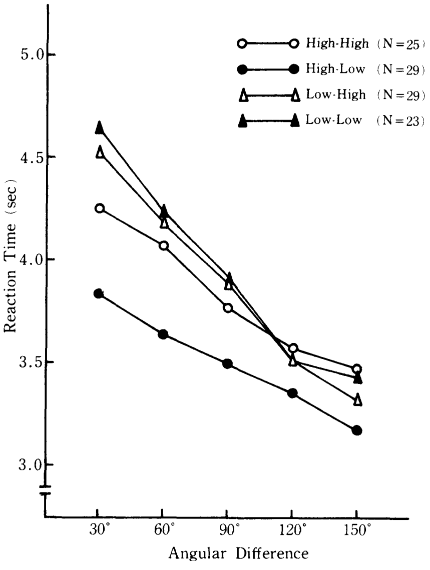

Fig. 5. Mean RT as a function of angular size difference for the four imagery groups on the Betts QMI-Flags in Experiment 2.

How can we explain this? We found that the H-L groups in common made the lowest mean estimates on all the four subjective measures, namely, faithfullness to the imagery instruction, clearness of the imagined clocks, imagery use frequency and confidence of correctness, and the L-H groups made the highest or second highest estimates, though a one-way analysis of variance revealed only a few significant differences, namely, for QMI-Flags on clearness 〔$F(3,102)=2.41$, $.05<p<.10$〕 and on frequency 〔$F(3,102)=2.76$, $p<.05$〕, and for QMI-Space Relations on clearness 〔$F(3,101)=2.20$, $.05<p<.10$〕 and on confidence 〔$F(3,101)=2.28$, $.05<p<.10$〕. On the total error responses, the H-L groups were commonly the most and the L-H groups were the least for the QMI-Flags and the QMI-TVIC, but both groups were not so for the QMI-Space Relations, though only one significant effect was revealed statistically, i.e., for the QMI-Flags 〔$F(3,102)=2.19$, $.05<p<.10$〕. It might be that the subjects who can visualize images vividly but cannot control them (in their respective senses of the tests) respond quickly to the mental size comparison task without recognizing the imagery process, and that the subjects who cannot visualize images vividly but can control them respond slowly and recognize the imagery process.

For the QMI-MPFB, the most quick reaction time was taken by the H-H group and the most slow one was by the L-L group. But that is not discussed here, since we could not find any relation of the MPFB itself to the task.

Conclusion

The present research examined how several imagery tests could predict the performance in two mental size comparison tasks, i.e., object comparison and clock comparison. The principal findings are as follows.

(a) In both tasks, a clear "symbolic distance effect" was revealed on reaction time, its dispersion and error responses, namely, they all decreased with the increase of size differences between the items.

(b) This effect, however, did not reflect the degree of subjective estimates on imagery use in both tasks. Probably the symbolic distance effect itself does not necessarily mean the imagery use in mental size comparison task.

(c) Subjects used imagery from their subjective point of view, since they gave mean estimates of about two thirds on the imagery use frequency in both tasks. According to the introspection, they judged directly without imagining when an item was known to have extremely large or small size or when it had been quite familiar to them for some reasons.

(d) The high estimators on imagery frequency responded more slowly in the clock task and made less errors in both tasks. Moreover, they gave higher estimates on all subjective measures than the low estimators. The subjective recognition concerning imagery use has a substantial relation to the performance in the mental size comparison task.

(e) In the clock task, men made a little more error responses at smaller angular differences and less at larger ones than women. Moreover, women estimated more moderately than men on the correctness in both tasks and on the clearness of images in the clock task.

(f) The Betts QMI was predictive of the response speed in both tasks, namely, the subjects with high vivid imagery responded more quickly, especially at small size differences. This finding suggests that imagery has an substantial function. Interestingly, the high vivid imagers estimated more highly than the low vivid imagers on imagery use frequency in the object task in which no imagery instruction had been given, but did not do so on the imagery use and also on the clearness of images in the clock task in which imagery instruction had been given. These findings might give a hypothesis that one of the factors underlying individual differences in imagery vividness can be due to the differences in the speed with which the image can be produced in some degree of clearness that makes it possible to be used.

(g) In both tasks, the subjects who got high scores on the Space Relations gave higher estimates on the correctness of their responses than the subjects who got low scores. In addition, in the clock task, they made higher estimates on the faithfulness to the imagery instruction. However, there were not any significant differences between the two groups in the actual performance. This test might predict the faithfulness to the imagery instruction and confidence in the performance.

(h) In the clock task, the subjects who got high scores on the Flags made less error responses. It might be said that this test predicted the ability to orientate accurately the directions or positions of the clock hands in the task.

(i) In the clock task, the subjects who estimated themselves as good at controlling visual images on the Gordon TVIC made higher estimates on the frequency of imagery use and made less errors in the performance. It might be that this questionnaire test predicted the manner of visualizing two clock faces one by one.

(j) In the clock task, for the QMI-Flags, for the QMI-Space Relations, and for the QMI-TVIC, the high vivid-low control subjects (in the respective senses of the tests) responded most quickly and the low vivid-high control subjects most slowly. It might be that, considering the subjective estimation data, the former subjects responded quickly to the task without recognizing the imagery process, and that the latter subjects responded slowly and recognized the imagery process.

(k) The MPFB, the VVIQ and the VVQ did not have any prominent predictive efficiency both in the objective performance and on the subjective estimates. The VVIQ should be expected to have predicted the tasks in the same manner as the Betts QMI, but it was not the case for some reason.

(l) The correlation coefficients between imagery tests showed low or no correlations with each other. Moreover, the predictive efficiency of each test in the two sorts of mental size comparison tasks were different from each other. Therefore, these tests would measure different aspects of imagery, if they would do, though they are all called and used as imagery tests.

(m) Individual differences in imagery ability played a more essential part in the clock task than in the object task, on the basis of predictive efficiency of the imagery tests.

References

Banks, W.P., & Flora, J. **1977** Semantic and perceptual processes in symbolic comparisons. *Journal of Experimental Psychology : Human Perception and Performance*, **3**, 278-290.

Bennet, G.K., Seashore, M.G., & Wesman, A.G. **1974** *Differential aptitude tests*. New York : Psychological Corporation.

Hatakeyama, T. **1981** Individual differences in imagery ability and mental rotation. *Tohoku Psychologica Folia*, **40**, 6-23.

Holyoak, K.J. **1977** The form of analog size information in memory. *Cognitive Psychology*, **9**, 31-51.

Holyoak, K.J., Dumais, S.T., & Moyer, R.S. **1979** Semantic association effects in a mental comparison task. *Memory and Cognition*, **7**, 303-313.

Jamieson, D.G., & Petrusic, W.M. **1975** Relational judgments with remembered stimuli. *Perception and Psychophysics*, **18**, 373-378.

Kirk, R.E. **1968** *Experimental design : Procedures for the behavioral sciences*. Belmont : Brook/Cole Publishing.

Kosslyn, S.M. **1980** *Image and mind*. Cambridge : Harvard University Press.

Kosslyn, S.M., Murphy, G.L., Bemesderfer, M.E., & Feinstein, K.J. **1977** Category and continuum in mental comparisons. *Journal of Experimental Psychology : General*, **106**, 341-375.

Likert, R., & Quasha, W.H. **1970** *Revised Minnesota paper form board test*. New York : Psychological Corporation.

Marks, D.F. **1973** Visual imagery differences in the recall of pictures. *British Journal of Psychology*, **64**, 17-24.

Moyer, R.S. **1973** Comparing objects in memory : Evidence suggesting an internal psychophysics.

Perception and Psychophysics, **13**, 180-184.

Moyer, R.S., & Bayer, R.H. **1976** Mental comparison and the symbolic distance effect. *Cognitive Psychology*, **8**, 228-246.

Paivio, A. **1975** Perceptual comparisons through the mind's eye. *Memory and Cognition*, **3**, 635-648.

Paivio, A. **1978**a Comparison of mental clocks. *Journal of Experimental Psychology : Human Perception and Performance*, **4**, 61-71.

Paivio, A. **1978**b Mental comparisons involving abstract attributes. *Memory and Cognition*, **6**, 199-208.

Paivio, A. **1978**c Imagery, language and semantic memory. *International Journal of Psycholinguistics*, **5**, 31-47.

Richardson, A. **1969** *Mental imagery*. London : Routledge & Kegan Paul.

Richardson, A. **1977** Verbalizer-visualizer : A cognitive style dimension. *Journal of Mental Imagery*, **1**, 109-125.

Richardson, J.T.E. **1979** Subjects' reports in mental comparisons. *Bulletin of the Psychonomic Society*, **14**, 371-372.

Thurstone, L.L., & Jeffrey, T.E. **1959** *Flags : A test of space thinking*. Chicago : Industrial Relations Center.

(Received November 30, 1984)