

Mental rotation of static and dynamic figures

GERALD E. LARSON

Navy Personnel Research and Development Center, San Diego, California

Previous studies comparing performance on standard (i.e., static) and dynamic spatial test items have concluded that the two item types measure different abilities. Such conclusions about the uniqueness of static and dynamic spatial abilities seem premature, however, since only a limited number of dynamic spatial tasks have been utilized in research and these have differed markedly from their static counterparts. In the present studies, tasks were designed to require a common mental operation (mental rotation) under static and dynamic conditions. Correlations between static and dynamic performance ranged from .80 to .90. This appears to suggest that the emergence of a unique dynamic ability factor depends on the utilization of certain specialized tasks (e.g., arrival time tasks) with mental operations much different than those required by conventional spatial tests. In other words, it is apparently the requirement for different cognitive processes and not the processing of stimulus motion per se that distinguishes performance on *some* dynamic tasks from performance on *some* standard static tasks.

Reviews of the literature on spatial ability (e.g., Cooper & Mumaw, 1985; Lohman, 1979; Pellegrino, Mumaw, & Shute, 1985) typically agree on most or all of the following conclusions: (1) A broadly defined spatial factor exists independently of verbal and quantitative factors, (2) spatial ability is an excellent predictor of performance in many occupations, but particularly in engineering and mechanical fields, and (3) while numerous spatial ability subfactors can be identified, these subfactors are interrelated and for the most part can be sorted into positions along a single *speed-power* or *simple-to-complex* continuum. At the speeded, cognitively simple end of the continuum are perceptual speed tests requiring only rapid visual comparisons of pictures. Tests requiring mental rotation of pictures are intermediate on the continuum, while tests requiring effortful mental transformations of detailed pictures define the continuum's complex or *power* end.

Since the findings cited above summarize much of the literature and evoke little controversy, it is tempting to conclude that spatial performance is adequately understood and documented (with the exception of specific controversies such as gender differences). Such a conclusion may be misleading, however, because much of the literature relies exclusively on a narrow set of paper-and-pencil test scores as data. This database is problematic, because, as Hunt and Pellegrino (1985) note, the paper-and-pencil format is a highly restrictive one for spatial tests. There is a poor correspondence between the dynamic, three-dimensional, colored world and the two-dimensional, static, black and

white printed test booklet. Thus, the existence of a consensus regarding spatial aptitude may stem more from the limitations of paper-and-pencil tests than from the accessibility of spatial cognition.

Some limitations of paper-and-pencil spatial tests have been demonstrated in studies by Hunt, Pellegrino, and their colleagues who, through use of computerized test presentations, have compared performance on standard and dynamic (i.e., animated) spatial test items (e.g., Hunt, Pellegrino, Frick, Farr, & Alderton, 1988; Pellegrino & Hunt, 1989). Their dynamic spatial tasks required that subjects remember the paths of moving objects, extrapolate motion from an observed path to an expected path, make judgments about the "arrival time" of one or more objects after being given a chance to observe initial speeds and trajectories, and "shoot down" an object traveling across the computer screen. Pellegrino and Hunt's analyses of their data indicated that the dynamic tasks tap processes that are correlated with, but not identical to, the processes or abilities underlying performance on the static tasks. The new dynamic spatial ability factor that they identified was largely related to dealing with relative visual motion, such as in the arrival time tasks.

Results from Pellegrino and Hunt (1989) appear to broaden our conceptualization of spatial ability by revealing one or more new dynamic abilities that are separate from standard spatial abilities. A still unanswered question, however, is whether a separate dynamic spatial factor will emerge for any set of tasks with stimulus motion, or whether evidence for a unique dynamic ability depends on the utilization of certain specialized tasks. For example, Pellegrino and Hunt's dynamic tasks required substantially different mental operations than the static tasks; the former required extrapolation of trajectories and estimation of arrival times, whereas the latter required mental operations such as rotation of images, cube folding, and so forth. It is not clear what results would have been obtained if standard and dynamic spatial tasks required essentially the

The opinions expressed are those of the author, are not official, and do not necessarily reflect the views of the Department of the Navy. The author wishes to thank David Alderton and Stephanie Booth-Kewley for their many helpful suggestions pertaining to this manuscript. Correspondence concerning this article should be addressed to G. E. Larson, Personnel Systems Department, Navy Personnel Research and Development Center, 53335 Ryne Road, San Diego, CA 92152-7250 (e-mail: glarson@nprdc.navy.mil).

same mental operations (such as mental rotation of images) and differed only in the existence of stimulus motion. Would a unique ability factor associated with dynamic stimuli still emerge, or would dynamic performance be fully predicted from static performance?

In the present study, a mental rotation task was modified to include both static items and items with two types of motion, and performance in the static condition was contrasted with performance in the dynamic conditions.

EXPERIMENT 1

Method

Subjects

Subjects were 693 randomly selected male Navy recruits (mean age, approximately 19 years old) tested on the 4th day of basic training at the Recruit Training Command, San Diego. The sample was 86% Caucasian, 8% African American, and 6% "other."

Instruments and Procedures

All subjects were tested in groups of approximately 40 individuals. During each session, half of the subjects were first administered two paper-and-pencil tests (Spatial Relations and Ravens Advanced Progressive Matrices) while the remaining subjects solved mental rotation items on a microcomputer. After about 50 min, the two groups traded places and completed the remaining tests. Group assignment was random and was not retained. The tests in the battery are described below.

Raven's Advanced Progressive Matrices (APM) test. The APM is a 36-item (paper-and-pencil) nonverbal test of general intelligence (Raven, 1962), administered with the standard 40-min time limit.

Primary Mental Abilities (PMA) Spatial Relations test. The PMA test (Thurstone, 1965) requires picture plane mental rotation. Thirty five-alternative (paper-and-pencil) items were administered with a 7-min time limit; each alternative must be evaluated to determine whether it is the *standard* figure in a new orientation or whether it is a mirror image. Individuals must identify test figures that match the standard; there are 72 matches (maximum score) and 78 non-matches.

Computerized mental rotation task. This task was a variation of the rotation task described in Alderton and Larson (1994). Given side-by-side asymmetric polygons presented on the video monitor, examinees had to determine whether the two figures were identical except for a picture plane rotation (response = "same") or rotated and mirror reversed (response = "different"). Figure 1a shows two items requiring a *same* response, and Figure 1b depicts two items requiring a *different* response. To make this determination, individu-

als mentally rotate the right-hand (or *test*) figure into congruence with the standard. The instructions told subjects that they would see pairs of shapes that would match unless one was drawn backward, and that both speed and accuracy of responding were important for their test scores. Several practice items were administered, and accuracy feedback was provided after each practice response. For each item in the test, accuracy and response time (in milliseconds) were recorded for analysis.

Test design. Seventy-two *same* response items were created from 12 asymmetric polygons or *standards* (four are shown in Figure 1). To create an item, each standard was paired with itself, and the right-hand copy, or test figure, was rotated clockwise 30°, 70°, 110°, or 150°. One fourth of the items were presented at each orientation, and some items were repeated (after numerous interpolated items) to achieve the desired test length. This procedure generated 72 *same* trial items. The set of 72 *different* items were created from the *same* response items by mirror reflecting the test figure following each rotation. This procedure yields a set of 144 total items.

Dynamic versus static item formats. One third of the test items were presented in each of the three formats shown below. (Item formats were randomly intermixed during the session.) For each format: (1) half of the items were *same* and half were *different*, and (2) one fourth of the test figures were shown at each orientation.

1. Standard presentations. Two static pictures were shown side by side.

2. Moving background. Each member of the side-by-side figure pair was shown against a background of 10 progressively sized concentric squares. The three largest squares completely enclosed the figures while the smaller squares overlapped the outlines of the figures. The sides of the squares were drawn with dots rather than solid lines. To create the illusion of continuous background motion, the dots with which each square's sides were drawn flowed in a clockwise direction around the square's perimeter.

3. Moving figures. The two members of each figure pair traveled in two adjacent circular paths. The effect was similar to having the figures ride in the buckets of adjacent ferris wheels, particularly since the figures maintained a constant orientation to the ground (i.e., they did not turn upside down). Each figure completed its circular path every 5 sec.

The computer-presented rotation tests were group-administered on Hewlett-Packard Integral microcomputers with a simplified keyboard: a plastic mask covered all but the designated response keys and a key labeled HELP that could be pressed during testing to suspend the program and request assistance. The "S," "F," "H," "K," and ";" keys were relabeled as "A," "B," "C," "D," and "E." The space bar was relabeled ENTER. The numeric keypad keys retained their meanings. The computers operate under UNIX and the tests were written in standard C.

Results

Preliminary analyses were conducted in order to ensure that all subjects had maintained an acceptable accuracy level on the computerized mental rotation task. Only 5 subjects in the sample of 693 had an overall accuracy rate of less than 90%, and these subjects were subsequently dropped from the analyses that follow.¹

Descriptive statistics for the various tasks are shown in Table 1. The average Raven Progressive Matrices score of 19.65 indicates that the approximate mean IQ for the sample was 105 on the basis of Raven/IQ conversion tables published in Jensen, Saccuzzo, and Larson (1988). The subjects were thus very close to a normal population "average" in terms of intelligence. Aggregate performance variables for the computer-presented rotation task are shown in the top half of the table; performance as a function of

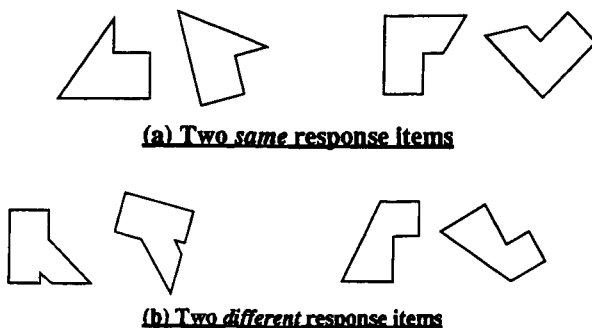


Figure 1. Examples of Mental Rotation Test items: (a) 75° *same* (left) and 225° *same* (right); (b) 285° *different* (left) and 135° *different* (right).

Table 1
Descriptive Statistics for Performance Scores

| | <i>M</i> | <i>SD</i> |
|---------------------|----------|-----------|
| Raven | 19.65 | 5.28 |
| Space relations | 39.82 | 13.70 |
| True RT | 2.28 | .70 |
| False RT | 2.79 | .94 |
| Accuracy (rotation) | 136.48 | 9.17 |
| By Item Type | | |
| Standard true RT | 2.13 | .74 |
| Standard false RT | 2.58 | .96 |
| Standard accuracy | 46.34 | 1.42 |
| Circling true RT | 2.19 | .69 |
| Circling false RT | 2.86 | .99 |
| Circling accuracy | 45.97 | 1.59 |
| Background true RT | 2.53 | .82 |
| Background false RT | 2.94 | 1.02 |
| Background accuracy | 46.23 | 1.52 |

Note—RT, response time.

item format is summarized in the bottom half of the table. For all of these analyses (and others to be described below) only the subjects' correct responses were analyzed. As Table 1 shows, average responses to *true* items were significantly faster than average responses to *false* items [$t(1,671) = 26.50, p < .01$]; this finding is typical for mental rotation data (see, e.g., Cooper & Shepard, 1973).

To analyze overall response time as a function of item format, responses were sorted into the three format types (*standard*, *circling*, or *background motion*) and then further sorted into *true* or *false* categories within each format type. Average *true* and *false* response time for each condition is shown in the bottom half of Table 1. Only *true* items, however, will be included in the analyses that follow, since *false* item data tend to be less reliable and are therefore often set aside in mental rotation studies (Alderton & Larson, 1994; Pellegrino & Kail, 1982). For *true* items, a repeated measures analysis of variance (ANOVA) revealed a main effect of item type [$F(2,1364) = 348.18, p < .01$]. Orthogonal contrasts indicated that responses to

standard items were significantly faster than responses to *circling* [$F(1,682) = 106.84, p < .01$] items and items with *background motion* [$F(1,682) = 526.87, p < .01$]; moreover, responses to *circling* items were significantly faster than responses to items shown against a moving background [$F(1,682) = 418.27, p < .01$].

Effect of Item Type on Rate of Mental Rotation

In rotation paradigms, response time (RT) typically increases in direct relation to angle of stimulus presentation. Consequently, it is possible to plot a regression line showing RT as a function of angle and to interpret the slope of the regression line as an index of rotation rate. Figure 2 shows the full-sample regression lines (i.e., using group-level data) for the three item types used in the present experiment. When slopes were calculated for each individual subject under each condition, a repeated measures ANOVA revealed a main effect of item type on RT slopes for individual subjects [$F(2,1366) = 72.52, p < .01$]. Orthogonal contrasts indicated that slopes for *circling* items were significantly different from slopes for either *standard* [$F(1,683) = 130.39, p < .01$] items or items with *background motion* [$F(1,683) = 121.74, p < .01$]; however, slopes for *standard* and *moving background* items were not significantly different from each other. Thus, if one interprets slope to be an index of mental rotation rate, then rotation rate with dynamic presentations seems to differ from the traditional static presentations only when stimulus motion per se (rather than background motion) is involved. Given that RTs to *moving background* items are nevertheless slower than RTs to *standard* items despite comparable slopes, this particular difference can probably be attributed to slowed encoding with background motion rather than slowed rotation.

Relationships Between Static and Dynamic Abilities

The preceding analyses indicate that RTs to rotation items are significantly lengthened by item motion. The present section addresses the question of whether perfor-

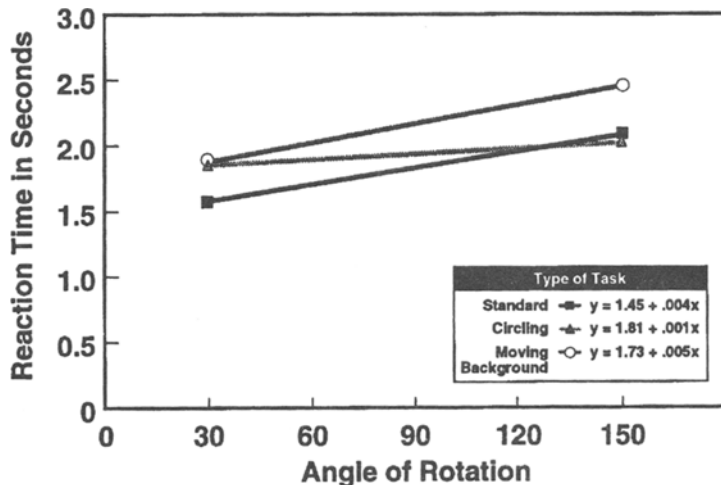


Figure 2. Mental rotation functions for three item types.

Table 4
Correlations for Experiment 2

| | 1 | 2 | 3 | 4 |
|--------------------|------|------|------|------|
| Standard RT | | | | |
| Circling RT | .90 | | | |
| Backmove RT | .88 | .86 | | |
| Rocking RT | .86 | .88 | .83 | |
| Assembling objects | -.26 | -.29 | -.28 | -.30 |

Note—All correlations are significant at the .01 level.

sults indicate quite strongly that the ability to deal with moving elements and dynamic spatial relations is separate from the abilities associated with reasoning about static spatial information" (p. 196). Pellegrino and Hunt go on to argue that dynamic spatial abilities are not only theoretically interesting but also practically useful, since they may provide improved prediction of dynamic real-world jobs such as machinery operation and piloting.

Such unambiguous conclusions about the uniqueness of static and dynamic spatial abilities seem premature, however, since only a limited number of dynamic spatial tasks have been utilized in research, and the dynamic tasks that have been used differ markedly from their static counterparts. For example, Pellegrino and Hunt's (1989) dynamic tasks required extrapolation of trajectories and estimation of arrival times, whereas their static tasks required mental operations such as rotation of images, cube folding, and so forth. Their newly discovered dynamic ability factor was largely related to dealing with relative visual motion, as in the task where subjects must make judgments about the "arrival time" of one or more objects after being given a chance to observe initial speeds and trajectories. The uniqueness of such tasks seems so pronounced, however, that some statistical independence from static tasks seems almost guaranteed and need not be treated as proof that the entire domain of dynamic performance abilities is separate from the domain of static performance abilities. This point is illustrated by the fact that even Pellegrino and Hunt reported a number of sizable intercorrelations between their static and dynamic tasks.

In the present studies, tasks were designed to require a common mental operation (mental rotation) under static and dynamic conditions. While the effects of item format on RT and slope were inconsistent across the two experiments (perhaps owing to different ability levels between military and college samples), strong and consistent evidence was found for a common rotation ability in all conditions. Static and dynamic tasks had nearly identical correlations with other spatial and reasoning tests along with extremely high intercorrelations among themselves. Indeed, all intercorrelations exceeded .80, and general factors accounting for 64% and 69% of the total variance in the combined set of rotation scores were found in Experiments 1 and 2, respectively. This appears to suggest that the finding of a unique dynamic ability factor depends on the utilization of certain tasks (e.g., arrival time tasks) with mental operations much different from those required by conventional spatial tests. In other words, it is apparently

the requirement for different cognitive processes and not the processing of stimulus motion per se that distinguishes performance on *some* dynamic tasks from performance on *some* standard static tasks. Hopefully, follow-on research will further expand our knowledge of the task characteristics that support ability differentiation arguments from the characteristics that do not. Also, follow-on studies should systematically address the specific cognitive processes that account for gender differences in dynamic spatial performance (Law, Pellegrino, & Hunt, 1993). The vast majority of subjects in the present research were male, precluding any analysis based on gender.

One final issue concerns the ecological validity of the dynamic tasks used in previous research and the rotation tasks administered in the present studies. Both types of tasks have important real-world counterparts and may therefore be of considerable practical interest. Arrival time tasks, for example, would seem to tap processes related to air traffic control, driving a car, and firing at moving targets. Real-world counterparts of dynamic mental rotation include utilization of visual landmarks during aircraft navigation and, to a lesser extent, navigation of surface craft. For example, a confused pilot who uses landmarks to orient during flight could find him-/herself mentally transforming the landscape to achieve different perspectives in conjunction with map consultation. Some of these landscape transformations may well involve rotations on scenes made dynamic by observer motion. In other cases, mental rotations are attempted on objects that are themselves in motion. One example is the attempt by a military pilot to identify other flying aircraft (as friend or foe) on the basis of mental comparison of viewed silhouettes with memorized aircraft profiles. Such comparisons are likely to require some mental manipulation of viewed aircraft, including mental rotation. The present research suggests that traditional static spatial tests may adequately predict performance in dynamic settings, as long as common mental operations are involved in both the test and the outcome measure.

REFERENCES

- ALDERTON, D. L., & LARSON, G. E. (1994). Cross-task consistency in strategy use and the relationship with intelligence. *Intelligence*, *18*, 47-76.
- COOPER, L. A., & MUMAW, R. J. (1985). Spatial aptitude. In R. F. Dillon (Ed.), *Individual differences in cognition* (Vol. 2, pp. 67-94). New York: Academic Press.
- COOPER, L. A., & SHEPARD, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 75-176). New York: Academic Press.
- HUNT, E., & PELLEGRINO, J. (1985). Using interactive computing to expand intelligence testing: A critique and prospectus. *Intelligence*, *9*, 207-236.
- HUNT, E., PELLEGRINO, J., FRICK, R. W., FARR, S. A., & ALDERTON, D. (1988). The ability to reason about movement in the visual field. *Intelligence*, *12*, 77-100.
- JENSEN, A. R., SACCUZZO, D. P., & LARSON, G. E. (1988). Equating the standard and advanced forms of the Raven Progressive Matrices. *Educational & Psychological Measurement*, *48*, 1091-1095.
- LAW, D. J., PELLEGRINO, J. W., & HUNT, E. B. (1993). Comparing the tortoise and the hare: Gender differences and experience in dynamic spatial reasoning tasks. *Psychological Science*, *4*, 35-40.

- LOHMAN, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature* (Tech. Rep. No. 8). Stanford University: School of Education Aptitude Research Project.
- PELLEGRINO, J. W., & HUNT, E. B. (1989). Computer-controlled assessment of static and dynamic spatial reasoning. In R. F. Dillon and J. W. Pellegrino (Eds.), *Testing: Theoretical and applied perspectives* (pp. 174-198). New York: Praeger.
- PELLEGRINO, J. W., & KAIL, R. (1982). Process analyses of spatial aptitude. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 311-366). Hillsdale, NJ: Erlbaum.
- PELLEGRINO, J. W., MUMAW, R. J., & SHUTE, V. J. (1985). Analyses of spatial aptitude and expertise. In S. E. Embretson (Ed.), *Test design: Developments in psychology and psychometrics* (pp. 45-76). New York: Academic Press.
- RAVEN, J. C. (1962). *Advanced progressive matrices*. London: H. K. Lewis.
- THURSTONE, T. G. (1965). *Primary mental abilities: Technical report*. Chicago: SRA.

NOTE

1. Since each individual response time (RT) variable was also screened for 90% accuracy, *N*s vary slightly across analyses.

(Manuscript received April 22, 1994;
revision accepted for publication July 19, 1995.)