

Methods of evaluating performance on spatial memory tasks

MATTHEW J. SHARPS and EUGENE S. GOLLIN
University of Colorado, Boulder, Colorado

Methods of evaluating adult spatial memory performance were examined in two spatial memory tasks. Displacement scores were shown to exhibit mathematical difficulties. A comparison was made of several measures of spatial memory performance that are mathematically adequate for parametric analysis. These were found to correlate with each other at very high levels. It was concluded that, provided the measure used in any given study of spatial memory satisfies the assumptions on which the statistics to be used are predicated, a variety of satisfactory measures of spatial memory can be employed, and that the criterion as to which should be used may depend primarily upon considerations germane to the specific research problem at hand.

Memory for spatial locations has become a focus of recent study (e.g., McCormack, 1982; Moore, Richards, & Hood, 1984; Park, Puglisi, & Lutz, 1982; Waddell & Rogoff, 1981). In a spatial memory task, a subject generally learns where items are located in a particular space, which might be a map, a three-dimensional array of some kind, or a larger scale environment such as a room, and then is asked to indicate where he or she remembers the items to have been. A variety of measures has been used to evaluate spatial memory performance. However, none have proven entirely satisfactory.

One popular measure has been the displacement score (e.g., Pezdek, 1983). This score is obtained by measuring the distance between the position occupied by the to-be-remembered item during study and the locus of placement by a respondent during recall. For example, if a subject remembers Item A as having been three units (feet, centimeters, squares on a chess board, etc.) from its actual position, and Item B as having been eight units from its actual position, then that subject's displacement score for Items A and B is the sum of the displacement units for A and B, $3+8=11$. Unfortunately, this measure is not completely compatible with the more commonly used statistical methods. Let us consider a 25-item spatial memory task. At recall, Subject I places each of the 25 items one unit from its correct position. Subject II places 20 items correctly, but places each of the last 5 items five units from its correct location. Each of these subjects receives a displacement score of 25. However, one is hard-pressed to demonstrate that placing each of

5 items five units from veridicality is equivalent in any meaningful sense to displacing each of 25 items by one unit. Therefore, displacement scores cannot be said to be metric in nature.

One of the basic requirements of the analysis of variance (ANOVA), the statistical method used in most recent studies of spatial memory (e.g., Light & Zelinski, 1983; Pezdek, 1983), is that the data be metric; data must be interval or ratio as opposed to ordinal or nominal. Since displacement scores cannot be demonstrated to possess the characteristics of interval or ratio data, ANOVA, applied to raw displacement scores, is potentially troublesome. In using the ANOVA on raw displacement scores, one is forced to trust in the relative insensitivity of ANOVA to mathematical departure from its basic assumptions. This could prove hazardous, because the robustness of the ANOVA is not infinitely elastic, and there is always a risk in relying on the untested assumption that any given analysis of displacement scores does not strain the mathematical plasticity of ANOVA. One could, of course, employ a variety of data transformations to ensure distributional compatibility with the ANOVA. However, if untransformed scores are of interest, the only viable alternative is to resort to the use of nonparametric methods, with the corresponding loss in statistical power that this generally entails (Winer, 1971).

Another method of evaluating spatial memory performance is to employ several kinds of scores possessing construct validity for the task, and then to make statistical comparisons among them (e.g., Light & Zelinski, 1983). The advantage of this approach is that one can choose ratio or interval measurement scales that are amenable to the ANOVA. However, if one employs separate univariate analyses to make statistical comparisons with more than one measure, the probability of erroneously finding a statistically significant effect mounts alarmingly (Harris, 1975). Multivariate methods are of course the treatments of choice in such cases, and are widely recommended in cases where several variables are to be compared (e.g.,

This research was supported in part by a grant from the Intermountain West Long-Term Gerontology Center; that support is gratefully acknowledged. The authors also wish to thank Phyllis Parrish and the staff and residents of the Golden West Manor, Boulder, Colorado, and Sue Erwin and the volunteers and staff of the Retired Senior Volunteer Program, for their help and participation in this research. Requests for reprints should be sent to E. S. Gollin, Department of Psychology, Campus Box 345, University of Colorado, Boulder, CO 80309.

Harris, 1975; Kerlinger & Pedhazur, 1973). Hummel and Sligo (1971) advocated a multivariate analysis of data followed by separate univariate analyses as the optimal analytic scheme in cases of multiple measures. However, multivariate methods pose several problems of a practical nature. In isolation, multivariate measures of statistical significance have been found to be very conservative (Hummel & Sligo, 1971). Multivariate analyses also require considerable quantities of computer time, as well as excursions into matrix algebra and relatively complex methods of interpretation. A simple univariate measure of spatial memory performance, if a highly representative one could be found, would be simpler, less costly, and generally more convenient. In the present study, we evaluated several measures which possess good construct validity for the evaluation of spatial memory, and which do not violate the mathematical requirements of the ANOVA.

METHOD

Subjects

Sixteen young adults with a mean age of 21 years (range 18-35 years) were recruited from the University of Colorado subject pool. Sixteen active, healthy elderly adults (mean age 78 years, range 68-84) also were recruited. Two age groups were employed to determine the generality of findings across age.

Materials

Two tasks, one large-scale and the other small-scale, were employed; this was intended to allow evaluation of the general applicability of potential results in markedly different task situations. One of these tasks employed an "environment" room in which subjects in the "environment" condition studied object locations. The environment was constructed in a 7 x 9.75-m classroom. File cabinets, piles of boards, tables, large equipment-storage lockers, and piles of carpet were organized into structures and placed around the room. Thirty small, common objects were placed in the room, on or near the structures.

Four sets of white cardboard cards, 4-cm square in size, were prepared. There were 30 cards in each set, each card bearing the name of one of the 30 stimulus items. These cards were used to test spatial memory in the environment condition.

A 1/6 scale schematic map, hereafter called the "study map," was drawn of the environment. A 1/17 scale map (the "test map") was also drawn and reproduced on 21.6 x 27.94-cm paper. The study map served as the stimulus array in the second task, the small-scale map task, and the test map was employed in testing spatial memory in this task.

Procedure

Subjects were tested for visual acuity. In the environment study condition, one of two presentation conditions (there were 8 young and 8 elderly subjects in each presentation condition), subjects were told that they would be taken into a room and shown a number of common objects. They were told to remember where the objects were placed in the room. The experimenter then led the subjects, 4 at a time, on a circuitous tour of the environment [participation in groups, as opposed to individual participation, made no difference to performance within this paradigm (Sharps & Gollin, 1986)]. The experimenter stopped at each of the 30 locations, pointed to the object at the given location and named it, waited 5 sec, and continued to the next item. At the conclusion of the tour, the subjects were led out of the room. Meanwhile, all of the to-be-remembered objects were cleared out of the room, and the subjects were led back to the room 2 min after having left it. They were then told that they would be given one card at a time, that the card would bear the name of one of the objects they had seen, and that they were to place each card, with the name facing down (to prevent other sub-

jects from seeing it) at the location where they had seen the object named on the card. On the back of each card in a subject's deck was a deck number (1 through 4) so that the responses of individual subjects could be scored appropriately. Four different orders of cards were used, to prevent subject convergence on the same spatial location at the same time.

In the map study condition, the other presentation condition, groups of subjects were informed that they would see a map on which a number of objects were pictured, and that they were to remember where these objects were located (the instructions were as identical as possible to those used in the environment condition). Subjects were then led into a room in which the 1/6 scale study map was displayed on a table. The experimenter pointed to the drawings of objects on the map in the same order used to point out the objects in the environment, naming them and waiting 5 sec between item presentations. (Previous research indicated that there was no intrinsic difference in memorability between the objects of the environment condition and the drawings of the map condition.) At the conclusion of the map study period, the study map was removed. Subjects began the spatial memory task 2 min after the conclusion of the presentation. Each participant was given a 1/17 scale test map, identical to the study map except in size. The experimenter then presented one of the sets of cards, one card at a time, to each subject. The subjects placed an "X" on the map in the precise location where they thought the object named on a given card had been, and then wrote the name of the item believed to have been there beside the "X." Five seconds elapsed between the presentation of consecutive items. After this procedure, the test maps were collected.

Four kinds of memory scores were analyzed. The first of these, the target score, was a measure of precise memory for object location. If an item was placed within 30 cm of its veridical position in the environment, or within 1.8 cm of its correct place on the map (these distances were to scale relative to one another), it was counted as correct. A simple count of these target placements for each subject constituted the target score.

The second most precise measure, the first quadrant measure, counted those item placements that occurred in the correct 1/4 section of the environment or the map. The least precise measure, the second quadrant score, counted those placements that were within the appropriate quarter of the map or environment.

The fourth measure taken was the structure score. This was based on the idea that subjects might use a structure, as discussed above, as a cue or group of cues to assist in remembering where an item located on or near it might be. If an item was placed by a subject on the correct structure, it was counted as a point on the subject's structure score, regardless of the quadrant in which it occurred; several structures overlapped into more than one quadrant.

RESULTS

The effects of age and task type (map or environment) on memory performance are reported elsewhere (Sharps & Gollin, 1986). Pearson product-moment correlations were applied to the four measures of spatial memory described above (see Table 1). All correlations between the four measures were significant (*p* ≤ .001). No effects of age, task type, or gender were significant on this

Table 1
Pearson Correlation Coefficients of the Four Measures of Spatial Memory

| | Score | | |
|--------------------|--------|--------------|--------------|
| | Target | 1st Quadrant | 2nd Quadrant |
| 1st Quadrant Score | .9719 | — | — |
| 2nd Quadrant Score | .9422 | .9748 | — |
| Structure Score | .9937 | .9711 | .9338 |

Note—All correlations were significant: *p* ≤ .001.

pattern, and within age group, task type, and gender, all correlations among the four measures remained significant ($p \leq .001$).

DISCUSSION

Precise measures of spatial memory, less precise measures, and measures based on association of to-be-remembered items with cues in the space under study correlate very highly with one another. Task type, age, and gender have no detectable effects on the pattern of correlation reported, indicating that any of the four kinds of scores examined would be an adequate measure of spatial memory, in that all four measures, having apparently high construct validity, also possess high concurrent validity as demonstrated by their strong intercorrelation. Although, for typical research purposes, the target score would seem to be the most convenient and useful measure of the four, any of these measures might therefore be used representatively as a mathematically tenable univariate measure of performance in any of the usual task frameworks employed in the analysis of spatial memory. Within task frameworks which differ radically from the typical kind of task discussed, however, an adaptation of the recommendations of Hummel and Sligo (1971) appears advisable. In such cases, multivariate and correlational analyses of the measures employed, followed by univariate analysis of the target score or other score specifically germane to the problem at hand, should minimize errors arising from misapplication of scoring procedures to the evaluation of spatial memory.

REFERENCES

- HARRIS, R. J. (1975). *A primer of multivariate statistics*. New York: Academic Press.
- HUMMEL, T. J., & SLIGO, J. R. (1971). Empirical comparison of univariate and multivariate analysis of variance procedures. *Psychological Bulletin*, *76*, 49-57.
- KERLINGER, F., & PEDHAZUR, E. J. (1973). *Multiple regression in behavioral research*. New York: Holt, Rinehart & Winston.
- LIGHT, L. L., & ZELINSKI, E. M. (1983). Memory for spatial information in old and young adults. *Developmental Psychology*, *19*, 901-906.
- MCCORMACK, P. D. (1982). Coding of spatial information by young and elderly adults. *Journal of Gerontology*, *37*, 80-86.
- MOORE, T. E., RICHARDS, B., & HOOD, J. (1984). Aging and the coding of spatial information. *Journal of Gerontology*, *39*, 210-212.
- PARK, D. C., PUGLISI, J., & LUTZ, R. (1982). Spatial memory in older adults: Effects of intentionality. *Journal of Gerontology*, *37*, 330-335.
- PEZDEK, K. (1983). Memory for items and their spatial locations by young and elderly adults. *Developmental Psychology*, *19*, 895-900.
- SHARPS, M. J., & GOLLIN, E. S. (1986). *Memory for spatial information in young and elderly adults: The role of stimulus factors*. Submitted for publication.
- WADDELL, K. J., & ROGOFF, B. (1981). Effect of contextual organization on spatial memory of middle-aged and older women. *Developmental Psychology*, *17*, 878-885.
- WINER, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.

(Manuscript received for publication August 30, 1985.)