Technical Commentary

USING CLUSTER ANALYSIS TO DISCOVER AND CHARACTERIZE COVERT STRATEGIES

Shui-I Shih[1] and George Sperling[2]

[1]*Saint Anselm College and* [2]*University of California, Irvine*

The present article focuses on a particularly useful application of cluster analysis: examining practice effects in psychological experiments, particularly with respect to individual differences in performance over sessions. The factor (e.g., practice) whose effect we wish to assess or to cancel is a *secondary variable*. Our aim is to determine whether the data that have been collected in an experiment are truly homogeneous with respect to a secondary variable and, if not, to obtain subsets of homogeneous data as suggested by cluster analysis. The homogeneous subsets can then be used to characterize the modes of performance.

The effects of a secondary variable, such as performance improvements over sessions, typically have been assessed by an analysis of variance (ANOVA). ANOVA answers such questions as "Are session-to-session variations in the dependent variable (typically accuracy or reaction time) due to chance, or is there some variation that can be attributed to performance changes between sessions?" However, there are two problems with ANOVA:

- **Multiple dependent variables:** ANOVA typically deals with a single dependent variable. We would like to be able to characterize performance changes that are more subtle than simply changes in either accuracy or speed. When the subject makes a complex response, as in a recall experiment, there may be a systematic change in the pattern of the recalled items. Even in experiments in which only accuracy and speed are measured, there may be systematic dependencies between speed and accuracy that are not captured by treating them independently.

- **Nature of variation:** Having determined that there might be session-to-session variation, we face a problem in determining just what this variation might be. That is, we would like to form relatively homogeneous subgroups of sessions and to compare performance between subgroups. ANOVA is not well adapted for the grouping task. In contrast, cluster analysis thrives on multiple dependent variables and is designed to produce homogeneous clusters.

Address correspondence to George Sperling, Human Information Processing Laboratory, Department of Cognitive Science, SST-6, University of California, Irvine, CA 92717; e-mail: sperling@uci.edu.

Whereas clustering methods typically do not offer useful tests of statistical significance, this is the forte of ANOVA. Therefore, it usually is advisable to use both a statistical model (ANOVA) and an analytic tool (cluster analysis). We now illustrate the value of applying these two tools in revealing and analyzing a practice effect in a short-term memory experiment, which was performed in the context of a large study of attentional shifts.

PROCEDURE

The experiment combined a partial-report procedure with rapid serial visual presentation (RSVP). The subject saw a stream of 3×3 letter arrays. A tonal cue, indicating one of the three rows, was presented at the onset of one of the arrays in the sequence. The subject was required to report the three letters that were (a) from the indicated row, i; (b) in the correct order (i.e., in the correct columns), j; and (c) from the earliest possible letter array, k, simultaneous with or subsequent to the onset of the tonal cue. Two subjects were run for 40 and 34 sessions. The performance for each session is characterized by the recall probability, $P(i, j, k; n)$, where i, j, k, and n indicate, respectively, row, column, array, and session; i and j both range from 1 to 3, k ranges from 1 to 6, and n ranges from 1 to 40 or 1 to 34. A total of 54 recall probabilities is required to describe the performance over the 54 spatiotemporal locations for each session.

ANOVA

An unreplicated two-way ANOVA (Table 1) showed a main effect for spatiotemporal location, but no main effect for session for either subject. However, a Tukey's test indicated highly significant interactions of session and spatiotemporal location for both subjects ($F[1, 2{,}066] = 176.69, p < .0001$, for one subject; $F[1, 1{,}748] = 26.96$, $p < .001$, for the other), indicating that there were complex, essentially unanalyzable, changes in performance over sessions.

CLUSTER ANALYSIS

Cursory inspection of the original data suggested strong systematic differences as a function of practice. The FASTCLUS procedure (SAS Institute, 1985) was applied to analyze the performance variation with respect

Table 1. *Unreplicated ANOVA for effects of session and spatiotemporal location*

Source of variation	*SS*	*df*	*MS*	*F*	*p*
	Subject S.S.				
Spatiotemporal location	47.5627	53	0.8974	83.3011	.0000
Session	0.5335	39	0.0137	1.2697	.1234
Residual	22.2680	2,067	0.0108		
	Subject J.S.				
Spatiotemporal location	8.5397	53	0.1611	29.8913	.0000
Session	0.1974	33	0.0060	1.1095	.3070
Residual	9.4278	1,749	0.0054		

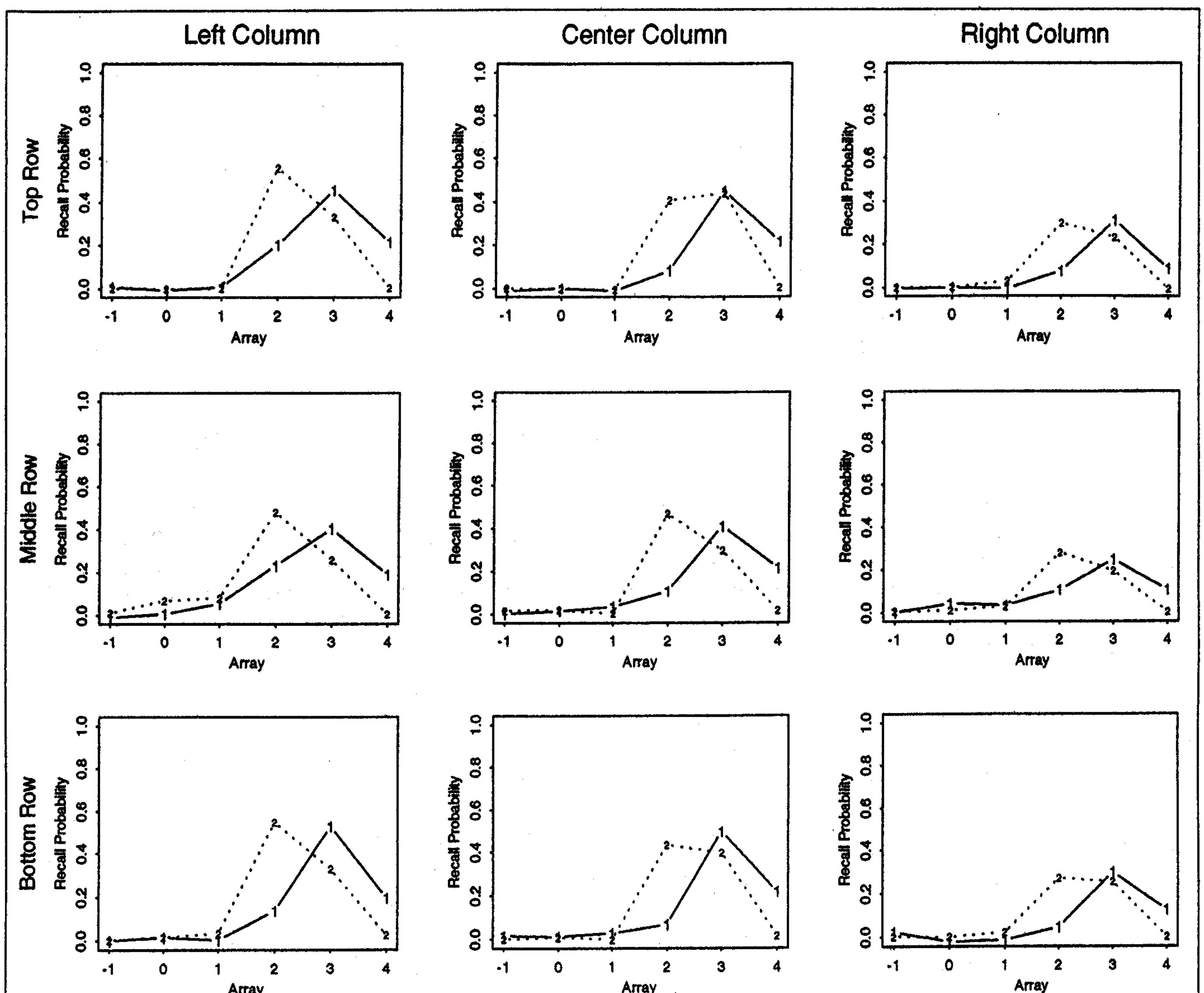

Fig. 1. The 54 recall probabilities of the experiment (9 panels × 6 points per panel). Data for subject S.S. The ordinate represents the recall probability; the abscissa represents the sequence number of an array. Each panel represents a different possible spatial location of reported items. Labels "1" and "2" indicate, respectively, Clusters 1 and 2 from the analysis.

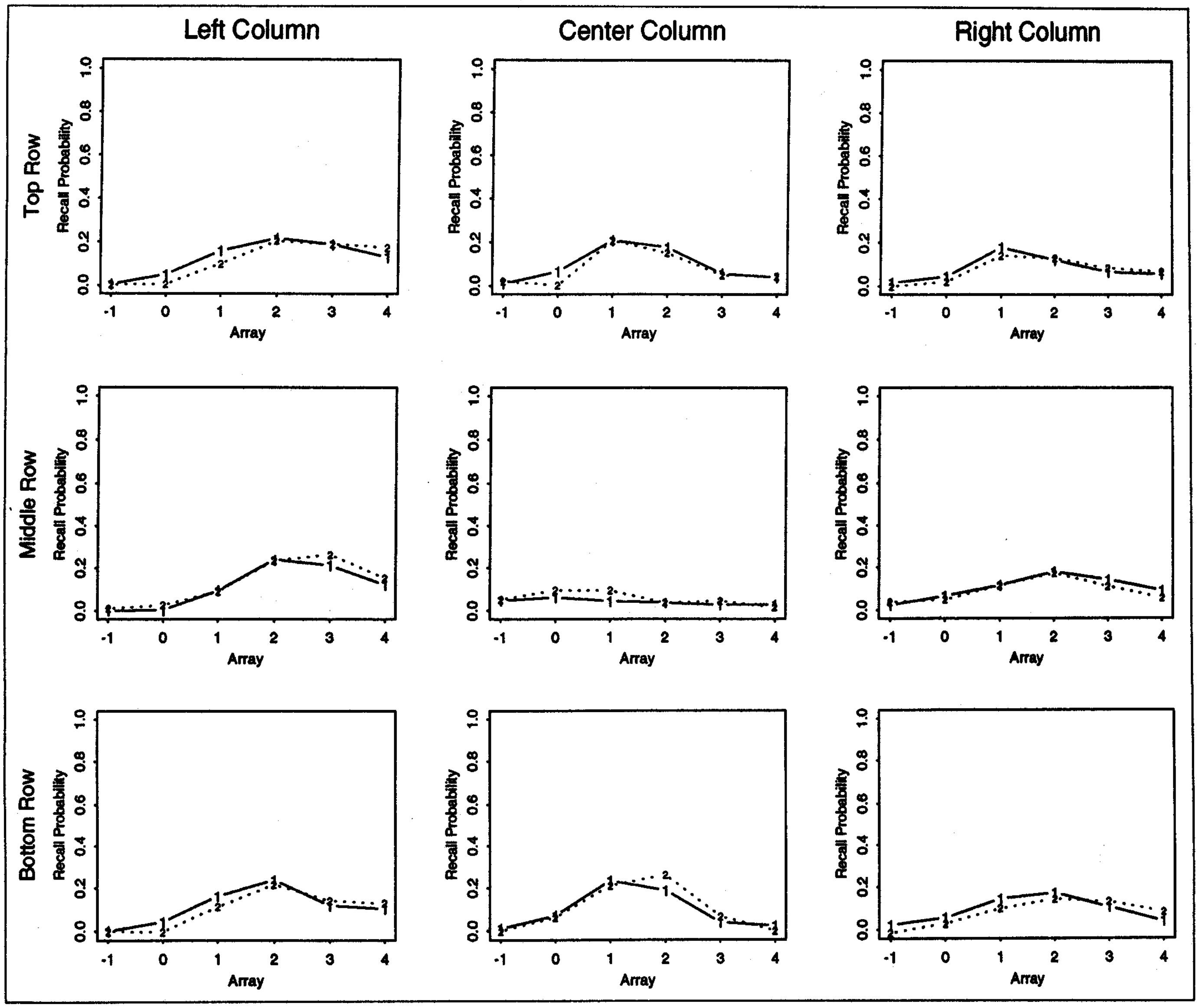

Fig. 2. Data for subject J.S. Details as in Figure 1.

to session number. This procedure combines sessions into a specified number of clusters so that the sum of squared standardized euclidean distances[1] of points from the means of their clusters is minimized. We ran this procedure for several choices of the number of clusters. For each resulting cluster c, we obtained the mean recall probability function, $\overline{P}_c(i, j, k)$. Because we do not have a statistical procedure to decide how many different clusters to accept, the number of partitions was determined by visual inspection of the $\overline{P}_c(i, j, k)$.

1. The standardized euclidean distance is the euclidean distance between each data point and the mean value for its cluster, divided by the standard deviation of the whole population.

RESULTS AND DISCUSSION

Figures 1 and 2 present the mean recall probability functions, given a choice of two clusters for each subject. Labels "1" and "2" indicate results for Clusters 1 and 2. Each of the nine panels in a figure corresponds to a spatial location (i, j), where $\overline{P}_c(i, j, k)$ is plotted against array k. The arrays are ranked relative to the onset of the tonal cue. The array presented simultaneously with the tonal cue is coded as Array 0.

Figure 1 shows clearly that, for each spatial location (i, j), Cluster 2 peaks one array earlier than Cluster 1. In other words, this particular subject could report items from earlier arrays in the sessions belonging to Cluster 2 that in the sessions belonging to Cluster 1. This performance difference is attributed to the practice effect because Cluster 1 consists of the 2nd through 9th sessions, whereas Cluster 2 consists of the 1st and the 10th through 40th sessions.

For the other subject (Fig. 2), there is little difference in performance between Clusters 1 and 2. In other words, the data collected across sessions can be regarded as homogeneous for this subject.

CONCLUSION

Although significant performance variations can be indicated by statistical tests, it is difficult to aggregate data with respect to a particular factor (e.g., practice or strategy) given multiple dependent variables. Cluster analysis allows us to determine whether data that have been collected over many experimental sessions (or trials) are truly homogeneous with respect to the factor of concern and, if not, to obtain subsets of homogeneous data. When there are different clusters, the clusters' properties offer an opportunity to characterize the corresponding performances, and the occasions on which particular data sets that enter into a cluster are obtained offer a further clue (such as early or late in the experiment). Thus, it is advisable to apply both a statistical model (ANOVA) and an analytic tool (cluster analysis) to discover and characterize the effects of practice and covert strategies. With this set of screening techniques, subjects need not be trained to some arbitrary criterion of performance prior to collecting data for an experiment—all the data are useful and informative. Attention to these details can yield interesting insights into the bases of performance.

Acknowledgments—This research was supported by Grant N00014-88-K-0569 from the Office of Naval Research, Perceptual Sciences Program, and by Grant 91-0178 from the Air Force Office of Scientific Research, Life Sciences, Visual Information Processing Program.

REFERENCE

SAS Institute Inc. (1985). *SAS® User's Guide: Basics, Version 5 Edition*. Cary, NC: Author.