Management of Computing

Gordon B. Davis Editor

An Experimental Evaluation of the Impact of Data Display Format on Recall Performance

NARAYAN S. UMANATH and RICHARD W. SCAMELL

ABSTRACT: *Recall, while an important topic in the study of learning and memory, has received relatively little attention as a dependent variable in studies that investigate alternative formats for presenting information. This paper describes two experiments, performed back to back, that examined the relationship between data display format and recall performance across different task categories. The results of Experiment 1 were reaffirmed by Experiment 2 and collectively suggest that a graphical presentation enhances recall when the task possesses a spatial orientation while the recall of specific facts is indifferent to data display format.*

1. INTRODUCTION

While psychologists have studied the memory functions of recall, recognition, and forgetting since the turn of the century [7, 8, 25], researchers in the information systems and administrative sciences fields have focused little attention on the apparently instinctive and spontaneous brain function of recall and its probable impact on human decision processes in organizational settings. Decision making requires identifying and configuring relevant facts in order to support hypothesized inferences. Detecting and using logical configurations of related facts is a difficult activity when external information sources such as texts must be used, but becomes an easier task when the source of information is the highly developed human memory mechanism. In particular, when a task requires information processing of greater sophistication than simple fact retrieval, such as that required for deductive reasoning or decision making, facts structured appropriately in memory are considered to be far superior to external information sources [12].

Mintzberg, discussing the real-time role of managers, states, "Managers' jobs at lower levels are more oriented towards issues that are current and specific, and are more focused—concentrated on a narrower range of issues, which span shorter periods of time." [19] Real-time management activities are often characterized by unexpected problems and opportunities that demand immediate attention. Given the value of human memory as an information resource for decision making plus the real-time focus of lower-level management issues, the investigation of recall ability as a contributing factor to effective management decision making merits attention. In addition, results from a recent survey (see section 3.6 for additional details) of 77 managers located at different levels in the organizational hierarchy in the Houston area provides strong support to the premise that a routine reliance on recall ability for on-the-spot (real-time) decisions of varying degrees of importance is veritably common. Since an increasing number of decision makers across functional areas use information obtained from computer-based management information or decision support systems, it is likely that the object of their recall may be a computer-generated report.

Mode of presentation of information has long been considered an important variable in information systems research [18]. In fact, a majority of studies that investigate the role of MIS in management decision making have used presentation format (tabular, graphical, etc.), presentation medium (hard copy, CRT, etc.) and, more recently, report features such as color [2, 31] to represent the MIS variable. In general, research involving the choice of either a graphical or tabular presentation of information for decision situations [15, 16, 22, 27, 30] has yielded equivocal results. The burgeoning addition to the state of conflict in this area of research has invoked genuine concern. However, a closer scrutiny of studies on visual displays that report contradictory results reveal that the studies differ in task content/characteristics, experimental conditions and procedures, environmental conditions and subject characteristics [4, 6].

A proper evaluation of display format as an MIS variable requires a consolidated effort through carefully designed experiments across a variety of task settings. One such setting concerns recall and recognition [2]. A recent study by Watson and Driver [27] investigates the effect of computer graphics on the degree of recall. Using a three-dimensional perspective projection and its tabular counterpart as treatment variations and a single measure of degree of recall, they reported no significant superiority in recall performance attributable to a graphical format of information presentation.

The ultimate objective of our research is to examine the role of recall in the real-time decision process and to evaluate the impact of such variables as mode of presentation of information, task characteristics and individual cognitive skills/styles on real-time decision making. This paper reports on the impact of data display format on recall performance from two laboratory experiments. Since the effectiveness of data display format is said to vary as a function of the task to be performed [4, 6], Experiment 1 examined the relationship between the two variables, data display format and recall performance, to first see if a relationship existed between them and if so, to find out if the relationship was mediated by recall task categories. Experiment 2 used the experience gained from Experiment 1 to improve the experimental design and controls in order to determine if the inferences suggested by the results of Experiment 1 could be verified.

2. CONCEPTUAL BACKGROUND

There are two major stages in the development of a cognitive skill that translate as the growth of two knowledge states—declarative knowledge and procedural knowledge. Declarative knowledge is defined as facts about the skill domain that are interpreted and encoded in a propositional network. Procedural knowledge exists when domain knowledge is directly embodied in procedures for performing the skill [1]. Both declarative knowledge and procedural knowledge are considered essential for problem solving [28].

Recall is a cognitive skill that operates in the realm of declarative knowledge [5]. Nonverbal imagery and verbal processes (including printed words) are said to impact recall differentially [21] since images are analog knowledge representations [11, 28] while verbal processes involve analytical encoding [11]. The general superiority of imaginal over verbal mediators is often explained by the parallel versus sequential processing hypothesis. Here, it is theorized that with imaginal mediators the associative information is stored in parallel whereas verbal mediators are stored sequentially. The difference in the type of organization provided by these two processing systems is correlated with differences in the ease of information processing. Assuming that a stimulus term functions as an effective retrieval cue on recall trials, with parallel processing the image can be quickly scanned and decoded to yield a verbal response. On the other hand, with sequential processing of verbal mediators, memory load during storage could be greater than with imagery, or redintegration[1] of the sequential verbal mediator may take longer and be more susceptible to error than image redintegration during recall [21]. Consequently, superior recall performance, in general, can be expected with an imagistic presentation even when the required response is verbal. Inasmuch as imaginal and verbal codes produced by the encoding process are functionally differentiated primarily in terms of their relative efficiency in information processing [21], it is conceivable that the differential effects may be a function of the task involved. Research suggests that people use mental imagery when tasks require a spatial manipulation of information. There are informal reports that people also use imagery in thinking about abstract relationships [28]. However, if the recall task dictates sequential processing, an imagistic presentation may be at a disadvantage that is compounded further by the fact that visual decoding of quantitative information [3] may be needed with imaginal stimuli. In the context of this research, imaginal and verbal presentations are represented by a graphical display and a tabular display respectively.

3. EXPERIMENT 1

3.1 Development of Hypotheses

Fry [10] states

> Graphs pack a high density of information into a small area . . . are more globally visible than are detailed, symbolic, and sequential . . . tend to show the 'big picture' or gestalt . . . Often interrelationships can be seen better with a graph than with a purely verbal or numerical presentation.

While detection of trends and comparison of multiple trends have been reported to be easier with graphical displays [24], numerical data presented in graphical form were perceived to represent greater complexity [17]. Preference for a tabular format over a bar chart presentation for accuracy and arrangement (orderly and precise) has also been reported [30]. Finally, of particu-

[1] Redintegration is the revival or restoration of a previous mental state.

lar relevance to this study, Washburne's [26] findings indicate that tables were most favorable to the recall of specific amounts, pictographs for simple comparisons, bar charts for complex comparisons and line graphs for trends.

In essence, the above evidence seems to favor the proposition that research results vary as a function of the task as well as the type of report format used. The purpose of this research is to investigate recall as a multi-dimensional construct varying as a function of a recall task category. In order to study this question a laboratory experiment was developed to examine whether certain data display formats enhance performance within certain recall task categories. A bar chart, which despite advances in graphics technology, remains a relatively simple, yet very useful and popular[2] graphics tool, was used as a competing treatment against a conventional tabular display. Three task categories were proposed a priori as independent dimensions of recall: (a) recall of directional order, (b) pattern recall, and (c) specific fact recall.

TABLE I. Hypotheses About Recall Performance

	Recall of Directional Order	Pattern Recall	Special Fact Recall
Immediate Time Domain	RG > RT	RG > RT	RT > RG
Delayed Time Domain	RG > RT	RG > RT	RT > RG

RG is recall performance using a graphical stimulus (bar chart) and RT is recall performance using a tabular stimulus.

Each cell represents a specific hypothesis. For example, the hypothesis indicated in row 1, column 1 suggests that a bar chart format is superior to a tabular format for recall of directional order in the immediate time domain.

Recall of directional order is the ability to recall the rank ordering of items in a set by a specified attribute. A situation where a sales manager must recall sales volume fluctuations and trends over the past several months is an example of recall of directional order.

Pattern recall in the ability to recall characteristics of a suggested pattern across multiple sets as opposed to trend detection (recall of directional order) within a single set of items. Production managers often recall seasonal patterns of excess capacities and overload situations across work centers from production control reports.

The third category of recall, *specific fact recall*, is the ability to recall specific information (point values) from a report. Financial accountants, for example, must frequently recall specific facts and figures from balance sheets and other financial reports.

Often, in decision situations, managers must invoke and rely upon their recall ability spontaneously and almost by rote soon after studying a report. On other occasions, they must recall information from reports reviewed several days or perhaps several weeks ago. Consequently, an investigation of recall ability in both the immediate and delayed time domains was included in Experiment 1.

In short, Experiment 1 attempted to unearth probable effects of data display format on recall ability, scrutinizing the variations of the impact of a single type of graphical display (a bar chart) on three task categories of recall in two time domains. Table I summarizes the six hypotheses associated with the experiment. Consistent with the theoretical framework and prior empirical findings, these hypotheses propose that while graphics may be relatively more effective than tabular presentations for the task categories of directional order and pattern recall, the effectiveness of graphics may not be pronounced over its tabular counterpart with respect to the specific fact recall task.

3.2 The Design

Subjects were divided into two randomized groups and exposed to a data display presented in either a tabular or graphical (i.e., bar chart) format. Subsequently, both groups were tested on their ability to recall information presented in the two display formats. The research methodology employed a multi-group, posttest only with repeated measures design [13].

3.3 Subjects

The 139 subjects in the sample included both graduate and undergraduate students in operations management and industrial engineering at the University of Houston and personnel from local industries. Almost all of the graduate (80 percent) and a large portion (63 percent) of the undergraduate students were employed in various capacities or had previous work experience in either manufacturing and/or service industries.

3.4 Experimental Procedure[3]

The report presented to each subject was a Work Center Load Profile in either a tabular or a graphical format. A bar chart format was selected for the graphical representation because it is commonly employed in work center loading reports [9]. The report was a sequential presentation of monthly load and capacity (in hours) for three work centers for a six-month period. Ives [14] has expressed concern over insufficient attention devoted to the equivalence of information and the relative quality of competing report formats in most studies involving modes of presentation. While the need for precise equivalence between competing formats is debatable [4], an effort was made to achieve (a) an equivalence of the information content in the competing treatments and (b) a good reproduction quality of the treatments. Since the information presented here pertained to an operations-oriented report, the

[2] Responses received to a survey, inquiring into the actual use of a graphical form of reports among 77 practicing managers, indicated a limited use of graphical reports. However, where used, only line graphs and bar charts were overwhelmingly prevalent.

[3] For a detailed account of the experiment, reproductions of the experimental material, description of the scoring procedure, data analysis, and discussion of findings, see [23].

contents happened to be precise and unambiguous. Consequently, the task of infusing a fair amount of equivalence between the competing formats was not difficult in this research setting. The treatments (graphical and tabular displays) were prepared in consultation with an experienced manufacturing engineer and a professional design draftsman.

Experimental data were collected during two sessions. In the first session (the immediate response session), the subjects were randomly assigned to one of two experimental groups, with one group studying the tabular report and the other studying the bar chart—each for a three-minute period. Each subject received a folder containing an instruction sheet which (a) described the experimental setting, (b) directed the subject to review the contents of either a tabular or graphical display and (c) stated that the subject would be answering a questionnaire pertaining to facts contained in the display based exclusively on recall of information just reviewed in the display. At the conclusion of the three-minute period, each subject was asked to respond to a questionnaire designed to test their recall performance in the three categories of recall with nothing to aid them except their memory. The subjects were not informed about a second session (the delayed response session) based on the premise that practicing managers might, from time to time, encounter such situations quite unprepared. The second session was conducted 48 hours later. At this time, the same questionnaire was given to them again in order to measure their delayed recall performance.

Each category of recall was operationalized and measured using a single exemplar (item of measure). The graphical and tabular displays and the questionnaire were pretested on a group of eight doctoral students and faculty in operations management and management information systems. The suggestions of this group were helpful and were incorporated in the final design of the experimental material.

3.5 Summary of Findings

A detailed account of the scoring procedure, data analysis and discussion of findings is available in [23]. Table II reflects recall performance indicators in the three proposed task categories, over the two time domains for the tabular and graphical treatment variations, transformed to a scale of 0 to 1 for ease of evaluation. The results indicate that recall performance was at acceptable levels in all but one condition—specific fact recall in the delayed time domain was poor for both tabular and graphical stimuli. Also, as theorized by Ebbinghaus [7], recall ability, in general, appeared to fade over time.

Observe that, as hypothesized, the graphical form of information presentation was superior to tables both for recall of directional order ($p = .04$) and for pattern recall ($p = .004$) in the immediate time domain; on the contrary, the effectiveness of the graphical format on specific fact recall in the immediate time domain was negligible ($p = .35$). Overall, the experimental findings provided reasonable support to the suggestion that the utility of graphics may be contingent upon the task at hand [4].

TABLE II. Recall Performance Results (By Task Category)

	Time Domain	Display Format		
		Graphical	Tabular	p
Recall of Directional Order	Immediate	.7333[a] ($n = 61$)	.6000 ($n = 60$)	.04
(Kendall's Tau, median score)	Delayed	.6000 ($n = 32$)	.4667 ($n = 25$)	.07
Pattern Recall	Immediate	.6516[b] ($n = 66$)	.4265 ($n = 68$)	.004
(Proportion correct, mean score)	Delayed	.6304 ($n = 46$)	.5405 ($n = 37$)	.17
Specific Fact Recall	Immediate	.4637[c] ($n = 68$)	.4460 ($n = 71$)	.35
(Proportion correct, mean score)	Delayed	.3578 ($n = 49$)	.2770 ($n = 45$)	.06[d]

[a] Recall of directional order was measured by asking the respondent to reproduce rankings contained in the data. Over 50 percent of the subjects had a value of Kendall's Tau ≥ .7333 while 75 percent had a value of Kendall's Tau ≥ .6000. Three (3) subjects were able to reproduce the rankings in the exact order.
[b] Pattern recall was measured by testing to see whether the respondents were able to detect a simple pattern in the data. Responses were categorized with respect to the degree of departure of the subject's answer from the correct answer. Forty-three (43) of the 66 subjects (65.15 percent) using the graphical format were able to detect the correct pattern.
[c] Specific fact recall was measured by asking each respondent to answer 15 questions—each of which required the use of a specific value contained in the report. On the average, the 68 respondents using the graphical format gave the correct answer 46.37 percent of the time.
[d] Not relevant since both the tabular and the graphical reports had poor impact on recall.

3.6 Ex-Post-Facto Evaluation

The findings of Experiment 1 were encouraging and offered sufficient grounds for continuing this line of research. Accordingly, specific aspects of the experiment were identified where design improvements plus the utilization of better controls would lead to the enhancement of internal validity and reliability.

Before further experimentation, a survey of managers in the Houston area was conducted in an effort to ratify the basic premise underlying this study: *recall plays an important role in the real-time decision process.* The results of this survey of seventy-seven managers indicated that

a. managers have a definite need to make on-the-spot decisions,
b. a majority of the on-the-spot decisions made are of an operational or management control nature,
c. managers frequently rely on experienced-based "gut feelings" and their memory for information (recall) for on-the-spot decision tasks, and
d. bar charts and line graphs are the most widely used graphical forms of reports among managers today.

Since the results of Experiment 1 suggest that the impact of data display format on recall performance

may be contingent upon task category, a more indepth examination of recall categories, as a function of the type of task to be performed, was conducted in Experiment 2. In addition, since the three categories of recall, defined a priori in Experiment 1, were operationalized by a single exemplar for each category, the categorization of recall in Experiment 1 suffered questionable validity/reliability. Therefore, in an attempt to empirically derive the recall categories, 11 test items were formulated based on the experience gained from Experiment 1.

Each of the 13 members of a carefully chosen panel[4] evaluated the 11 test items. Suggestions for other test items and/or rewording of existing test items to promote understanding were solicited from the panel members. While no new test items were proposed, several were revised based on the panel's responses. The 11 test items, seven self-report questions regarding confidence level, preferences and format familiarity, and demographic questions on the subject's gender, education, experience, etc., were assembled as a three-part questionnaire.[5]

Another important improvement made in Experiment 2 involved the replacement of the three recall task categories defined a priori in Experiment 1 with two recall task categories derived from experimental observations using factor-analytic procedures. A review of the factor-analytic derivation of the recall task categories is presented in Section 4.1.

Two other groups participated in a pretest of the experimental material. In a pilot test, a set of six doctoral students critically evaluated the reports, the questionnaire and the experimental procedure. In order to empirically assess the quality and interpretability of the experimental material, the experiment was administered on a naive audience.[6] An evaluation of the performance of this group verified the adequacy of the general quality of the competing data display formats.

Since business graphics are considered novel [4], the sampling procedure in Experiment 1 endeavored to control for an unfamiliarity bias by sampling only industrial engineering and operations management students. But, empirical evidence to substantiate this effort was not gathered. Therefore, the questionnaire designed for Experiment 2 asked the subjects to rate their degree of familiarity with tabular and graphical display formats.

[4] The panel consisted of four professionals from industry with a minimum of 15 years of industrial/management experience and graduate level education, three doctoral students and six faculty members from the College of Business Administration, University of Houston.

[5] Parts I and II comprised the recall test items, while Part III contained self-report inventories and demographic questions. A copy of the questionnaire is available from the authors upon request.

[6] 51 undergraduate (junior and senior) full-time students were used for this purpose, and were considered to be a representative sample of a naive audience with respect to the experimental setting (work center load profile). The students were asked to answer the questionnaire by referring to the tabular or graphical report while answering. If interpretability of the reports and questions can be established with a naive audience, it should be reasonable to assume that the experimental material would be adequately interpretable to an audience familiar with the report formats and task content.

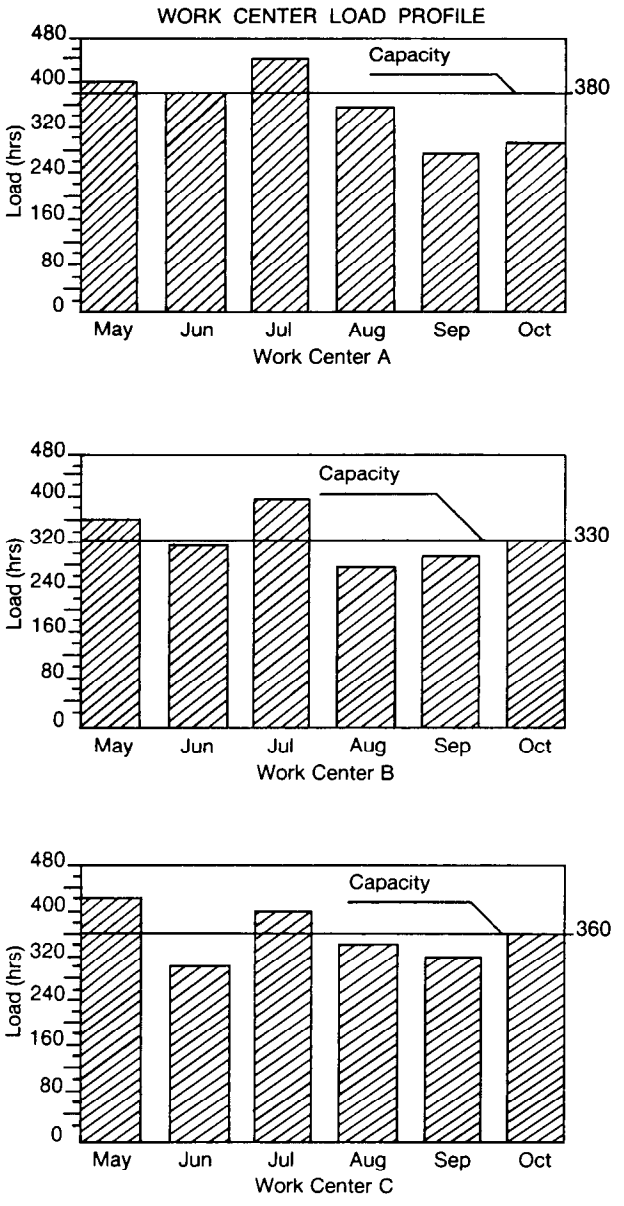
Another important concern that emerged from the experience of Experiment 1 was that while the subject had an opportunity to study the displays, the level of cognitive encoding of information was unknown. Since the recall performance of interest to this study did not pertain to recall of nonsensical data from rote memory, it was necessary to induce semantic encoding of information (learning) in the subjects. In organizations, such cognitive encoding at the appropriate level (visual, auditory, or semantic) from various information sources may be generic to the job or position held by the individual. Therefore, an exercise in problem solving and decision analysis was designed as a pretest treatment. It was expected that through subject participation in such an exercise, learning would be unobtrusively induced. Further, performance on the decision problem also serves as an indicator of the comprehensibility in addition to the previously established interpretability of the reports.

Finally, both the bar chart and its tabular counterpart (shown in Figure 1) were improved in several ways based on feedback from Experiment 1.

4. EXPERIMENT 2

Since the results of Experiment 1 supported the common sense notion of loss of recall over time [7], further investigation of delayed recall was relegated to future research. Therefore, Experiment 2 employed a multigroup, posttest only design that contained the enhanced controls discussed earlier without the delayed recall session. The 108 subjects that participated in the experiment were all MBA students at the University of Houston. More than 86 percent of the participants were also employed full-time and the average industrial/business experience of the experimental group was six years. Participants were assigned at random to one of two experimental groups. The subject ratings in the questionnaire indicated adequate familiarity with both graphical and tabular forms of information presentation. Furthermore, regarding familiarity with the particular treatment each experienced, the two groups were found statistically equivalent.

Six questions were used to measure recall performance and were compiled from the 11 test items. Except for one question, the scoring scheme underlying the questions involved measurement of the degree of agreement/departure of the subject's response from the correct answer. Several scoring schemes were evaluated for the recall performance measure using fictitious data. The singular objective of this evaluation was to arrive at a good measure for the degree of agreement/departure of the subject's response from the correct answer. Some of the scoring schemes discarded after examination were Spearman's rank correlation, Kendall's Tau, just the number of discordances and just the number of concordances. Spearman's rank correlation was considered since Watson and Driver [27] employed it to measure recall performance. Kendall's Tau was preferred over Spearman's Rho even in Experiment 1, since the distribution of Kendall's Tau approaches nor-



WORK CENTER LOAD PROFILE

	May	Jun	Jul	Aug	Sep	Oct
Capacity (hrs)	380	380	380	380	380	380
Load (hrs)	400	380	440	360	280	300

Work Center A

	May	Jun	Jul	Aug	Sep	Oct
Capacity (hrs)	330	330	330	330	330	330
Load (hrs)	360	320	400	280	300	330

Work Center B

	May	Jun	Jul	Aug	Sep	Oct
Capacity (hrs)	360	360	360	360	360	360
Load (hrs)	420	300	400	340	320	360

Work Center C

FIGURE 1. Graphical and Tabular Treatment Variations

mality quite rapidly; thus the normal approximation is better for Kendall's Tau than it is for Spearman's Rho when the null hypothesis of independence between the subject's response and the correct answer is true. Then it was discovered that correlation, while quantifying a linear relationship between two sets, need not necessarily reflect the degree of agreement/departure between the sets. Number of concordances and discordances were eliminated from consideration because they are involved in the computation of Kendall's Tau and therefore tend to parallel Kendall's Tau. Intuitively it appeared that the number of simple permutations required to transform the subject's response to the correct answer may be a reasonably sound measure of the degree of agreement/departure between the two. In order to ease computation and understanding an approximation of this scheme was chosen to score recall performance.[7]

4.1 Experimental Procedure

The administration of Experiment 2 was quite similar to that of Experiment 1 except that (a) all the participants were exposed to a pretest treatment—the exercise in decision analysis discussed earlier and (b) measurements were taken at only one point in time.

[7] Details of the computational procedure used are available from the authors upon request.

A statistical analysis of subject scores in the decision analysis exercise[8] revealed that the performance was quite good for both the tabular and graphical treatment groups, thus indirectly favoring the supposition that the displays were "comprehensible." The tabular and graphical displays (See Figure 1) used as competing treatment variations were improved versions of the treatments used in Experiment 1.

With a view to rigorously explicate recall task categories, the subjects' scores on the six measures of recall performance were factor analytically examined. Approximately one half of the data (50 observations) collected in Experiment 2 was randomly selected for exclusive use in a factor-analytic derivation of recall task categories. Even though the subjects to variables ratio was adequate,[9] Bartlett's test was employed to determine whether "enough" statistically significant correlations were present among the variables to yield factors that represent something other than the intercorrelation of random data [29]. The chi-square value (37.55) indicated presence of "enough" common variance among the variables ($p = .005$). The principle components method was employed for factor condensation from which two factors emerged. Table III shows the pattern of factor loadings. Labeling categories derived from factor analysis in generalizable terms is risky because such labels are often conveniently but erroneously interpreted as meaningful beyond the confines of the research context. With this word of caution, the derived recall task categories are given below:

Pattern Integration Recall refers to recall of information that requires identification of temporal and set-integrative patterns. Four of the six measures gathered belong to this category.

Simple Fact Recall refers to recall that requires memory for specific point values of data and simple comparisons. The other two of the six measures comprise this category.

4.2 Statement of Hypotheses

Under the assumption that pattern integration requires a spatial manipulation of data whereas simple fact retrieval is more of a sequential process, the theory suggesting the superiority of imagery for tasks requiring spatial manipulation and lack of superiority of imaginal mediators for sequential processing tasks (see Section 2) lead to the following two hypotheses:

A graphical display format is significantly superior to its tabular counterpart for pattern integration recall.

A tabular display format is significantly superior to its graphical counterpart for simple fact recall.

[8] A simple well-structured forecasting problem was presented to the subjects in the form of a one-page case. The two experimental groups using the tabular and graphical report respectively, made a forecasting decision. The decision and the decision process were evaluated based on the content of the decision analysis response form completed by each subject. A detailed account of this exercise can be obtained directly from the authors.

[9] Nunnally [20] suggests a subjects to variables ratio of 10 as moderately large and prescribes a value of 5 as the least acceptable. In this factor analysis, the subjects to variables ratio was 8.

TABLE III. Principal Components Analysis of Recall Measures (Varimax Rotated Factor Pattern)

Recall* Measure	Factor 1	Factor 2	Communality Estimates
Q11	0.76548	0.29088	0.670569
Q12	0.73133	0.25875	0.601804
Q23	0.62285	−0.32105	0.706837
Q21	0.58571	−0.07921	0.349328
Q13	−0.10254	0.83446	0.507613
Q22	0.16707	0.69260	0.491018
Variance Explained	1.890234 32%	1.436935 24%	3.327169 56%

Note: Factor condensation based on eigenvalue > 1.0
* Q11 is the subject score on Question 1 of Questionnaire—Part I
Q12 is the subject score on Question 2 of Questionnaire—Part I
⋮
Q23 is the subject score on Question 3 of Questionnaire—Part II

4.3 Method of Analysis

Since 50 observations from the experiment were used for the factor-analytic derivation of recall task categories, only the remaining 58 observations were utilized for the data analysis. A careful examination of the central tendencies, frequency distributions, and data plots followed by the Kolmogorov Goodness-of-Fit test indicated that the two groupings of scores analyzed exhibited a significant departure from normality. However, since the exploratory data analysis showed that the sample distribution about the median displayed a reasonable level of symmetry in both cases, the conservative and robust nonparametric Wilcoxon Rank Sum Test (Mann–Whitney Test) was used to test for the equality of medians between the recall performance indicators using the tabular and graphical treatments across the two recall categories.

4.4 Discussion of Findings

The findings associated with the two hypotheses examined are presented in Table IV. The results indicate the graphical display to be significantly more conducive to information recall than the tabular display when the task required memory for temporal and set-integrative patterns. But, contrary to the hypothesis, recall of simple facts, such as point values and simple comparisons,

TABLE IV. Recall Performance Results (Wilcoxon Rank Sum Test)

	Pattern Integration Recall Performance H: RG > RT			Specific Fact Recall Performance H: RT > RG		
Display Format	Sample Size	Median Score	p	Sample Size	Median Score	p
Graphical	25	.7208	.006	27	.7500	.993
Tabular	25	.5695		27	.7500	

RG is recall performance using the graphical stimulus (bar chart).
RT is recall performance using the tabular stimulus.

was indifferent to variations in presentation format. Since the sample size was reasonably large, a parametric one-way analysis of variance (ANOVA) was performed as a follow-up investigation. While confirming the above findings, the results of the one-way ANOVA indicated that display format explained 17 percent of the variance in pattern integration recall performance whereas the variance in simple fact recall attributable to display format was negligible (.1 percent).

Findings regarding the expressed level of confidence about recall performance, (format) preference, and (format) familiarity may at best be labeled interesting and used for future hypothesis generation since the constructs evaluated were operationalized with single item self-report measures and thus can generate only weak inferences. Analysis of covariance indicated that actual recall performance had the most impact on confidence level. In other words, overall, subjects who performed better displayed a higher level of confidence in their performance. Compared to performance, the effect of preference and familiarity with the format on confidence level was negligible.

Another interesting observation that provides an unobtrusive measure of the "cognitive strain" encountered by the subjects was the extent of their spontaneous use of external media while working on the decision analysis exercise. The data analysis revealed that 56 percent of the subjects in the tabular treatment group resorted to "scribbling on the displays" with such notation as summations/differences, symbols (+'s, −'s), and rough graphical sketches, while only 20 percent of the subjects in the graphical treatment group scribbled at all. Under the assumption that tables induce sequential processing, use of tabular displays for problem solving/decision making could induce cognitive strain which may be avoidable with the use of other forms of information presentation. Further, the reaction to cognitive strain may reduce the quantity of information assimilated and consequently impair recall performance.

5. CONCLUDING REMARKS

Evidence supporting the importance of recall to real-time decision making is quite favorable. Experiment 1 was an endeavor to identify a direction for research concerning recall ability as it applies to decision making. The principal independent variable of interest here was data display format. Since it was suspected that the impact of data display format on recall performance might be contingent upon type of task involved, the experiment was designed to investigate the relationship between data display format and recall performance in three recall tasks categories that were defined a priori. The inferences of Experiment 1 offered support to the contention that a graphical format of information presentation may have a unique value in the context of recall enhancement.

The experience gained from Experiment 1 was instrumental in intensifying the focus of this research. Thus, in Experiment 2, serious effort was expended for a more rigorous explication of recall categories. While the empirical derivation resulted in the development of two recall categories, it can be seen that the three recall task categories (directional order, pattern and specific fact) defined a priori in Experiment 1, may be logically reduced to the pattern integration/specific fact dimension derived in Experiment 2. In this light, the results of Experiment 2 support the inferences arrived at in Experiment 1. As a result, the cumulative effect of the two experiments lends strength to the conclusion regarding the conduciveness of graphics to recall tasks involving pattern recognition as well as the indifference of recall ability to report format for simple fact retrieval tasks.

The conclusions made at this point require qualification. Individual differences and demographic characteristics (for example, age, experience, gender) may affect recall and have not been included in this investigation. The so-called "list length" (amount of information present) of the reports was arbitrary, and thus it is not known whether the competing treatments contained too little or too much information. The time of exposure of the subjects to the stimuli (three minutes) was based on the pretest on the pilot group. However, during the administration of Experiment 1, it was observed that very few subjects studied the reports for more than two minutes. While insufficient time may impair proper assimilation of the report contents, excessive time may also affect the subject in the form of anxiety and boredom.

The principal contribution of this research, it is hoped, is the reinstatement of the sparsely studied subject of recall as a topic for management and MIS research. This intriguing memory function is often taken for granted by people in their personal and professional lives. This raises the question of how much does recall ability aid decision making. The survey of 77 managers suggests that recall ability is an important factor in real-time decision making. But, further research is essential to affirm or negate the intensity of the influence of recall ability on decision making.

Of particular interest to those involved in real-time decision making, the experimental findings suggest that a graphical display of information is at least as good as tables for recall in general. Furthermore, where tasks involve the identification of trends and/or pattern integration across units, use of a graphical display in place of tabular reports should be considered. Of course, the benefits of graphics will have to be balanced against the cost of generating them. With the advent of computer-based managerial graphics and the prevalent competitive vendor environment, a declining trend in cost can be anticipated. As is typical with any novel managerial tool, confidence can be developed with prolonged use. The fact that almost all of the 108 experimental subjects were practicing managers, with an average experience of six years, lends credibility to our findings. Yet, these implications should be viewed with a touch of skepticism because without consistent replication of findings in several different contexts and experimental conditions, these results should be viewed as tentative.

An important area for future research pertains to the design variables. In addition to their identification, it is crucial to isolate the key variables that exhibit significant influence on recall performance. Mode of presentation, task properties, order effects (primary/recency), list length (information overload or underload), scale(s) of graphical treatments, individual differences, and demographic characteristics are all candidates for analysis. Finally, since many decision makers are routinely exposed to the same form of management report(s) repeatedly, studies of multi-trial recall performance may be another future productive area for research.

REFERENCES

1. Anderson, R.J. Acquisition of cognitive skill, *Psychological Review 89*, 4 (1982), 369–406.
2. Benbasat, I., and Dexter, A.S. An experimental evaluation of graphical and color-enhanced information presentation, *Management Science, 31*, 11 (Nov. 1985), 1348–1364.
3. Cleveland, W.S., and McGill, R. Graphical perception: Theory, experiments and application to the development of graphical methods, *Journal of American Statistical Association 79*, 387 (Sept. 1984), 531–554.
4. DeSanctis, G. Computer graphics as decision aids: Directions for research, *Decision Sciences 15*, 4 (Apr. 1984), 463–487.
5. DeSanctis, G., and Jarvenpaa, S.L. An investigation of the "tables versus graphs" controversy in a learning environment. In *Proceedings of the Sixth International Conference on Information Systems* (Chicago 1985), Soc. Inf. Manage., 134–144.
6. Dickson, G.W., DeSanctis, G., and McBride, D.J. Understanding the effectiveness of computer graphics for decision support: A cumulative experimental approach. *Commun. ACM, 29*, 1 (Jan. 1986), 40–47.
7. Ebbinghaus, H. *Memory*. (H.A. Ruger and C.E. Bussenius, trans.). Dover, New York, 1964. (Originally published in *1885*).
8. Erdelyi, M.H., and Kleinbard, J. Has Ebbinghaus decayed with time: The growth of recall (Hypermnesia) over days, *Journal of Experimental Psychology: Human Learning and Memory 4*, 4 (Apr. 1978), 275–289.
9. Fogarty, D.W., and Hoffman, T.R. *Production and Inventory Management*. South-Western, Palo Alto, Calif., 1983.
10. Fry, E. Graphical Literacy, *Journal of Reading 24*, 5 (May 1981), 389–392.
11. Glass, A.L., Holyoak, K.J., and Santa, J.L. *Cognition*. Addison-Wesley, Reading, Mass., 1979.
12. Hayes-Roth, B., and Walker, C. Configural effects in human memory: The superiority of memory over external information sources as a basis for inference verification. *Cognitive Science 3*, 2 (Apr.–June 1979), 119–139.
13. Huck, S.W., Cormier, W.H., and Bounds, W.G., Jr. *Reading Statistics and Research*. Harper and Row, New York, 1974.
14. Ives, B. Graphical user interfaces for business information systems. *MIS Quarterly* Special Issue *12*, 16 (Dec. 1982), 15–47.
15. Lucas, H.C. An experimental investigation of the use of computer-based graphics in decision making. *Management Science 27*, 7 (Jul. 1981), 757–767.
16. Lucas, H.C., and Nielson, N.R. The impact of information presentation on learning and performance. *Management Science 26*, 10 (Oct. 1980), 982–993.
17. Lusk, E.J., and Kersnick, M. The effect of cognitive style and report format on task performance: The MIS design consequences, *Management Science 28*, 8 (Aug. 1979), 787–798.
18. Mason, R.O., and Mitroff, I.I. A program for research :n management information systems. *Management Science, 19*, 5 (May 1973), 475–485.
19. Mintzberg, H. *The Nature of Managerial Work*. Prentice-Hall, Englewood Cliffs, N.J., 1980.
20. Nunnally, J.C. *Psychometric Theory*. McGraw-Hill, New York, 1978.
21. Paivio, A. *Imagery and Verbal Processes*. Lawrence Erlbaum Associates, Hillsdale, N.J., 1979.
22. Remus, W. An empirical investigation of the impact of graphical and tabular data presentations of decision making, *Management Science, 30*, 5 (May 1984), 533–542.
23. Scamell, R.W., Umanath, N.S., and Das, S.R. An assessment of the effectiveness of report graphics on recall ability. Working Paper No. 172, University of Houston, 1986.
24. Schutz, H.G. An evaluation of formats for graphical trend displays—Experiment II. *Human Factors*, 3, 2 (Feb. 1961), 99–107.
25. Tulving, E. Theoretical issues in free recall. In *Verbal Behavior and General Behavior Theory*, T.R. Dixon and D.L. Horton, Eds. Prentice-Hall, Englewood Cliffs, N.J., 1968.
26. Washburne, J.N. An experimental study of various graphic, tabular and (textural) methods of presenting quantitative material. *Journal of Educational Psychology 18*, 6 (1927), 361–376.
27. Watson, C.J., and Driver, R.W. The influence of computer graphics on the recall of information. *MIS Quarterly 7*, 1 (Jan. 1983), 45–53.
28. Weber, E.S. Cognitive process involved in solving information systems (IS) design problems. In *Proceedings of the Sixth International Conference on Information Systems* (Chicago, 1985), Soc. Inf. Manage., 305–312.
29. Weiss, D.J. Factor analysis and counseling research. In *Methods and Analysis in Organizational Research*, R.S. Bateman and G.R. Ferris, Eds. Reston, Reston, Va., 1984.
30. Zmud, R.W. An empirical investigation of the dimensionality of the concept of information. *Decision Sciences 9*, 2 (Feb. 1978), 187–195.
31. Zmud, R.W., Blocker, E., and Moffie, R.P. The impact of color graphic report formats on decision performance and learning. *Proceedings of the Fourth International Conference on Information Systems* (Chicago, 1983), Soc. Inf. Manage., 179–193.

CR Categories and Subject Descriptors: H.1.2 [**Models and Principles**]: User/Machine Systems; 13.m [**Computer Graphics**]: Miscellaneous; E.m [**Data**]: Miscellaneous

General Terms: Human Information Processing

Additional Key Words and Phrases: information systems, recall performance

Received 1/85; revised 11/87; accepted 12/87

Authors' Present Addresses: Richard W. Scamell, Department of Decision and Information Sciences, College of Business Administration, University of Houston, Houston, Texas 77004; Narayan S. Umanath, Accounting and Management Information Systems, College of Business Administration, The Pennsylvania State University, University Park, Pennsylvania 16802.