INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

Order Number 8900283

**A contrast of guideline recommendations and Tullis's prediction
model for computer displays: Should text be left-justified?**

Fontenelle, Gail Ann, Ph.D.

Rice University, 1988

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy.
Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy ✓
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received _____
16. Other _____

U·M·I

RICE UNIVERSITY

A CONTRAST OF GUIDELINE RECOMMENDATIONS
AND TULLIS'S PREDICTION MODEL FOR COMPUTER DISPLAYS:
SHOULD TEXT BE LEFT-JUSTIFIED?

by

GAIL FONTENELLE

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

DOCTOR OF PHILOSOPHY

APPROVED THESIS COMMITTEE:

Dr. William C. Howell
Professor of Psychology, Chairman

Dr. David M. Lane
Associate Professor of Psychology

Dr. Kenneth R. Laughery
Professor of Psychology, Luce Chair

Dr. Robert M. Stein
Associate Professor of Political Science

Houston, Texas
May, 1988

A CONTRAST OF GUIDELINE RECOMMENDATIONS
AND TULLIS'S PREDICTION MODEL FOR COMPUTER DISPLAYS:
SHOULD TEXT BE LEFT-JUSTIFIED?

GAIL FONTENELLE

ABSTRACT

Two experiments investigated the effect of layout complexity for performance at varying levels of practice on four types of information extraction tasks. Layout complexity is defined as the number of unique horizontal and vertical starting positions of items in the display (Tullis, 1984). Tullis investigated the ability of display-formatting variables (overall density, local density, number of groups, size of groups, number of items, and layout complexity) to predict human performance and preference. Although layout complexity was the best predictor of subjective ratings, it did not contribute to the prediction of search time beyond what could be predicted by overall density, local density, number of groups, and size of groups. This is a particularly interesting finding since vertically aligning lists and left-justifying items, practices strongly recommended in guidelines for display formatting (e.g., Engel & Granada, 1976; Smith & Mosier, 1986), are important factors in Tullis's definition of layout complexity. Thus, the guidelines and Tullis's model lead to conflicting predictions concerning the effect of left-justifying text on user search time. In the first study, layout complexity was manipulated by either left-justifying or not left-justifying text. Although the text was not left-justified, the starting positions of the text were ordered rather than random. In the second study, subjects

viewed a third experimental screen that displayed the starting positions of items in a completely unpredictable pattern. Subjects performed all four tasks (find label, scan data, compare labels, and compare data) in four one-hour sessions. Moderate violations of the typical guideline recommendations did not increase user search time across all four tasks in either the first or second study. However, when subjects compared multiple data values, the random format did increase user search time. Though performance using the three experimental screens was comparable across the four tasks with only one exception, subjective ratings demonstrated differences between the three formats. Subjects disliked the random format and degraded their performance using the random screens.

TABLE OF CONTENTS

	<u>Page</u>
Abstract	i
Introduction	1
Experiment 1	38
Method	38
Results	46
Discussion	46
Experiment 2	49
Method	49
Results	52
Discussion	55
General Conclusions	58
References	60
Appendix A. Experimental Screens	66
Appendix B. Analysis of Variance Tables	91
Appendix C. Subjective Rating Forms	94
Appendix D. Raw Error Data for Experiment I & II	119

LIST OF TABLES

	<u>Page</u>
TABLE 1. Summary of Documents Containing Screen Guidelines.	6
TABLE 2. Selected Empirical Studies of Screen Design.	11
TABLE 3. Definitions of the Six Conceptually Distinct Dimensions Underlying Display Formatting (Tullis, 1984).	13
TABLE 4. Means, Standard Deviations, and Ranges of Six Dimensions for the Screens Used by Tullis (Experiment1; Tullis, 1984).	19
TABLE 5. Best-fitting Multiple Regressions Predicting User Search Time and Subjective Ratings (Tullis, 1984).	20
TABLE 6. Correlation Matrix of the Six Format Dimensions for the Displays (N=520) Used by Tullis(1984).	23
TABLE 7. Correlation Matrix of the Four Format Dimensions for the Displays (N=6) Used by Schwartz (1986).	24
TABLE 8. Means, Standard Deviations, and Ranges of Six Dimensions for the Left-justified and Predictable Screens Used in Experiment 1.	42
TABLE 9. Means, Standard Deviations, and Ranges of Six Dimensions for the Random Screens Used in Experiment 2.	51

LIST OF FIGURES

	<u>Pages</u>
Figure 1. An example of two screens used by Tullis (1984).	39
Figure 2. An example of an experimental screen contrasting left-justified (lower panel) and predictable (upper panel) formats.	41
Figure 3. Examples of questions for each of the four information extraction tasks.	43
Figure 4. Means of median reaction time (seconds) in Experiment I as a function of task and layout complexity.	47
Figure 5. An example of an experimental screen in random format.	50
Figure 6. Means of median reaction time (seconds) in Experiment II as a function of task and layout complexity.	53

Introduction

Perspective on the problem: design of alphanumeric displays

Recent developments in information technology continue to reduce hardware costs and constraints while producing an explosion in software expenditures and capabilities. Consequently, interaction with computer technology is no longer restricted to a small group of specialists. It is becoming increasingly necessary for members of the modern society to interact with computer systems. The link between the information resident within the computer and the human user is known as the human-computer interface. One facet of the typical human-computer interface is a software-generated display which allows the user to interact with a computer-based system. In a survey conducted by Smith and Mosier (1984), software design professionals estimated that 21 to 47 percent of all operational software is designed to implement the user interface. These estimates suggest that the investment of time and money in the human-computer interface represents a sizeable portion of the total system software design effort. Industry is becoming increasingly aware that the design of the human computer interface in any interactive system is crucial for its efficiency, its acceptability, and therefore its commercial potential (Monk, 1984).

At the present, there are few tools or methods available to the software designer to aid in decisions about the human computer interface. For example, one of the central issues in the design of the human computer interface is the design of software generated displays. The software generated display presents information to be acted upon by the user; consequently, the design of the display has substantial

impact on the usefulness and efficiency of the human computer interface. Poor screen design can have a detrimental effect on human performance by increasing the time needed to extract information from the screen, provoking errors, and complicating machine operations (Galitz, 1985; Keister & Gallaway, 1983; Tullis, 1981). It is also recognized that some ways of presenting and coding information are superior to others (e.g., color, graphics, highlighting); moreover, the amount and nature of the advantage is highly dependent on the perceiver's task (Hitt, 1961; Schwartz & Howell, 1985). Yet, one of the more basic decisions for screen designers--how to arrange information spatially on a screen in order to facilitate an operator's performance on a specific task--has received little attention. Little is known about the effect of spatial characteristics on the most common type of display format, the alphanumeric display, or how these spatial characteristics interact with different types of tasks. While display capabilities in information technology continue to advance (e.g., three dimensional graphics, media rooms, super cockpits), the development of techniques and methodologies for empirically evaluating the usefulness of these new technologies lags far behind. The challenge for human factors specialists is to develop HCI design methods and tools that provide effective support to screen designers (Smith & Mosier, 1985).

The present research focuses on the importance of spatial characteristics in screen design and their interaction with task characteristics. It is restricted to an examination of alphanumeric displays presented on single screens. While other features such as color, windows, graphics, and the like are equally valid candidates for study, the fact that so little is known about the how properties of the more fundamental

alphanumeric display affect user performance and acceptances argues for according it priority as a research topic. The following section discusses the state of the art in screen design and reviews the available literature.

Resources for addressing the problem

The literature on screen design may be divided into two general categories: guidelines and empirical studies. The following two sections summarize information about the design of alphanumeric displays that can be gleaned from these two sources. Several recent reviews on the more general topic of screen design serve as particularly good resources and deserve special mention. Smith and Mosier (1986) compiled an extensive list of specific guidelines for interface design; furthermore, they supported these guidelines with numerous references to earlier documents and empirical studies. Tullis (1986) provided the most recent and thorough discussion of the narrower topic of screen design. A discussion of more general human factors issues concerning the human-computer interface can be found in Shneiderman (1987), Monk (1984) and Salvendy (1984).

Guidelines

Human engineering standards and design handbooks have been of little use to the screen designer. The popular human factors design handbook by Woodson (1981) is typical. It includes only three pages of material on information processing and there is no reference to computers in the index. MIL-STD-1472B, the major human engineering design standard for military procurement, was concerned almost

exclusively with hardware and physical safety. In 1981, a revised "c" version of the standard was published which included nine pages dealing with interface design. Later this material was expanded to 19 pages in a later revision of MIL-STD-1472C (1983). This was a beginning but much more is needed. More recently, professionals have addressed the problems of screen design by summarizing the state of the art in handbooks and guidelines written exclusively for that purpose. While these guidelines are of some use, their main contribution was to heighten awareness concerning the importance of screen design and to direct attention to problem areas.

Many guidelines are contracted technical reports consisting of standards written to establish specification requirements for a particular software system. A few examples are the IBM report by Engel and Granada (1975), and the MITRE reports by Smith and his colleagues (Smith & Aucella, 1983; Smith & Mosier, 1984, 1986). The most recent MITRE report (Smith & Mosier, 1986) provides 161 references to earlier documents and empirical studies supporting 944 guidelines.

Another type of guideline is the in-house standard developed to guide interface design within a particular commercial organization; for example, Apple (1985), Burroughs (1986) and Lockheed (1983). The in-house standards vary considerably. Some are relatively application-independent (e.g., Lockheed); on the other hand, a few present detailed descriptions of design rules adopted within particular application domains (e.g., Apple). Guidelines can also be found in books and articles which cover very specific topics such as design of alphanumeric displays (e.g., Tullis, 1983) or data entry displays (e.g., Galitz 1985). In addition, very general discussions of human computer interaction often contain helpful discussions of broad conceptual

design issues (e.g., Martin, 1973; Shneiderman, 1987).

Smith and Mosier (1984) provided an interesting discussion of guidelines and their usefulness. They surveyed a group of professionals involved with system design at a national convention. While more than 80% (N = 83) of the respondents believed that guidelines would have facilitated the interface software design in which they had been involved; 63% (N = 87) reported that no guidelines were used. For instances in which the use of design guidelines were reported, most respondents described the establishment of ad hoc system specific guidelines based on published sources and in-house system design experience.

Table 1 presents a revised listing of guidelines from Tullis (1986). These guidelines vary on several dimensions which can be helpful in assessing their usefulness. These are the presence of screen examples, references to empirical data, and whether the document primarily addresses alphanumeric or graphics.

The same general comments are made repeatedly about guidelines: They are too specific to a particular type of display or system to be applicable to a particular problem, or they are too general to be directly applicable to the design of a given display (Smith & Mosier, 1984; Schneider, 1987; Streveler & Wasserman, 1984; Tullis, 1984, 1986). Most guidelines are entirely subjective, offering little information about their applicability to different user classes or tasks, or about how to evaluate alternative interfaces objectively.

Some guidelines are broad and vague, for example: " (a) Display a small amount of information at one time, (b) Have one idea per display..." (Martin, 1973; p. 316-317). Others guidelines are very narrow, for example:

Table 1. Summary of Documents Containing Screen Design Guidelines*

Author(s)	Summary	Screen Examples?	References to Empirical Data?	Application Domain	# Refs
Apple (1985)	Description of the standard Macintosh user interface, including icons, palettes, pointers, selected objects, windows, scrolling, pull-down menus, dialog boxes, buttons, check boxes and radio buttons, dials, and alerts	Many partial	None	Graphic	None
Bailey (1982) (p. 328-350)	Discussions of grouping techniques, standardization, messages, format, labeling, multi-screen displays, presentation of coded information, screen organization, feedback, highlighting, and cursor design	Several partial	None	Mostly alpha-numeric	**
Burroughs (1986)	About 60 guidelines or standards developed for InterPro software series addressing a standard layout, terminology, letter case, justification, headings, menus, field alignment, check-off lists, and field labeling.	Several complete, Many partial	Some	Alpha-numeric	25
Cropper & Evans(1968)	Discusses color, grouping consistency, and density	None	A few	Alpha-numeric	16
Danchak (1976)	Discusses screen design for power plants, including coding (numeric, textual, shape, color, blinking), density, word length, grouping, and preferred quadrant	Several complete	Some	Alpha-numeric & graphic	12
Engel & Granada (1975)	About 60 guidelines addressing display formats (e.g., highlighting, layout) and content (e.g., feedback, labels, messages)	Several partial	None	Alpha-numeric	32
Galitz (1985)	Several hundred guidelines addressing general screen design (e.g., placement, fonts, messages), data entry screens (e.g., headings, alignment), inquiry screens (e.g., organization, justification), menus, and color	Several complete (insurance)	Some numeric	Mostly alpha-	85

Table 1 (continued).

Author(s)	Summary	Screen Examples?	References to Empirical Data?	Application Domain	# Refs
Lockheed (1983)	About 72 guidelines or standards addressing display areas, conventions, alphabets, numerics, layout, lists, abbreviations, labels, terminology, and use of color	Partial only	None	Mostly alpha-numeric	22
Marcus, Arent, & Brown (1985)	Discussions of formats and grids, typefaces, color, screen, metaphors, sign-on, windows, menus, icons and cursors, animation, sound, and corporate identity	Extensive	None	Graphic	95
MIL-STD-1472C (1981)	About 37 guidelines addressing standard formats, grouping, updating, coding, density, and data presentation	None	None	Alpha-numeric	None
Pakin & Wray (1982)	Discussions of format elements, color, content, layout, style, and word use	Pair of "good" & "bad"	None	Alpha-numeric	None
Peterson (1979)	Discussions of menus, prompts, information presentation, data formats, text, and messages	A few partial	None	Alpha-numeric	None
Pew & Rollins (1975)	About 35 guidelines developed for the U.S. Dept. of Agriculture addressing standard format, menus, data entry forms, output data, and error messages	Many complete	None	Alpha-numeric	None
Smith & Mosier (1984)	About 953 guidelines addressing displays of different data types (text, forms, graphics), density, format, (e.g., organization, grouping), coding (e.g., highlighting, brightness, color), display generation, framing, updating	Several complete	Extensive	Mostly alpha-numeric	161
Stewart (1976)	Discusses 6 screen design factors (logical sequencing, spaciousness, relevance, consistency, grouping, and	Pair of "good" & "bad"	A few	Alpha-numeric	19

Table 1 (continued)

Author(s)	Summary	Screen Examples?	References to Emperical Data?	Application Domain	# Refs
	simplicity) and 6 coding techniques (alphanumeric, color, brightness, spatial, shape, and flashing)				
Tullis (1983)	Review of the literature on four basic characteristics of alphanumeric displays: overall density, local density, grouping, and layout complexity; presents measures for each and applies them	Several complete	Extensive	Alpha-numeric	91
Williges & Williges (1981)	Compilation of about 110 guidelines from 13 other documents addressing information coding (color, shape, brightness), density, labels, format (prompts, tables, graphics, text), and layout	None	When original document did	Mostly alpha-numeric	73

* Adopted from Tullis (1986)

** Screen design references included with references from remainder of book not related to screen design

"In presenting text collections on a small screen, there should be a maximum of 50-55 characters on each line. On larger screens, break up text into two (or more) columns of 30-35 characters per line. Separate columns by at least 5 spaces if the text is not right justified and by 3-4 if the text is right and left-justified." (Engel & Granada, 1975; p. 10)

It is not that the guidelines are wrong. Rather, the problem is that one does not know how to apply them in specific context--for particular users performing particular tasks.

Empirical Studies

There are still relatively few empirical studies of screen design. Screen design is the traditional domain of practitioners and computer programmers. Their strategy is to develop techniques to obtain fast results to design problems, for example, participatory design (e.g., Gould & Lewis, 1985) or rapid prototyping (e.g., Wasserman & Shewmake, 1985).

increasingly, academic and industrial researchers are discovering that the scientific method can be employed fruitfully to the study of screen design. Classic experimental methods of psychology are being refined to deal with the complex cognitive tasks of human performance with computer displays. The fact that basic cognitive theory has application to the area is evidenced in recent articles by Barnes (1987), Howell, Lane, Harvey, and Holden (1986), and Tullis (1984). There is also a large body of evidence from the visual search literature that generalizes to

complex real-world displays. (For a review, see Teichner & Krebs, 1974 and Teichner & Mocharnuk, 1979).

Table 2 is adapted from Tullis (1986) and summarizes a selection of empirical studies on screen design. The table includes information about the type of display variables manipulated (e.g., grouping, density), the type of application (e.g., Space Shuttle, NCR software package), and the key findings. The findings describe any difference in performance or preference measures obtained for the different types of screens. The most common performance measure is search time (i.e., time to extract a specific data item from the screen), and subjectively reported "ease of use" is the most common preference measure.

Most of the studies listed in Table 2 compared performance for two or more methods of displaying the same basic information. Some of the studies manipulated a single physical dimension systematically (e.g., Ringel & Hamner, 1964, Callan, Curren, & Lane, 1977), while others compared screens that differed on several different dimensions (e.g., Dodson & Shields, 1978; Tullis, 1981). Also included in the latter category are studies that compared an existing version of a screen to a redesigned version (e.g., Burns, Warren, & Rudisill, 1986; Keister & Gallaway, 1983). Typically the redesign involved a wide variety of changes to the screens; thus, it is difficult to attribute changes in performance to specific changes in screen design. Several other alternative explanations may also account for performance changes in these studies such as negative transfer (Crowder, 1976) or the Hawthorne effect (Howell & Dipboye, 1986).

To date the most coherent body of empirical evidence for the effect of spatial

Table 2. Selected Empirical Studies of Screen Design.*

Reference	Type of Screens**	Key Findings
Benpasat, Cexter, & Todd (1986)	Tabular (A/N) and graphical presentations of marketing data in color or monochrome; measured decision-making performance	Graphical presentations are more useful when evaluating data in order to detect promising directions; tabular presentations are more useful when the task requires precise computations. Color resulted in fewer iterations to complete the task
Burns, Warren, & Rudisill (1986)	Current and reformatted A/N screens for the Space Shuttle; changes included grouping of related data, use of indentation to indicate subordinate relationships, and consistent use of abbreviations	Reformatted screens resulted in 30% reduction in search time for non-experts and no change in search time for experts; both groups showed greater accuracy with the reformatted screens
Callan, Curran, & Lane (1977)	Six A/N formats for presenting Navy tactical information; varied the number of display items from 6 to 40	Search time increased linearly with the number of items ($r = .95$)
Card (1982)	A/N menus with different semantic bases for groups; alphabetical, functional, or random	Time to select an item from the menu was fastest for the alphabetical grouping and slowest for the random grouping
Dodson & Shields (1978)	Three A/N formats for Spacelab displays: 30%, 50%, and 70% overall density (although other factors varied as well, such as grouping)	Search time increased significantly as overall density increased
Keister & Gallaway (1983)	Original and redesigned A/N screens for an NCR software package; redesigns involved changes in screen format and content, consistency, data entry procedures, error messages, and on-line help	Redesigned screens resulted in 25% less time to complete transactions, 30% less human entry time, 25% reduction in error rate, and 32% reduction in time spent correcting errors
Kolers, Duchnicky, & Ferguson (1981)	Single- and double-spaced displays of text	Single spacing required more eye fixations per line, resulted in fewer words read per fixation, and required longer total reading time
Ringel & Hammer (1964)	Nine A/N formats for tables presenting the status of battlefield units; varied the number of lines of data and the amount of space between the lines	Search time increased with the number of lines of data and decreased slightly with less space between the lines

Table 2 (continued)

Reference	Type of Screens**	Key Findings
Schwartz (1986)	Six A/N formats for presenting the status of systems within a hypothetical power plant; varied local density, grouping, and item alignment; subjects scanned each display and made decisions based on five critical data items	Size of groups of characters and alignment of data items were the best predictors of scanning time; number and size of groups were the best predictors of perceived ease of use
Schwartz & Howell (1985)	Graphic and numeric representations of hurricane tracking data	Under time pressure, subjects reached a better and faster decision using graphic rather than numeric data; under self-pacing, there was no difference
Streveler & Wasserman (1984)	Seven pairs of A/N formats for presenting a variety of data; studied effects of grouping, alignment of data, position on screen, and ordering of lists	Search time decreased significantly with clearly defined groups and well aligned items. Found shortest search times for targets in upper-left quadrant and longest times for targets in the lower-right
Tullis (1981)	Four designs for results of tests on telephone lines: "narrative" A/N, "structured" A/N (with better-defined groups of data), monochrome graphic (with a simple schematic), and color graphic	After practice, time to answer questions about the display was 40% shorter for the "structured" A/N and both graphic formats than for the "narrative" A/N format. Subjects expressed preference for the color graphic format
Tullis (1984)	52 A/N formats for presenting airline and lodging information and 15 A/N formats for presenting information about books; measured 6 characteristics of the formats: overall density, local density, number of groups, size of groups, number of items and item alignment	Multiple regressions using the format variables as predictors accounted for 51% of the variance in search time and 30% of the variance in rated ease of use; best predictors of search time were number and size of groups, and of rated ease of use were local density
Wolf (1986)	Five A/N formats for presenting lists of the names of host computers available in a network; manipulated order of names (alphabetical by row or column and grouping	Shortest search times and best ratings were for a format with clearly defined columns of names, alphabetically ordered within each column
Worchak, Allison, & Dodd (1984)	Fixed 2-dimensional graphical display and simulated user-defined 3-dimensional display of satellite altitude and coverage area	Subjects preferred the 3-D displays, although there were no differences in speed or accuracy of display interpretation

*Adopted from Tullis (1986)

** "A/N " stands for alphanumeric

characteristics is organized around the Tullis (1984) model. Tullis identified six physical dimensions of alphanumeric displays: overall density, local density, number of groups, size of groups, number of items, and layout complexity. The definition of each of these dimensions is given in Table 3. The model represents a significant contribution to the area, synthesizing previous research findings and analytic techniques to provide a model that could be empirically tested. Tullis developed and validated two regression equations which predicted user performance and subjective ratings in two empirical studies. Tullis is not the first to discuss analytic techniques for assessing the effect of spatial characteristics of alphanumeric displays. For example, Streveler and Wasserman (1984) suggested three techniques--e.g., boxing,

Table 3. Definitions of the Six Conceptually Distinct Dimensions Underlying Display Formatting (Tullis, 1984)

<u>Dimension</u>	<u>Definition</u>
Overall Density	Number of filled characters spaces as a percentage of the total spaces available.
Local Density	Average number of filled character spaces near each character.
Grouping	Number of well-defined perceptual groups formed by display characters.
Size of Groups	Average visual angle subtended by the perceptual groups.
Number of Items	Number of data items where a data item is a set of characters on a display line separated by no more than one blank space.
Layout Complexity	The complexity, as defined in information theory, of the distribution of unique horizontal and vertical starting positions of items in the display.

hot spot, and alignment analysis--and their respective quantitative measures--e.g., physical, associative, and aesthetic measures. These measures are conceptually similar to Tullis' measures of grouping, density, and layout complexity, respectively. However, only Tullis (1984) provided empirical evidence to support the validity of his model. The specific methodology used to develop Tullis' model is discussed in the next section. The remainder of this section highlights key findings in the empirical work which supports the model before presenting a detailed discussion of the methodology.

Tullis (1984) reviewed guidelines and empirical studies pertaining to the design of alphanumeric displays. From these sources he identified six physical dimensions of alphanumeric displays (i.e., overall density, local density, number of groups, size of groups, number of items, and layout complexity) and potential analytic techniques for measuring these same dimensions. After synthesizing and refining several display analysis techniques, Tullis wrote a computer program to analyze alphanumeric screens using six metrics which represent the six dimensions mathematically. In two empirical studies, he investigated the extent to which these metrics individually and collectively were able to account for display differences in user performance and subjective ratings. Separate regression models for search time and subjective ratings were developed on the basis of one experiment and validated in another. Tullis also compared the model's predictions with actual data from other published studies. The cross-validated level of prediction was very high ($\approx .8$).

Although layout complexity was the best predictor of subjective ratings, it did not contribute to the prediction of search time beyond what could be predicted by

overall density, local density, number of groups, and size of groups. This is a particularly interesting finding since vertically aligning lists and left-justifying items, practices strongly recommended in guidelines for display formatting (e.g., Engel & Granada, 1976; NASA, 1980; Peterson, 1979), are important factors in Tullis's definition of layout complexity. Thus, the guidelines and Tullis's model lead to conflicting predictions concerning the effect of left-justifying text on user search time.

In a follow-up study, Schwartz (1986) examined the usefulness of four of these metrics (local density, layout complexity, and the number and size of groups) across four task conditions (pattern detection, scanning, multiple cue judgment, and process control). Display variables did indeed affect performance; however, the combinations of dimensions that produced the best performance were not common to all tasks. Exploring the data further, Schwartz partialled out the variance due to scanning time from the performance data of the other three tasks; as a result, the differences among the prediction equations for the three tasks vanished. Display variables no longer predicted performance for pattern detection, multiple cue judgment, or process control. Schwartz concluded that display variables only predict the simple information extraction component of each task. The same dimensions that predicted scanning time--size of groups and layout complexity--predicted subjective ease of use for the experimental displays. Schwartz's findings provide additional evidence (though contradictory to Tullis's findings) concerning the effect of layout complexity on user search time and support the notion that display variables may interact with tasks characteristics.

While the independent causal contribution of density and grouping dimensions is supported by other studies (e.g., Dodson & Shields, 1978; Ringel & Hammer, 1964), there is no empirical evidence directly addressing the effect of layout complexity on human performance besides the Tullis and Schwartz studies. The guidelines suggest that increasing layout complexity may have a detrimental effect on performance. Schwartz's work also presents evidence that layout complexity is an important predictor of performance. However, the evidence provided by Tullis contradicts both of these sources and suggests that layout complexity does not contribute to the prediction of performance. It is clear that the layout complexity dimension is one component of the Tullis model which warrants further investigation. This paper presents two empirical studies investigating conditions which may limit the usability of the Tullis model. The studies investigate the effect of layout complexity on human performance across four information extraction tasks with long-term practice.

The next section discusses the methods used in the two previous studies on the Tullis model (Tullis, 1984; Schwartz, 1986) and highlights the procedural difference which may account for their contradictory findings. The remainder of the introduction presents further justification for investigating the effect of layout complexity, task characteristics, and long term practice on human performance. Finally, the chapter ends with a brief summary of the overall research strategy and design used to address these issues.

Tullis Model: The Development and Refinement of
a Testable Model for Screen Design

Tullis' work

Tullis (1981, 1984) conducted three studies concerned with the development of a "display evaluation tool." Tullis's specific goals were to (1) identify characteristics of displays that had an effect on the user's ability to process information; (2) develop objective measures of these characteristics, and (3) determine regression equations experimentally to predict user performance and preference. The overall objective was to develop a tool that could assess the usability of a computer display without the cost of collecting performance data.

In the first study, Tullis (1981) demonstrated that screen redesign could improve user performance. The redesign strategy involved a variety of techniques (e.g., eliminating unnecessary information and grouping related data). Tullis collected reaction time data for Bell System employees performing realistic query tasks using original and reformatted displays. Employees performed significantly better on the reformatted display than on the original display. However, this strategy--typical of redesign studies in the applied human-computer interface literature (e.g., Burns et al., 1984; Keister & Gallaway, 1983)--made it impossible to determine which redesign technique created the performance differences.

In a second and third study, Tullis (1984) used six display characteristics to predict performance in a realistic database query task. He distilled four dimensions from the guidelines and empirical studies on the design of alphanumeric displays:

Overall density, local density, grouping, and layout complexity. He selected these four dimensions on the following criteria: The characteristics had to be (1) related to the spatial array of characters on a display, (2) objectively defined or have the potential for such a definition, and (3) applicable to any alphanumeric display. Tullis (1984) noted that he found some empirical evidence supporting the effect of the first three dimensions on human performance; however, there was virtually no empirical evidence directly addressing the effect of layout complexity on human performance.

Tullis refined the measurement techniques which were associated with each dimension and combined them in a computer program that mathematically derived a metric for each dimension. Tullis derived two separate metrics for the grouping and layout complexity dimension: number as distinct from size of groups, and number of items as distinct from item uncertainty, respectively. Table 3 gives the definition of each of the resulting six dimensions and Table 4 gives their means, standard deviations, and ranges. To demonstrate the potential usefulness of the six metrics, Tullis applied them to the alphanumeric displays he tested in the first study. Since the two displays differed on all six metrics, he concluded that the six dimensions did indeed capture important aspects of display formats. He then proceeded to conduct a more comprehensive test of the six dimension model.

In the second study, his main objective was to determine regression equations that used the six dimensions to predict search time and subjective ratings. The size and number of groups were the best predictors of search time, while item uncertainty and local density best predicted subjective "ease of use" ratings. The third study validated the two regression equations using a different set of subjects and stimuli; in

Table 4. Means, Standard Deviations, and Ranges of Six Dimensions for the Screens Used by Tullis (Experiment 1; Tullis, 1984)

	Mean	Standard Deviation	Minimum	Range Maximum
Overall Density (%)	27.9	12.0	8.4	62.1
Local Density (%)	50.2	11.7	29.2	71.9
No. of Groups	21.3	22.2	1.0	107.0
Size of Groups (deg)	7.7	5.2	2.0	26.0
No. of Items	86.1	37.0	30.0	202.0
Item Uncertainty (bits)	7.2	0.8	5.9	9.8

addition, the equations were applied to data published in the literature (Callan, Curren & Lane, 1977; Dodson & Shields, 1978; Ringel & Hammer, 1964; Tullis, 1981). Table 5 gives the two regression equations developed in the second study which best fit the actual search time and subjective ratings collected in the third study. Although layout complexity was the best predictor of subjective ratings, it did not contribute to the prediction of search time beyond what could be predicted by overall density, local density, number of groups, and size of groups. With regard to this apparent contradiction, Tullis concluded that "a display that optimizes search time is not necessarily a display that optimizes subjective rating, and vice versa". He suggested that a system designer may need to decide which prediction to optimize, or how to assign relative weights to the predictions.

Table 5. Best-fitting Multiple Regressions Predicting User Search Time and Subjective Ratings (Tullis, 1984).

	t
User search time =	3.38**
1.6996	
-.0209 x Overall Density	-2.28*
+.0286 x Average Local Density	2.43*
+.0201 x Number of Groups	4.01**
+.0914 x Average Visual Angle of Groups	4.95**
$R^2 = .488$	
Subjective Ratings =	-3.79**
-3.1627	
-.0801 x Overall Density	-3.23**
+.0522 x Average Local Density	4.64**
+.0184 x Number of Groups	3.75**
+.0729 x Average Visual Angle of Groups	3.80**
+.0194 x Number of Items	2.37*
+.4121 x Total Layout Complexity	3.60**
$R^2 = .805$	
** $p < .01$ (two-tailed)	
* $p < .05$ (two-tailed)	

Procedural differences between the Tullis and Schwartz study may explain why layout complexity was an important predictor of search time in Schwartz's study but did not contribute to the prediction of search time in Tullis's study. This discussion highlights two key factors in Tullis's procedure; the stimuli and the task. A detailed comparison of the two methodologies is presented in the next section which discusses Schwartz's work. Tullis made no attempt to create a set of displays that would provide orthogonal combinations of the six dimensions. He argued that a high degree of correlation occurs among these six dimensions in naturally occurring displays. Instead Tullis designed his displays to keep the correlation among the six measures to a minimum. (See Table 6.) There were 520 displays (26 formats with 10 examples for each format) in the second study and 150 displays (15 formats with 10 examples for each format) in the third study. The stimulus screens presented three types of data: airplane, motel, and book listings. Subjects viewed a unique display on each trial, but they consistently performed the same task. They searched each display for the correct response to a complex query--e.g., "What is the rate for a room with a double bed at the Topeka Howard Johnson?"--which involved the integration of numerous items on the display. The task was more difficult than a simple search task--e.g., retrieving only one item from the display--, but did not involve the use of a conceptual rule or complex decision making. In the study discussed in the next section, Schwartz (1986) attempted to generalize Tullis' results to other types of tasks that were more complex.

Schwartz's work

Schwartz (1986) examined the usefulness of four of the Tullis metrics (local density, item uncertainty, and the number and size of groups) across four different task conditions. His main objective was to extend the Tullis model to more complex tasks. To accomplish this, Schwartz created a four cell matrix of tasks which he defined by factorially combining two levels of complexity (low/high) with two levels of monitoring load (absence/presence). This conceptualization produced four task conditions: scanning, pattern detection, process control, and multiple cue judgment. Schwartz dichotomized complexity into low and high levels on the basis of the number of operations the task required and the level of detail at which the display stimuli were analyzed. He defined low complexity tasks as requiring simple information extraction and response selection operations (i.e., where the physical characteristics of display stimuli were sufficient to define an appropriate response). In contrast, he defined highly complex tasks as requiring more than simple information extraction and response selection (i.e., where the conceptual nature of displayed stimuli defined the appropriate response). He equated the presence of monitoring load with the requirement that the subject perform the task using a dynamic display--i.e., the data values changed during the monitoring period. When the monitoring load was absent, subjects performed the task using a static display.

Schwartz regressed search time and subjective ratings onto the four dimensions. Display dimensions did affect performance, but the combinations of the dimensions which resulted in the best performance were not common to all four tasks. Moreover, when the variance due to scanning performance was partialled out of the

Table 6. Correlation Matrix of the Six Format Dimensions for the Displays (N=520) Used by Tullis (1984).

	Overall Density	Local Density	No. of Groups	Size of Groups	No. of Items
Local Density	.65	.	.	.	.
No. of Groups	-.04	-.64	.	.	.
Size of Groups	.51	.73	-.59	.	.
No. of Items	.97	.58	.04	.45	.
Item Uncertainty	.70	.46	-.05	.54	.73

total variance accounted for by the various predictive models for the other three task, the differences among the prediction equations for the three tasks disappeared. Display variables no longer predicted performance for pattern detection, multiple cue judgment, or process control. Schwartz concluded that display variables only predicted the simple information extraction component of each task. The best predictors of scanning time were size of groups and item uncertainty. These same dimensions were also the best predictors of subjective "ease of use" for the scanning task.

Several differences in the methodology used by Schwartz and Tullis may be responsible for the conflicting predictions concerning the effect of layout complexity on human performance. The most obvious explanation is the varying degrees to which item uncertainty correlated with the other measures in Tullis' stimulus set as compared with Schwartz stimulus set (Compare Table 6 with Table 7). For several

Table 7. Correlation Matrix of the Four Format Dimensions for the Displays (N = 6)
Used by Schwartz (1986).

	Local Density	No. of Groups	Size of Groups
No. of Groups	-.46	.	.
Size of Groups	.75	-.56	.
Item Uncertainty	.05	.09	.26

reasons (e.g., fewer screens, more control over the formats), Schwarziz was able to minimize effectively the simple correlations between item uncertainty and the other three metrics examined in his study. His stimulus set consisted of only six different displays while Tullis' stimulus set consisted of 520 displays for the second study and 150 displays in the third study.

Conflicting predictions might also be attributed to different procedures. In the Schwartz study, subjects used the same six formats to perform all four tasks during three one hour sessions. This procedure contrasts with Tullis's procedure in which each subject viewed a unique display for each trial. It is difficult to separate the effect of layout complexity from the effect of practice with a particular display in Schwartz's results. The effect of layout complexity on human performance for scanning tasks may not extend to displays which are unfamiliar to the user. Schwartz concluded that item uncertainty played an important role in predicting performance for scanning task displays in which the relative fixed locations of the data items provided reliable information for experienced users. He suggested that item uncertainty may not be

particularly important in the Tullis study because the procedure focused on search for display items in changing locations. Because the location of the data items changed in Tullis' study, their relative locations did not provide much useful information. Finally, Schwartz suggested that it may be worthwhile to base the item uncertainty measure primarily on data items rather than data and label items for the scanning task.

The analysis of methodological differences between the Tullis and Schwartz studies highlights two important variables--task and user characteristics--which may affect the extent to which different display dimensions can predict human performance. The study presented in this paper examines how different characteristics of information extraction tasks (e.g., number of items searched, searching for labels or data values) and user characteristics (e.g., level of practice with a particular display) affect the causal contribution of layout complexity to human performance. The following section discusses each of these variables and also presents a detailed discussion of the layout complexity dimension.

Layout complexity

One of the most common themes found in the guidelines is that the display format should minimize the complexity of the layout, or maximize the visual predictability. It is repeatedly recommended that, based on a knowledge of where some items appears on the screen, one should be able to predict the location of others. Guidelines have encouraged designers to minimize display complexity and maximize predictability of items by using tabular formats (Bailey, 1982, p. 345; NASA, 1980, p. 3-20; Peterson, 1979, p. 34), left-justifying alphanumeric data, and

right-justifying numeric data on the decimal point (Baily, 1982, p. 346; Engel & Granada, 1976, p. 7; Galitz, 1980, p. 113; Lockheed, 1983, p. 2-6 & 2-10; Sidorsky et al., 1984, p. 211-13; Smith & Mosier, 1986, p. 127).

The effect of all these recommendations is to increase the predictability of the locations of items on the display. For example, in a tabular format, the location of any entry in the table can be predicted from a knowledge of the column and row positions. By contrast, the location of any item in a textual format cannot be predicted as efficiently. The guidelines focus on reducing layout complexity by aligning data items vertically and horizontally.

Despite these recommendations, Tullis (1984) pointed out that there were virtually no empirical data regarding the utility of tabular formats. However, Tullis (1984) did identify one objective method for measuring layout complexity (Bonsiepe, 1968) and revised the equation for inclusion in his computer analysis program. Bonsiepe (1968) described the complexity of typographically designed pages as a function of (1) the unique widths and heights of item groups (used to derive a measure of "system order") and (2) the unique horizontal and vertical starting positions of these groups from some arbitrary reference point (used to derive a measure of "distribution order"). Specifically, Bonsiepe defined layout complexity as the product of the number of groups and the average location uncertainty based on system and distribution order.

Tullis noted some problems using Bonsiepe's measure of system order. First, the grouping of display items seemed somewhat arbitrary. Secondly, the width and height of groups was not clearly defined by Bonsiepe. Although he dropped system order from further consideration, Tullis did adopt the concept of distribution order—it made

intuitive sense that the speed of search and encoding of display items would be determined, at least in part, by their starting points. In order to circumvent problems in defining the perimeters of groups, distribution order was modified by focusing on individual display items (labels and data) which Tullis defined operationally as adjacent characters with fewer than two blank spaces between them. The basic entropy equation:

$$H = -\sum_{n=1}^m (p_n \log_2 p_n)$$

where:

- H = Average location uncertainty of individual display items,
expressed in bits
- m = Number of event classes (i.e., no. of unique vertical or
horizontal starting positions)
- p_n = Probability of occurrence in the nth event class (based on the
frequency of events within that class)

was applied to both vertical and horizontal starting positions. Of course, this equation only defines a measure of average item uncertainty rather than total layout complexity. As a result, Tullis's (1984) measure of layout complexity was composed of two components--number of items and item uncertainty. Number of items is a measure of how many data items or labels are on the display. Item uncertainty is the average uncertainty of any one item's location, based on the number of unique positions of items on the display and the frequency of occurrence of each of those positions. As used by Bonsiepe (1968), these two components were multiplied together to yield the total amount of information (in bits) conveyed by the layout of all the items on the display.

In Tullis' study, the two components were measured separately to determine which component was a better predictor of display usability.

As discussed previously, layout complexity was the best predictor of subjective ratings in Tullis's study; however, it did not contribute to the prediction of search time beyond what could be predicted by overall density, local density, number of groups, and size of groups. In contrast, Schwartz (1986) identified layout complexity as one of the best predictors of scanning time and subjective ratings. These two studies provide correlational evidence in support of numerous guideline recommendations that layout complexity is an important spatial characteristic of alphanumeric displays. However, due to the manner in which stimulus sets were designed in the Tullis (1984) and Schwartz (1986) studies, it is difficult to determine the unique causal contribution of layout complexity to human performance and preference independent of its correlation with the other dimensions in these two studies. Moreover, closer examination of the procedural and task differences between the two studies suggest that the effect of layout complexity may be task dependent. Clearly, the layout complexity dimension is one component of the Tullis model which should be manipulated independently if its contribution to human performance is to be determined. It was to this end that the present research was directed. Layout complexity was manipulated by creating screens of identical amounts of information which had equivalent values for all dimensions with the exception of layout complexity. Thus it is possible to determine the causal contribution of layout complexity on human performance. The effect of layout complexity was also investigated across four different tasks in long-term practice conditions. The next sections discuss the importance of task considerations.

Task Characteristics

The moderating influence of task characteristics on many design variables is becoming increasingly clear (e.g., Hammond, 1980; Hitt, 1961; Kahneman & Treisman, 1984). It is recognized that some ways of organizing, presenting, and coding information (e.g., color, graphics, highlighting) are superior to others; however, the amount and nature of the advantage is highly dependent on the perceiver's task (Hitt, 1961; Schwartz & Howell, 1985). Unfortunately, even human factors professionals often neglect the impact of task variables. Such neglect is evident in most design guidelines (e.g., Bailey, 1982; Engel & Granada, 1975; Smith & Mosier, 1986). Screen designers must rely on guidelines or empirical findings without full knowledge of their generalizability. This uninformed design strategy can lead to serious design errors. Similarly, the use of Tullis' prediction model to determine design criteria in other than data base query tasks may lead to serious operational consequences or at best wasted effort. Schwartz (1986) provided evidence that the effect of display variables as defined by Tullis' model may be task dependent. In Schwartz's study, display variables only predicted human performance for the information extraction component of tasks. Schwartz's findings support the notion that display variables may interact with task characteristics and that the generalizability of Tullis' findings may be limited to information extraction tasks. Furthermore, Schwartz's predictions concerning the effect of layout complexity on human performance for information extraction tasks contradict Tullis' predictions. Seemingly unimportant aspects of tasks and procedure appear to have interacted with display variables producing conflicting predictions. A closer examination of relevant task dimensions for information extraction tasks seems

warranted.

The taxonomic approach has long been the traditional methodology for exploring the task-specificity of human performance (Gagne, 1962; Melton & Briggs, 1960; Miller, 1962). Melton and Briggs (1960) called for the development of a general task taxonomy that would permit the integration of human performance data into a useful, predictive tool. However, the response to their plea was not a single general taxonomy, rather many different taxonomies each highlighting different parameters or dimensions by which to analyze and categorize tasks. (See Companion and Corso (1982), Fleishman and Quaintance (1984), and Meister (1985) for summaries.)

Discrepancies between the various approaches arose in the selection of the critical dimensions. For example, tasks have been categorized on the basis of (1) the inherent abilities or traits they require, (2) the specific behaviors they require, (3) the pattern of stimuli to which humans respond, (4) the explicit behaviors which are observed while the task is performed, and (5) the cognitive level of functioning required by the operator to perform the task. Although each approach had certain merits, no single taxonomy answered the challenge set forth by Melton and Briggs--i.e., to develop a taxonomy that integrates human performance into a useful tool. Companion and Corso (1982) noted that most efforts were devoted solely to the development of a taxonomy thereby producing nothing more than a useless trivia list. They proposed that a classification system for tasks is a useful predictive tool only when it includes information concerning the way the task affects operator performance and the way the task is affected by stimulus and response constraints. These authors concluded that more effort should be devoted to constructing empirical relationships which can later be

integrated into a taxonomy.

The majority of taxonomies which address human-computer interaction fail to provide a general taxonomy which integrates human performance data. The major reason most taxonomies fail is that there is little human performance data available for human computer tasks. Although not a single taxonomy could be found which integrates performance data concerning human-computer interaction, one taxonomy (Lenorovitz, Phillips, Ardrey & Kloster, 1984) did at least provide a general framework within which to begin examining information extraction tasks. Lenorovitz, Phillips, Ardrey, and Kloster (1984) proposed a taxonomy of users tasks for a wide range of applications consisting of four separate categories: computer internal, computer output, user internal, and user input. The taxonomy extended the earlier work of Berliner, Angell, and Shearer (1964), directly incorporating many of their tasks and explicitly defining each task. The user-internal taxonomy arranged three major classes of tasks central to the user into a hierarchical structure: perception, mediation, and communication. Perception was defined as the process of getting information into the human system as well as some initial level of recognition, classification or identification of that information. Several of the tasks included in the perception category appeared to be central to the type of information extraction tasks which are the focus of this paper, for example:

Search	Purposeful exploration or looking for specified item(s).
Scan	Glance over quickly, usually looking for overall patterns or anomalous occurrences (not details).
Cross-Reference	Accessing or looking up related information usually by

means of indexing organized structured scheme set up for that purpose.

Another useful approach to the classification of tasks was to describe underlying dimensions. For example, the Farina and Wheaton (1973) taxonomy separated tasks components (goals, response, procedure, stimulus, and stimulus-response) into a number of task characteristics for each component (e.g., complexity of output units, response rates) which could be used to describe any task. This taxonomy can be used to determine why a task presents difficulty to the operator and permits comparisons of tasks from different systems in terms of underlying commonalities. However, it could not identify specific tasks, and consequently, must always be used in combination with another descriptive taxonomy such as Lenorovitz et al. (1984).

The general task framework suggested by Lenorovitz et al. (1984) combined with the underlying dimensions approach (Farina and Wheaton, 1973) loosely represents the strategy used to create the four information extraction tasks examined in this study. The focus of the study was information extraction tasks similar to the search, scan, and cross reference tasks defined by Lenorovitz et al (1984) and the look-up and scan tasks used by Tullis (1984) and Schwartz (1986). The strategy was to identify basic dimensions of tasks which may interact with the effect of layout complexity on human performance. In his conclusion, Schwartz pointed to two potentially important task dimensions--familiarity with the display and type of target. Schwartz concluded that layout complexity may only predict scanning performance for fixed displays, i.e., displays in which the relative fixed location of data items provide

reliable information for experienced users. Furthermore, he suggested that because the relative location of data items alone seemed to influence scanning performance, it may be worthwhile to base the layout complexity measure primarily on data items rather than data and labels for fixed formats.

The importance of the interplay between these task dimensions--type of target (label or data) and familiarity with the stimulus displays (level of experience)--was also supported by Burns et al. (1985). These authors investigated the effect of user experience, type of question and display redesign on response time for two Space Shuttle displays. They compared experts, who worked frequently with Space Shuttle displays (flight controllers, trainers, and astronauts), with non-experts, who had no previous experience with these displays (technicians and clerks). Both expert and nonexpert users performed two types of tasks using two original and two redesigned Space Shuttle displays: to search for a label and report a data value associated with the label (label to data), and to search for a data value and report the label associated with that value (data to label). The redesign of the two screens included several changes (e.g., functional grouping, alignment of numbers by decimal point, highlighting). On the original displays, nonexperts responded more quickly to data to label questions than label to data questions while experts demonstrated the opposite pattern. However, there were no differences between the two question types on the reformatted displays for either nonexperts or experts. These findings point to the importance of considering level of experience with a particular display format and its potential impact on performance for different types of tasks. It is important to note that in the Burns et al. study level of experience not only indicated familiarity or practice with a particular display, but also

After considering the evidence, type of target (label versus data items) was selected as one important task dimension. The fixed vs. variable formats dimension as distinguished by Schwartz was reconceptualized as a user characteristic (level of experience or amount of practice with a particular display) in a manner described in the next section. In addition to type of target item, two other considerations were selected: number of target items (e.g., Drury & Clement, 1978; Kahneman & Triesman, 1984) and task complexity, i.e., search or comparison (e.g., Hitt, 1961; Neisser, 1967). With these task considerations in mind four information extraction tasks were created. They are defined as follows:

- (1) Find labels: given a label report a data value
- (2) Scan data: given a data value report a label
- (3) Compare labels: given two labels compare their data values and report the correct label that fits a specified rule
- (4) Compare data: given a data value, compare all appropriate data values and report the correct label that fits a specified rule.

Findings from many different areas of basic research (e.g., visual search, attention, memory, learning) have demonstrated the importance of long-term practice across a variety of tasks and task combinations(e.g., Bahrnick, Noble, & Fitts, 1954;

Hirst, Spelke, Reaves, Coharach, & Neisser, 1980; Neisser, Norick, & Lozar, 1964; Schneider & Shiffrin, 1977). Researchers have attributed the effect of long term practice to many different processes and mechanisms--e.g., automatic processing (Schneider & Shiffrin, 1977), reduced resource demand (Wickens, 1984), or the "motor program" --e.g., automaticity of the response end of processing (Keele, 1973). Regardless of the process or mechanism cited, the results are the same: human performance improves with more practice.

While the effect of long term practice is well documented in the basic literature, few studies have investigated its effects on human performance with computer displays. Several studies on screen design have contrasted level of experience, novice vs. expert users (Burns et al., 1985; Siegel & Fischel, 1971); but, these studies have typically confounded practice on a particular display with semantic or conceptual understanding of the information presented on the display. There is one exception, however. In a series of studies, Christ and Corso (1983) investigated the specific effect of extensive practice on the relative effectiveness of different visual display codes. They conducted nine experiments in which highly practiced subjects performed different tasks using various coding dimensions (letters, digits, familiar geometric shapes, and colored dots). Subjects used single code and dual code displays in three isolated tasks (choice reaction, search-locate, and identification-memory) and in a multiple task that combined the three tasks in an irregularly alternating sequence. Across all nine experiments, extensive practice minimized differences in performance that were initially due to code and that could have been attributed to exposure duration, response interval, luminance, and task uncertainty. These findings have far-reaching

implications for screen design. The authors suggested that when selecting the appropriate visual code, the designer should first decide whether the goal is the enhancement of long-term or short-term task performance. If short-term performance enhancement is the objective and inexperienced operators are to be used, manipulation of visual codes improve human performance. However, if long-term performance enhancement is the goal then the choice of visual code may not make a significant difference. Furthermore, Christ and Corso recommended conducting a thorough analysis of task requirements before selecting visual display codes because different underlying psychological processes for different tasks often dictate the selection of different visual codes.

Evidence such as that provided by the basic literature and the Christ and Corso study points to the importance of considering long term practice in screen design and the need for further research. The two studies discussed in this paper provide more evidence regarding the effect of long term practice. The next section presents a brief summary of the overall research design used in the two studies to investigate the effect of long term practice, layout complexity, and task characteristics.

Research design

The two studies manipulated the layout complexity dimension creating screens of identical amounts of information which had equivalent values for all dimensions in the Tullis model with the exception of layout complexity. Three types of experimental screens with varying degrees of layout complexity were used to determine the effect of layout complexity on human performance. The first study compared human

performance using a left-justified format with a format that was not left-justified but predictable. In the second study, subjects also used a third type of screen on which the starting position of items was arranged randomly within constraints. To increase the generalizability of the research, four different information extraction tasks were used in long-term practice situations.

Experiment I

Method

Experimental Screens. Eight screens were chosen from Tullis' (1984) motel stimuli set. The motel screens contained listings of motels for a particular city and state. Cities were listed alphabetically, and within each city the motels were listed alphabetically by name. Other kinds of information included were area code, telephone number, single room rate and double room rate. Two examples of the motel screens are shown in Figure 1.

In Tullis' stimulus set, there were 26 different formats and 10 examples of each format containing different kinds of data (i.e., different cities and towns, hotel names and prices). The 26 formats represented different combinations of a range of values for each dimension. These different combinations created 26 distinct spatial layouts. Table 4 shows the means, standard deviations, and ranges of the values on the six dimensions for the 520 screens used in Tullis' study.

In the present study, eight examples of screen formats were chosen by randomly selecting eight formats from Tullis' set of 26, and then randomly choosing one of each selected format. To create two types of experimental screens representing two different levels of layout complexity, the eight screen examples were redesigned. The objective of the redesign was to increase the layout complexity value yet maintain equivalent values on the other dimensions and the identical amount of information. This was accomplished by increasing the unique number of vertical (i.e., columnar) starting positions for data and label items within previously defined perceptual groups (as defined by Tullis' program). The starting positions of items in the original screen

Maryland

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Laurel	Holiday Inn	301	498-0900	\$40	\$48
	Quality Inn Colony 7		725-5100	\$34	\$40
Oakland	Alpine Village	301	387-5534	\$36	\$39
	Lake Breez		387-5564	\$36	\$40
	Lake Side Motor Court		387-5566	\$38	\$42
	Oak-Land		334-2171	\$21	\$26
	Town		334-3955	\$25	\$30
	Will o' the Wisp		387-5503	\$47	\$59
Ocean City	Carousel Hotel	301	524-3464	\$79	\$95
	Quality Inn Boardwalk		289-4401	\$74	\$90
	Quality Inn Stardust		289-6444	\$75	\$75
	Sahara		289-8101	\$52	\$62
	Sheraton Fontainbleau		524-3535	\$35	\$45
	Surf and Sands		289-7161	\$28	\$60

Maryland

Laurel	Area Code: 301				
Holiday Inn	Phone: 498-0900	Single: \$40	Double: \$48		
Quality Inn Colony 7	Phone: 725-5100	Single: \$34	Double: \$40		
Oakland	Area Code: 301				
Alpine Village	Phone: 387-5534	Single: \$36	Double: \$39		
Lake Breez	Phone: 387-5564	Single: \$36	Double: \$40		
Lake Side Motor Court	Phone: 387-5566	Single: \$38	Double: \$42		
Oak-Land	Phone: 334-2171	Single: \$21	Double: \$26		
Town	Phone: 334-3955	Single: \$25	Double: \$30		
Will o' the Wisp	Phone: 387-5503	Single: \$47	Double: \$59		
Ocean City	Area Code: 301				
Carousel Hotel	Phone: 524-3464	Single: \$79	Double: \$95		
Quality Inn Boardwalk	Phone: 289-4401	Single: \$74	Double: \$90		
Quality Inn Stardust	Phone: 289-6444	Single: \$75	Double: \$75		
Sahara	Phone: 289-8101	Single: \$52	Double: \$62		
Sheraton Fontainbleau	Phone: 524-3535	Single: \$35	Double: \$45		
Surf and Sands	Phone: 289-7161	Single: \$28	Double: \$60		

Figure 1. An example of two screens used by Tullis

examples were left-justified within groups. Although items on the redesigned screens were not left-justified, the starting positions of the items were ordered rather than random. The ordered pattern of the label and data items provided a degree of certainty; that is, starting positions were not left-justified but they were predictable. Knowing the starting position of any particular label and data item provided information about the starting position of an adjacent item. Thus, two types of experimental screens were created left-justified and predictable screens. (See Figure 2 for examples.) The only appreciable difference between the two types of screens was the value of the layout complexity dimension. In addition to maintaining equivalent values on the other four dimensions (e.g., overall density, local density, number and size of groups), "number of items" was identical for each type of screen. Table 8 shows the means, standard deviations, and ranges on each of the six dimensions for the eight left-justified screens and the eight predictable screens. (See Appendix A for the complete set of screens.) The average increase in the layout complexity value was 1.91 bits. This average was calculated using the difference between the value for the predictable screen type minus that for the left justified screen type for each of the eight screen examples.

Experimental Tasks. Each subject performed four information extraction tasks: find labels, scan data, compare labels, and compare data. Each experimental task involved a different type of question which required the subject to locate and report different types of information (label or data) using the screens. (See Figure 3 for examples of each type of question.) "Find" questions gave the label (e.g., the name of a hotel in a certain city) and asked him to report the appropriate data value (e.g., single room rate). "Scan" questions gave a data value (e.g., a double room rate) and asked the

Maryland

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Laurel	Holiday Inn	301	498-0900	\$40	\$48
	Quality Inn Colony 7		725-5100	\$34	\$40
Oakland	Alpine Village	301	387-5534	\$36	\$39
	Lake Breez		387-5564	\$36	\$40
	Lake Side Motor Court		387-5566	\$38	\$42
	Oak-Land		334-2171	\$21	\$26
	Town		334-3955	\$25	\$30
	Will o' the Wisp		387-5503	\$47	\$59
Ocean City	Carousel Hotel	301	524-3464	\$79	\$95
	Quality Inn Boardwalk		289-4401	\$74	\$90
	Quality Inn Stardust		289-6444	\$75	\$75
	Sahara		289-8101	\$52	\$62
	Sheraton Fontainbleau		524-3535	\$35	\$45
	Surf and Sands		289-7161	\$28	\$60

Maryland

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Laurel	Holiday Inn	301	498-0900	\$40	\$48
	Quality Inn Colony 7		725-5100	\$34	\$40
Oakland	Alpine Village	301	387-5534	\$36	\$39
	Lake Breez		387-5564	\$36	\$40
	Lake Side Motor Court		387-5566	\$38	\$42
	Oak-Land		334-2171	\$21	\$26
	Town		334-3955	\$25	\$30
	Will o' the Wisp		387-5503	\$47	\$59
Ocean City	Carousel Hotel	301	524-3464	\$79	\$95
	Quality Inn Boardwalk		289-4401	\$74	\$90
	Quality Inn Stardust		289-6444	\$75	\$75
	Sahara		289-8101	\$52	\$62
	Sheraton Fontainbleau		524-3535	\$35	\$45
	Surf and Sands		289-7161	\$28	\$60

Figure 2. An example of an experimental screen contrasting left-justified (lower panel) and predictable (upper panel) formats

Table 8. Means, Standard Deviations and Ranges of the Six Dimensions for the Left-justified and Predictable Screens Used In Experiment I.

<u>Left-justified Screens</u>				
	Means	Standard Deviations	Minumum	Range Maximum
Overall Density (%)	22.7	5.9	11.9	30.7
Local Density (%)	50.0	7.9	36.4	63.4
No. of Groups	16.0	6.7	4.0	24.0
Size of Groups (deg)	6.5	2.3	3.4	10.6
No. of Items	65.5	16.3	37.0	89.0
Layout Complexity (bits)	6.6	0.4	5.9	7.01
<u>Predictable Screens</u>				
	Means	Standard Deviations	Minumum	Range Maximum
Overall Density (%)	22.7	5.9	11.9	30.7
Local Density (%)	48.7	7.7	35.9	62.0
No. of Groups	15.8	6.7	4.0	24.0
Size of Groups (deg)	6.5	2.4	3.5	10.8
No. of Items	66.5	16.0	37.0	89.0
Layout Complexity (bits)	8.5	0.7	7.2	9.5

Task	Definition and Example
<u>Finding:</u>	<p>given a single label, reports its data value.</p> <p><u>ex.</u> What is the single room rate of the Alpine in Oakland?</p>
<u>Scanning:</u>	<p>given a single data value, report the correct label.</p> <p><u>ex.</u> What hotel in Ocean City has a double room for \$40 a night?</p>
<u>Comparing Labels:</u>	<p>given two labels, compare their data values and report the label that fits the specified rule.</p> <p><u>ex.</u> Which hotel in Oakland has the cheaper single room Town or Alpine?</p>
<u>Comparing Data:</u>	<p>given a data value, compare all data values and report the label that fits the specified rule.</p> <p><u>ex.</u> Which hotel in Ocean City has a double room for under \$50?</p>

Figure 3. Examples of questions for each of the four information extraction tasks.

subject to report the label (e.g., the hotel) associated with that data value. A "compare label" question gave two hotel names in a certain city and asked the subject to compare these two labels and the data values associated with the label in accordance with a given rule (e.g., which hotel is cheaper or more expensive). The "compare data" questions asked the subject to scan data values in a certain city, search for the correct value as determined by a given rule (e.g., the most expensive or the cheapest), and report the label associated with the correct value.

There were eight questions for each type of task for a total of 32 questions for each of the eight formats. The same set of task questions was used for each type of screen (left-justified or predictable) since the information content was identical.

Subjects and Apparatus

Eight undergraduates from Rice University earned \$4.00 per hour for their participation as subjects. All had normal or corrected-to-normal vision.

Apple Macintosh computers with 512k memory were used for stimulus presentation and response recording. All stimuli were presented in the Macintosh's 12-point Monaco textfont which is constantly spaced and has a height to width ratio (2:1) identical to the textfont used by Tullis (1984). Finally, the computer program used for stimulus presentation and data collection was written in Basic supplemented by programs from the CLR ToolLib™ assembly language library.

Design and Procedure

The experimental design consisted of a factorial combination of two types of experimental screens (left-justified or predictable), four tasks (find, scan, compare label, or compare data) and four practice sessions with repeated measures on all

factors.

Each of the eight subjects was paired with a different screen example to create eight unique subject-screen pairs. In each of four one-hour sessions conducted on four different days within a seven day period, subjects performed 16 trials for each of four tasks using one type of a particular screen example (left-justified or predictable) and then performed 16 trials for each task using the other type of the screen. Subjects performed a total of 128 trials per session. Different data values were randomly generated for each trial so that the subject could not memorize the correct answer to each task question. One of eight different questions for each task was selected randomly for each trial. The order of tasks and presentation of the two types of experimental screens were counterbalanced across subjects.

Subjects were instructed to stress accuracy and take rest breaks as needed. The subject read a brief set of instructions, hit the space bar as instructed, and the first trial began. Each trial consisted of three parts. First the task question appeared. When the subject understood the question he/she hit the space bar. The question disappeared, and an experimental screen appeared. The subject searched the information presented for the answer to the question and then hit the space bar a second time upon locating the answer. This second space bar press caused the screen to disappear and the question to reappear, along with two multiple-choice answers. The subject selected one of the answers with the mouse and the next trial began. Each session was completed in approximately 15 to 20 minutes. Search time was measured from the appearance of the screens until the subject pressed the space bar causing the screen to disappear.

Results

Subjects performed at a 97.7% accuracy rate. Due to the high degree of accuracy, error trials were neither analyzed nor eliminated. (See Appendix D for raw error data.) Median search times were analyzed using an ANOVA design with repeated measures. (See Appendix B for the summary of the ANOVA tables.)

As can be seen in Figure 4, search times for left-justified screens ($M = 3.99$) were not faster than for predictable screens ($M = 4.06$). There was also no tendency for type of screen to interact with practice session.

There were significant differences in search time for the four different tasks. "Find" was the fastest ($M = 2.32$ sec) followed by "scan" ($M = 2.70$ sec), "compare labels" ($M = 4.93$ sec), and "compare data" ($M = 6.16$ sec), $F(3, 21) = 8.64$, $p = .001$. In a post hoc analysis using Tukey's method (Winer, 1971), "find" was significantly different from "compare Labels" and "compare data", but not from "scan" at $p = .05$. "Scan" was significantly different from "compare data" but not from "compare labels" at $p = .05$.

Subjects improved for all tasks across the four sessions, but there was no significant interaction between practice and tasks (first session, $M = 4.84$; second session, $M = 4.85$; third session, $M = 3.28$; fourth session, $M = 3.13$), $F(3, 21) = 3.84$, $p = .05$.

Discussion

Violation of the guidelines that recommend left-justified format for alphanumeric displays did not increase user search time in any of four basic information extraction tasks. These results are consistent with Tullis' predictions and

Figure 4. Means of median reaction time (seconds) in Experiment I as a function of task and layout complexity.

support the contention that layout complexity is not a causal factor in determining search time for alphanumeric displays. Though layout complexity did not predict search time in Tullis' prediction model, it did correlate with search time when considered alone in the Tullis study. Since independently manipulating layout complexity (i.e., holding the values of the other dimensions constant) did not affect search time, the findings from Experiment I suggest that the simple correlation of layout complexity and search time in the Tullis study may be due to layout complexity's correlation with the other five dimensions.

Experiment II

The next logical step was to attempt to replicate the present results using a third type of experimental screen that displayed the starting positions of items in a completely unpredictable fashion. The basic design was similar to Experiment I. Subjects performed the same four tasks in four practice sessions. However, in addition to the two types of screens used in Experiment I, subjects viewed a third type of experimental screen on which the starting positions of items were arranged randomly within constraints. (see Figure 5.)

In Experiment II, subjects performed 64 additional trials per session and received more massed practice (i.e., more consecutive trials performing the same task using the same type of screen) with each type of screen than in Experiment I. In addition, subjects made subjective ratings on all three types of experimental screens for all eight screen examples. (See Appendix A for the complete set of experimental screens.)

Method

Experimental Screens. Three types of experimental screens were used in Experiment II: the left-justified and predictable screens used in Experiment I and a third "random" screen on which the starting positions of items were arranged randomly within constraints. Since the objective was to maintain the identical amount of information in the same physical space, the entire range of columns was not available as possible starting positions for the randomization process. Starting positions were assigned randomly within the constraint that the string of information (i.e., label and data items) would fit within the 80 column width of the screens. (See Figure 5.) Table

Maryland

City	Motel/Hotel	Area Code	Phone	Rates Single Double	
Laurel	Holiday Inn	301	498-0900	\$40	\$48
	Quality Inn Colony 7		725-5100	\$34	\$40
Oakland	Alpine Village	301	387-5534	\$36	\$39
	Lake Breez		387-5564	\$36	\$40
	Lake Side Motor Court		387-5566	\$38	\$42
	Oak-Land		334-2171	\$21	\$26
	Town		334-3955	\$25	\$30
	Will o' the Wisp		387-5503	\$47	\$59
Ocean City	Carousel Hotel	301	524-3464	\$79	\$95
	Quality Inn Boardwalk		289-4401	\$74	\$90
	Quality Inn Stardust		289-6444	\$75	\$75
	Sahara		289-8101	\$52	\$62
	Sheraton Fontainbleau		524-3535	\$35	\$45
	Surf and Sands		289-7161	\$28	\$60

Figure 5. An example of an experimental screen in a random format.

Table 9. Means, Standard Deviations and Ranges of the Six Dimensions for the Random Screens Used In Experiment II.

	Means	Standard Deviations	Minumum	Range Maximum
Overall Density (%)	23.7	4.8	17.4	31.9
Local Density (%)	45.2	9.0	29.7	62.0
No. of Groups	17.5	8.7	1.0	29.0
Size of Groups (deg)	10.0	5.8	5.7	22.7
No. of Items	65.5	16.0	39.0	91.0
Layout Complexity (bits)	8.8	0.5	8.0	9.5

9 shows the means, standard deviations and range for each of the six dimensions for the eight "random" screens.

Subjects and Apparatus. Sixteen undergraduates from Rice University earned \$4.00 per hour for their participation as subjects. All had normal or corrected-to-normal vision. The same Apple Macintosh computers, screen dimensions, fonts, and programs were used as in Experiment I. The program was revised to accommodate procedural changes.

Design and Procedure. The experimental design consisted of a factorial combination of three types of screens (left-justified, predictable, or random), four tasks (find, scan, compare label, or compare data), and three practice sessions with repeated measures on all factors.

Each of 16 subjects was paired with one of eight screen examples to create 16 subject-screen pairs. Subjects participated in four one-hour sessions within a seven

day period. The first day was a training session. The procedure was identical to Experiment I except that subjects also performed 16 trials for each task using the "random" screen. Data from the first session were not included in the analyses.

During each of the second, third, and fourth practice sessions, subjects used a different type of screen (aligned, predictable, or random) for the the entire session. In each session, they performed 16 trials for each of the four tasks, repeating that sequence three times for a total of 292 trials per session using the same screens. The same task questions were used and each trial proceeded in the same fashion as in Experiment I.

At the end of the fourth session, subjects rated the entire set of 24 stimuli (8 screen examples x 3 types of screens) using two seven-point Likert scales. On one scale subjects rated how quickly the tasks could be performed using a particular type of screen (e.g., 1--very slowly; 7--very quickly). On the other scale they rated how much they liked the way the screen was arranged (e.g., 1--strongly dislike; 7--strongly like).

Results

Search time. Subjects performed at a 98.7% accuracy rate. Due to the high degree of accuracy, error trials were neither analyzed nor eliminated. (See Appendix D for raw error data.) Median search times were analyzed using an ANOVA design with repeated measures. (see Appendix B for the summary of the Anova Tables.)

As can be seen in Figure 6, search times for the three types of screens were not significantly different; left-justified screens ($M = 2.85$), predictable ($M = 2.68$) or random screens ($M = 2.99$). There was no tendency for type of screen to

Figure 6. Means of median reaction time (seconds) in Experiment II as a function of task and layout complexity.

interact with practice session.

There were significant differences in search time for the four different tasks. "Find" was the fastest ($M = 1.73$) followed by "scan" ($M = 2.03$), "compare label" ($M = 3.43$) and "compare data" ($M = 4.17$), $F(3, 45) = 33.66$, $p = .001$. In a post hoc analysis using Tukey's method (Winer, 1971), "find" was significantly faster than "compare Labels" and "compare data" but not significantly different from "scan" at the $p = .05$ level. "Scan" was also significantly faster than "compare labels" and "compare data" at the $p = .05$ level. "Compare labels" did not differ significantly from "compare data".

There were also significant differences for search time of the four tasks using different types of screens, $F(6, 90) = 6.93$, $p = .001$. While the search times for the find, scan, and compare label tasks were not significantly different using the three types of screens, search times for the compare data task using the left-justified ($M = 4.02$) and the predictable ($M = 3.73$) screens were significantly faster than using the random screen ($M = 4.76$), $F(2, 30) = 6.77$, $p = .004$.

Subjects improved for all tasks across the four sessions, but there was no significant interaction between practice and tasks (first session, $M = 3.02$; second session, $M = 2.76$; third session, $M = 2.79$), $F(2, 30) = 8.84$, $p = .001$.

Subjective Ratings. Both types of subjective ratings were analyzed using a one-way repeated measures ANOVA. (See Appendix B for the summary of the ANOVA tables for like and quickness ratings respectively.) Subjects preferred the left-justified screen ($M = 6.13$) as compared to the predictable ($M = 5.4$) or random ($M = 2.27$) screens, $F(2, 28) = 33.55$, $p = .0001$. Subjects also rated that their

performance would be faster on the left-justified screens ($M = 5.87$) as compared to the predictable ($M = 5.67$) or random ($M = 3.87$) screens, $F(2, 28) = 18.02$; $p = .001$.

Discussion

Moderate (predictable) violations (i.e., the predictable format) of the widely acceptable principle that alphanumerics should be left-justified did not increase user search time across the four tasks, thereby replicating the findings from Experiment I. Furthermore, even severe violations of the left-justified format recommendation (i.e., random assignment of starting position within constraints) did not increase user search time for three tasks (find label, scan data, and compare labels). However, when subjects compared multiple data values and selected a value according to a conceptual rule, the random format did increase user search time. Though performance using the three experimental screens was comparable across the four tasks with only this one exception, subjective ratings demonstrated pronounced differences among the three formats. Subjects disliked the random format and felt that they performed poorly using the random screens. Unfortunately, the ratings were not made separately for each of the four tasks; as a result it is impossible to compare performance on a particular task directly with the subjective ratings for a particular task.

While increasing the number of unique starting positions of items on predictable screens increased the layout complexity index, the uncertainty of the starting positions was not increased by presenting items in an ordered pattern. The context of the ordered pattern of starting positions may have made it possible for subjects to predict the starting position of one item from the starting position of

adjacent items. This alternative explanation could explain the equivalent user search times for the four tasks using the left-justified and predictable formats in Experiment 1. However, the results of Experiment 2 demonstrated that for three out of four tasks even random starting positions did not increase user search time.

Performing the "compare data" task using the random format did increase user search time in comparison to performing the "compare data" task using the left-justified or predictable format. The compare task was the most difficult task requiring more than simple information extraction and response selection. It required scanning multiple data values in accordance with a specified rule. The conceptual nature, rather than the physical nature, of the items defined the appropriate response. This task is very similar to the scanning task used in the Schwartz (1986) study for which layout complexity was the best predictor of performance. Schwartz's scanning task required the subject to scan data values for several critical display items and compare them according to a conceptual rule. The look-up in Tullis's task required subjects to search for a single label value and report the corresponding data value. Layout complexity was not an important predictor of search time in the Tullis study. The findings from Experiment 1 and 2 suggest that the different findings regarding the importance of layout complexity in the Schwartz and Tullis studies may be due to differences in task characteristics.

When evaluating the practical utility of these findings, it is important to remember that task characteristics (type and number of targets) only interacted with layout complexity when starting positions were totally unpredictable. When starting positions were arranged in a predictable pattern, subjects' performance on the four

tasks did not differ depending on the type of screen. While an ordered pattern of starting positions is plausible, it is unlikely that a screen designer would choose to present information using random starting positions. However, it is not difficult to imagine a situation for which it may be necessary to use a particular grouping strategy (e.g., a hierarchical structure based on semantic properties of the information where indenting indicates group membership) which does not lend itself to a left-justified format. The findings from Experiment 1 and 2 suggest that it would not be necessary to left-justify the information. One limit of this conclusion is the use of a totally unpredictable pattern when the screen will be used to scan multiple data values.

Another practical consideration is that in the present study subjects stressed accuracy and were motivated to perform as quickly and efficiently as possible. This may not always fit the profile for all user populations of alphanumeric displays. It is also possible that performing different tasks with different types of screens required different degrees of effort. Different degrees of mental effort may affect performance in work situations in which the user's attention may be directed to more than one tasks at a time. The answers to these questions can only be obtained through further research.

General Conclusions

Manipulating layout complexity independently of the other Tullis dimensions demonstrated that for all but one task layout complexity is not a causal factor in determining search time for alphanumeric displays. Even severe violations of the guidelines that recommend a left-justified format did not increase search time for three tasks (find label, scan data, and compare label). However, when subjects compared multiple data values, totally unpredictable starting position sfor data items increased user search time. Practice improved performance across all four tasks but did not interact with task characteristics or layout complexity.

In general the findings support the predictions of the Tullis model that layout complexity is not an important factor in the prediction of human performance but is an important factor in the prediction of human preference. It is not necessary to left-justify items on alphanumeric displays to be used to perform information extraction tasks. However if user acceptance of the display is an important consideration or if the display is to be used to scan multiple data values, it is recommended that data be arranged in some ordered, predictable pattern. This recommendation affords screen designers some latitude for presenting information which is not readily organized in a left-justified tabular format and gives some context for trade-off decisions which pit user performance against user preference.

The conclusion is not that the guidelines are wrong, but that without knowledge of how different display principles are affected by different task or user characteristics it is difficult to know how to apply the guidelines to specific situations. The most useful guidelines are applied design principles worked out for a particular context. The

emphasis must be on the identification of design principles and the specification of the contexts in which they are applicable. To the extent to which we understand what aspects of tasks and user characteristics are important to performance and preference and how a given user processes available information in performing those tasks, we are in a better position to develop and test methods for improving display design.

References

- Apple Computer Inc. (1985). Macintosh user interface guidelines. In Inside macintosh (Promotional Edition). Cupertino, CA: Apple Computer, Inc.
- Bahrack, H. P., Noble, M., & Fitts, P. M. (1954). Extra task performance as a measure of learning a primary task. Journal of Experimental Psychology, 48, 298-302.
- Bailey, R. W. (1982). Human performance engineering: A guide for system designers. Englewood Cliffs, NJ: Prentice-Hall.
- Barnes, M. J. (1987). Human information processing guidelines for decision aiding displays. Naval Weapon Center. China Lake, CA. Report No. AD-A124858.
- Benpasat, I., Dexter, A. S., & Todd, F. (1986). The influence of color and graphical information presentation in a managerial decision simulation. Human Computer Interaction, 2, 65-92.
- Berliner, C., Angell, D., & Shearer, J. N. (1964). Behaviors, measures, and instruments for performance evaluation in simulated environment. Proceedings: Symposium and Workshop on the Quantification of Human Performance. Albuquerque, N.M.
- Bonsiepe, G. (1968). A method of quantifying order in typographic design. Journal of Typographic Research, 2, 203-220.
- Burn, M. J., Warren, D. L., & Rudisill, M. (1985). Expert and nonexpert search task performance with current and reformatted Space Shuttle displays. (Report No. NASA-JSC-20967) Houston, TX: Lockheed Engineering and Management Services Company, Inc. (LEMSCO-22353).
- Burroughs Corporation (1986). Interpro (TM) user interface standards. Version E (March 20, 1986). (Unpublished technical report available from Manager, Human Factors Section, System Products Group, Burroughs Corp., 19 Morgan, Divine, CA 90718).
- Card, S. K. (1982). User perceptual mechanisms in the search of computer command menus. Proceedings: Human Factors in Computer Systems. Gaithersburg, MD. New York: Association for Computer Machinery.
- Cakir, A., Hart, D. J., & Stewart, T. F. M. (1980). Visual display terminals: A manual covering ergonomics, workplace design, health and safety, and task organization. Chichester, England: John Wiley & Sons.
- Callan, J. R., Curran, L. E., & Lane, J. L. (1977). Visual search times for Navy

- tactical information displays (Report No. NPRDC-TR-77-32). San Diego, CA: Navy Personnel Research and Development Center.
- Christ, R. M., & Corso, G. M. (1983). The effects of extended practice on the evaluation of visual display codes. Human Factors, 25, 71-84.
- Companion, M. A., & Corso, G. M. (1982). Task taxonomies: A general review and evaluation. International Journal of Man-Machine Studies, 17, 459-472.
- Cropper, A. G., & Evans, S. J. W. (1968). Ergonomics and computer display design. The Computer Bulletin, 12, 94-98.
- Crowder, R. G. (1976). Principle of Learning and Memory. N.J.: John Wiley.
- Danckak, M. M. (1976). CRT displays for power plants. Instrumentation Technology, 23, 29-36.
- Dodson, D. W., & Shields, N. L., Jr. (1978). Development of user guidelines for ECAS display design. (Vol. 1). (Report No. NASA-CR-150877). Huntsville, AL: Essex Corp.
- Drury, C., & Clement, M. (1978). The effect of area, density, and number of background characters on visual search. Human Factors, 20, 597-602.
- Engel, S. E., & Granda, R. E. (1975). Guidelines for man/display interfaces (Report No. TR 00.2720). Poughkeepsie, NY: IBM.
- Farina, A. J., & Wheaton, G. R. (1973). Development of a taxonomy of human performance: The task characteristics approach to performance prediction. (American Institute for Research Tech Rep.) JSAS Catalog of Selected Documents in Psychology, 3, 26-27 (Ms. No. 323).
- Fleishman, E. A., & Quaintance, M. K. (1984). Taxonomies of human performance: The description of human tasks. Orlando, FL: Academic Press.
- Gagné, R. M. (Ed.). (1962). Psychological principles in system development. New York: Holt.
- Galitz, W. C. (1980). Human factors in office automation. Atlanta, GA: Life Office Management Assoc.
- Galitz, W. C. (1985). Handbook of screen format design (Second Edition). Wellesley Hills, MA: QED Information Sciences.
- Gould, J. D., & Lewis, C. (1985). Designing for usability: Key principles and what designers think. Communications of ACM, 28, 300-311.

- Hammond, K. R. (1980). The integration of research in judgment and decision theory (Report No. CRJP 226). Boulder, CO: University of Colorado, Institute of Behavioral Science.
- Hitt, W. D. (1961). An evaluation of five different abstract coding methods. Human Factors, 3, 120-130.
- Hirst, W., Spelke, E. S., Reaves, C. C., Cahavack, G., & Neisser, U. (1980). Dividing attention without alternation or automaticity. Journal of Experimental Psychology: General, 109, 98-117.
- Howell, W.C., & Dipboye, R.L. (1986). Essentials of Industrial and Organizational Psychology. Chicago, Ill.: Dorsey Press.
- Howell, W. C., Lane, D. M., Harvey, R. J., & Holden, K. L. (1986). Human cognition and information display in C3I System Tasks (Report No. MDA 903-85-C-0347). Houston, TX: Rice University.
- Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman and D. R. Davies (Eds.), Varieties of attention. Orlando, FL: Academic Press.
- Keele, S.W. (1973). Attention and human performance. Pacific Palisades, CA: Goodyear.
- Keister, R. S. & Gallaway, G. R. (1983). Making software user friendly: An assessment of data entry performance: Proceedings of the Human Factors Society 27th Annual Meeting. 1031-1039, Santa Monica, CA: Human Factors Society.
- Kolers, P. A., Duchnicky, R. L., & Ferguson, D. C. (1981). Eye movement measurement of readability of CRT displays. Human Factors, 23, 517-527.
- Lenorovitz, D. R., Phillips, M. D., Audrey, R. S., & Kioester, G. V. (1984). A taxonomic approach to characterizing human-computer interfaces. In G. Salvendy (Ed.), Human-Computer Interaction: Proceedings of the First USA-Japan Conference on Human Computer Interaction. Amsterdam: Elsevier Science Publishers B.V.
- Lockheed Missiles and Space Company (1983). Human factors engineering standards for information processing systems. Sunnydale, CA: Lockheed Missiles and Space Co.
- Martin, J. (1973). Design of man-computer dialogues. Englewood Cliff, NJ: Prentice-Hall, Inc.
- Meister, D. (1985). Behavioral analysis and measurement methods. New York: Wiley.

- Melton, A. W., & Briggs, G. E. (1960). Engineering psychology. Annual Review of Psychology, 11, 7-98.
- Miller, R. B. (1962). Task description and analysis. In R. M. Gagné (Ed.), Psychological principles in system development. New York: Holt.
- MIL-STD-1472C. (May, 1981). Military standard: Human engineering design criteria for military systems, equipment, and facilities. Washington, DC: Department of Defense.
- Monk, A. (1984). Fundamentals of human computer interaction. London: Academic Press.
- NASA (1980). Spacelab display design and command usage guidelines. (Report MSFC-PROC-711A). Huntsville, AL: George C. Marshall Space Flight Center.
- Neisser, U., Novick, R., & Lazar, R. (1964). Searching for word targets. Perceptual and Motor Skills, 19, 427-432.
- Peterson, R. J., Banks, W. W., & Gertman, D. I. (1982). Performance-based evaluation of graphic displays for nuclear power plant control rooms. In Human Factors in Computer Systems, 182-189.
- Ringel, S., & Hammer, C. (1964). Information assimilation from alphanumeric displays: Amount and density of information presented (Tech Report TRN141). Washington, DC: US Army Personnel Research Office. (NTIS No. AD 601973).
- Salvendy, G. (Ed.) (1984). Human Computer Interaction: Proceedings of the First USA-Japan Conference on Human Computer Interaction. Amsterdam: Elsevier Science Publishers.
- Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing. Psychological Review, 84, 1-66.
- Schwartz, D. (1986). Formatting effects on the use of computer-generated alphanumeric displays: The moderating effects of task characteristics. Unpublished dissertation. Rice University, Houston, Texas.
- Schwartz, D. R., & Howell, W. C. (1985). Optional stopping performance under graphic and numeric CRT formatting. Human Factors, 27, 433-434.
- Shneiderman, B. (1987). Designing the user interface: Strategies for effective human computer interaction. Reading, MA: Addison-Wesley.
- Sidorsky, R. C., Parrish, R. N., Gates, J. L., & Munger, S. J. (1984). Design guidelines for user transactions with battlefield automated systems: Prototype

- for a handbook. (ARI Research Product 84-08). Alexandria, VA: U.S. Army Research Institute. (NTIS No. AD A513 231).
- Siegel, A. I., & Fischl, M. A. (1971). Dimensions of visual information displays. Journal of Applied Psychology, 55, 470-476.
- Smith, S. L., & Ancella, A. F. (1983). Design guidelines for the user interface to computer-based information systems. (F19628-82-C-0001). Bedford, MA: The Mitre Corp.
- Smith, S. L., & Mosier, J. N. (1984). Design guidelines for user-system interface software (ESD-TR-84-190). Bedford, MA: The Mitre Corp.
- Smith, S. L., & Mosier, J. N. (1986). Guidelines for user-interface software (ESD-TR-86-278). Bedford, MA: The Mitre Corp. (NTIS. AD-A177 198).
- Stewart, T. F. M. (1976). Displays and the software interface. Applied Ergonomics, 7, 137-146.
- Streveler, D. J., & Wasserman, A. I. (1984). Quantitative measures of the spatial properties of screen designs. Proceedings of Interact '84 Conference on Human-Computer Interaction. London, England.
- Teichner, W. H., & Krebs, M. J. (1974). Laws of visual choice reaction time. Psychological review, 81, 15-28.
- Teichner, W. H., & Mocharnuk, J. B. (1979). Visual search for complex targets. Human Factors, 21, 259-275.
- Tullis, T. S. (1981). An evaluation of alphanumeric, graphic, and color information displays. Human Factors, 23, 211-230.
- Tullis, T. S. (1983). The formatting of alphanumeric displays: A review and analysis. Human Factors, 25, 657-682.
- Tullis, T. S. (1984). Predicting the usability of alphanumeric displays. Unpublished doctoral dissertation, Rice University, Houston, TX.
- Tullis, T. (1986). Screen design. In M. Helander (Ed.), Handbook of human computer interaction. In press.
- Wasserman, A. I., & Shewmake, D. T. (1985). The role of prototypes in the user software engineering (USE) methodology. In R. Hartson (Ed.), Advances in human-computer interaction 1. Norwood, NJ: Ablex Publishing Corporation.
- Wickens, C. D. (1984). Engineering psychology and human performance. Columbus, OH: Merrill Publishing.

- Winer, B.J. (1971). Statistical Principles in Experimental Design. New York: McGraw Hill.
- Willigers, B. H. & Willges, R. D. (1981). User considerations in computer-based information systems (Tech. Rep. CSiE-81-2). Blacksburg, VA: Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University (NTIS No. AD A106 194).
- Wolf, C. E. (1986). BNA "HN" command display: Results of user evaluation. Unpublished technical report. (Avaliable for Cynthia Wolf Connally, Burroughs Corporation, 19 Morgan, Irvine, CA, 92718).
- Yorchak, J. P., Allison, J. E., & Dodd, U. S. (1984). A new tilt on computer generated space situation displays. Proceedings of the Human Factors Society 28th Annual Meeting, 894-898. Sania Monica, CA: Human Factors Society.

Appendix A
Experimental Screens

North Carolina

Motel/Hotel	Telephone	Single	Double
Asheboro Area Code: 919			
Holiday Inn	629-2101	\$32	\$39
Sir Robert	625-5171	\$18	\$24
Asheville Area Code: 704			
American Court	253-4427	\$20	\$24
Best Western	298-5562	\$32	\$32
Buena Vista	274-1646	\$22	\$23
Days Inn Airport	684-2281	\$27	\$31
Days Inn East	298-5140	\$27	\$31
Econo Lodge	298-5519	\$23	\$29
Forest Manor Lodge	274-3531	\$40	\$55
Holiday Inn I-40 East	298-5611	\$40	\$50
Holiday Inn West	667-4501	\$40	\$48
Howard Johnsons South	274-2300	\$50	\$60
Nakon	667-9103	\$20	\$30
Ramada Inn	254-7451	\$38	\$42
Sheraton Motor Inn	253-1851	\$36	\$47

Figure 1. Screen A--predictable

North Carolina

Motel/Hotel	Telephone	Single	Double
Asheboro Area Code: 919			
Holiday Inn	629-2101	\$32	\$39
Sir Robert	625-5171	\$18	\$24
Asheville Area Code: 704			
American Court	253-4427	\$20	\$24
Best Western	298-5562	\$32	\$32
Buena Vista	274-1646	\$22	\$23
Days Inn Airport	684-2281	\$27	\$31
Days Inn East	298-5140	\$27	\$31
Econo Lodge	298-5519	\$23	\$29
Forest Manor Lodge	274-3531	\$40	\$55
Holiday Inn I-40 East	298-5611	\$40	\$50
Holiday Inn West	667-4501	\$40	\$48
Howard Johnsons South	274-2300	\$50	\$60
Nakon	667-9103	\$20	\$30
Ramada Inn	254-7451	\$38	\$42
Sheraton Motor Inn	253-1851	\$36	\$47

Figure 1. Screen A--left-justified

North Carolina

Motel/Hotel	Telephone	Single	Double
Asheboro	Area Code: 919		
Holiday Inn	629-2101	\$32	\$39
Sir Robert	625-5171	\$18	\$24
Asheville	Area Code: 704		
Court	253-4427	\$20	\$24
Best Western	298-5562	\$32	\$32
Buena Vista	274-1646	\$22	\$23
Days Inn Airport	684-2281	\$27	\$31
Days Inn East	298-5140	\$27	\$31
Econo Lodge	298-5519	\$23	\$29
Forest Manor Lodge	274-3531	\$40	\$55
Holiday Inn I-40 East	298-5611	\$40	\$50
Holiday Inn West	667-4501	\$40	\$48
Howard Johnsons South	274-2300	\$50	\$60
Nakon	667-9103	\$20	\$30
Ramada Inn	254-7451	\$38	\$42
Sheraton Motor Inn	253-1851	\$36	\$47

American

Figure 1. Screen A--random

Maryland

City	Motel/Hotel	Area Code	Phone	Single	Double
Laurel	Holiday Inn	301	498-0900	\$40	\$48
	Quality Inn Colony 7		725-5100	\$34	\$40
Oakland	Alpine Village	301	387-5534	\$36	\$39
	Lake Breez		387-5564	\$36	\$40
	Lake Side Motor Court		387-5566	\$38	\$42
	Oak-Land		334-2171	\$21	\$26
	Town		334-3955	\$25	\$30
	Will o' the Wisp		387-5503	\$47	\$59
Ocean City	Carousel Hotel	301	524-3464	\$79	\$95
	Quality Inn Boardwalk		289-4401	\$74	\$90
	Quality Inn Stardust		289-6444	\$75	\$75
	Sahara		289-8101	\$52	\$62
	Sheraton Fontainbleau		524-3535	\$35	\$45
	Surf and Sands		289-7161	\$28	\$60

Figure 1. Screen B--predictable

Maryland

City	Motel/Hotel	Area Code	Phone	Rates	
Laurel	Holiday Inn	301	498-0900	Single \$40	Double \$48
	Quality Inn Colony 7		725-5100	\$34	\$40
Oakland	Alpine Village	301	387-5534	\$36	\$39
	Lake Breez		387-5564	\$36	\$40
	Lake Side Motor Court		387-5566	\$38	\$42
	Oak-Land		334-2171	\$21	\$26
	Town		334-3955	\$25	\$30
	Will o' the Wisp		387-5503	\$47	\$59
Ocean City	Carousel Hotel	301	524-3464	\$79	\$95
	Quality Inn Boardwalk		289-4401	\$74	\$90
	Quality Inn Stardust		289-6444	\$75	\$75
	Sahara		289-8101	\$52	\$62
	Sheraton Fontainbleau		524-3535	\$35	\$45
	Surf and Sands		289-7161	\$28	\$60

Figure 1. Screen B---left-justified

Maryland

City	Motel/Hotel	Area Code	Rates	
			Phone	Single Double
Laurel	Holiday Inn	301	498-0900	\$40 \$48
	Quality Inn Colony 7		725-5100	\$34 \$40
Oakland	Alpine Village	301	387-5534	\$36 \$39
	Lake Breez		387-5564	\$36 \$40
	Lake Side Motor Court		387-5566	\$38 \$42
	Oak-Land		334-2171	\$21 \$26
	Town		334-3955	\$25 \$30
	Will o' the Wisp		387-5503	\$47 \$59
Ocean City	Carousel Hotel	301	524-3464	\$79 \$95
	Quality Inn Boardwalk		289-4401	\$74 \$90
	Quality Inn Stardust		289-6444	\$75 \$75
	Sahara		289-8101	\$52 \$62
	Sheraton Fountainbleau		524-3535	\$35 \$45
	Surf and Sands		289-7161	\$28 \$60

Figure 1. Screen B---random

Pennsylvania

City	Motel/Hotel	Area	Rates	
			Code Phone	Single Double
Lancaster	Best Western	717	768-3636	\$31 \$40
	Cherry Lane Motor Inn		687-7646	\$42 \$44
	Historic Strasburg Inn		687-7691	\$53 \$69
	Holiday Inn East		299-2551	\$35 \$62
	Holiday Inn North		393-0771	\$28 \$34
	Howard Johnsons		397-7781	\$46 \$50
	Quality Inn		898-2431	\$29 \$44
	Quiet Haven		397-6231	\$32 \$34
	Red Caboose Lodge		687-6646	\$27 \$33
	Sheraton Lancaster		393-0731	\$42 \$48
Lebanon	Willow Valley Farm		464-2711	\$50 \$53
	Lantern Lodge	717	866-6536	\$30 \$42
	Lebanon Treadway Inn		273-6771	\$34 \$40
	Tulpehocken Manor		866-4926	\$30 \$35
Lewisburg	Colonial Crest	717	523-1106	\$22 \$26

Figure 1. Screen C--predictable

Pennsylvania

City	Motel/Hotel	Area		Rates	
		Code	Phone	Single	Double
Lancaster	Best Western	717	768-3636	\$31	\$40
	Cherry Lane Motor Inn		687-7646	\$42	\$44
	Historic Strasburg Inn		687-7691	\$53	\$69
	Holiday Inn East		299-2551	\$35	\$62
	Holiday Inn North		393-0771	\$28	\$34
	Howard Johnsons		397-7781	\$46	\$50
	Quality Inn		898-2431	\$29	\$44
	Quiet Haven		397-6231	\$32	\$34
	Red Caboose Lodge		687-6646	\$27	\$33
	Sheraton Lancaster		393-0731	\$42	\$48
	Willow Valley Farm		464-2711	\$50	\$53
Lebanon	Lantern Lodge	717	866-6536	\$30	\$42
	Lebanon Treadway Inn		273-6771	\$34	\$40
	Tulpehocken Manor		866-4926	\$30	\$35
Lewisburg	Colonial Crest	717	523-1106	\$22	\$26

Figure 1. Screen C--left-justified

Pennsylvania

City	Motel/Hotel	Area		Rates	
		Code	Phone	Single	Double
Lancaster	Best Western	717	768-3636	\$31	\$40
	Cherry Lane Motor Inn		687-7646	\$42	\$44
	Historic Strasburg Inn		687-7691	\$53	\$69
	Holiday Inn East		299-2551	\$35	\$62
	Holiday Inn North		393-0771	\$28	\$34
	Howard Johnsons		397-7781	\$46	\$50
	Quality Inn		898-2431	\$29	\$44
	Quiet Haven		397-6231	\$32	\$34
	Red Caboose Lodge		687-6646	\$27	\$33
	Sheraton Lancaster		393-0731	\$42	\$48
Lebanon	Willow Valley Farm		464-2711	\$50	\$53
	Lantern Lodge	717	866-6536	\$30	\$42
	Lebanon Treadway Inn		273-6771	\$34	\$40
Lewisburg	Tulpehocken Manor		866-4926	\$30	\$35
	Colonial Crest	717	523-1106	\$22	\$26

Figure 1. Screen C---random

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Asheboro	Holiday Inn	919	629-2101	\$32	\$39
	Sir Robert		625-5171	\$18	\$24
Asheville	American Court	704	253-4427	\$20	\$24
	Best Western		298-5562	\$32	\$32
	Buena Vista		274-1646	\$22	\$23
	Days Inn Airport		684-2281	\$27	\$31
	Days Inn East		298-5140	\$27	\$31
	Econo Lodge		298-5519	\$23	\$29
	Forest Manor Lodge		274-3531	\$40	\$55
	Holiday Inn I-40 East		298-5611	\$40	\$50
	Holiday Inn West		667-4501	\$40	\$48
	Howard Johnsons South		274-2300	\$50	\$60
	Nakon		667-9103	\$20	\$30
	Ramada Inn		254-7451	\$38	\$42
	Sheraton Motor Inn		253-1851	\$36	

Figure 1. Screen D--predictable

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Asheboro	Holiday Inn	919	629-2101	\$32	\$39
	Sir Robert		625-5171	\$18	\$24
Asheville	American Court	704	253-4427	\$20	\$24
	Best Western		298-5562	\$32	\$32
	Buena Vista		274-1646	\$22	\$23
	Days Inn Airport		684-2281	\$27	\$31
	Days Inn East		298-5140	\$27	\$31
	Econo Lodge		298-5519	\$23	\$29
	Forest Manor Lodge		274-3531	\$40	\$55
	Holiday Inn I-40 East		298-5611	\$40	\$50
	Holiday Inn West		667-4501	\$40	\$48
	Howard Johnsons South		274-2300	\$50	\$60
	Nakon		667-9103	\$20	\$30
	Ramada Inn		254-7451	\$38	\$42
	Sheraton Motor Inn		253-1851	\$36	\$47

Figure 1. Screen D--left-justified

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Asheboro	Holiday Inn	919	629-2101	\$32	\$39
	Sir Robert		625-5171	\$18	\$24
Asheville	American Court	704	253-4427	\$20	\$24
	Best Western		298-5562	\$32	\$32
	Buena Vista		274-1646	\$22	\$23
	Days Inn Airport		684-2281	\$27	\$31
	Days Inn East		298-5140	\$27	\$31
	Econo Lodge		298-5519	\$23	\$29
	Forest Manor Lodge		274-3531	\$40	\$55
	Holiday Inn I-40 East		298-5611	\$40	\$50
	Holiday Inn West		667-4501	\$40	\$48
	Howard Johnsons South		274-2300	\$50	\$60
	Nakon		667-9103	\$20	\$30
	Ramada Inn		254-7451	\$38	\$42
	Sheraton Motor Inn		253-1851	\$36	\$47

Figure 1. Screen D--random

West Virginia

City	Motel/Hotel	Area		Rates	
		Code	Phone	Single	Double
Beckley	Best Western Lodge	304	252-0671	\$26	\$27
	Holiday Inn		255-1511	\$32	\$37
	Honey in the Rock		252-7391	\$24	\$29
	Laurel Lodge		255-2161	\$25	\$28
	Pagoda		253-7373	\$22	\$24
Bluefield	Ramada Inn		252-8661	\$32	\$36
	Twin Falls Lodge		294-4000	\$26	\$30
	Brier Red Carpet Inn	304	325-9111	\$25	\$28
	Econo-Travel		327-8171	\$21	\$25
	Holiday Inn		325-5421	\$40	\$44
Buckhannon	Bicentennial	304	472-5000	\$22	\$28
	Centennial		472-4100	\$23	\$34
Charleston	El Rancho	304	727-2201	\$21	\$29
	Executive Inn		345-8820	\$34	\$38
	Holiday Inn Broad St		343-4661	\$40	\$44
	Holiday Inn Civic Ctr		345-0600	\$38	\$42
	Holiday Inn Downtown		344-4092	\$44	\$48
	Holiday Inn Elk River		323-4521	\$36	\$42

Figure 1. Screen E---predictable

West Virginia

City	Motel/Hotel	Area	Rates	
			Single	Double
Beckley	Best Western Lodge	304	\$26	\$27
	Holiday Inn		\$32	\$37
	Honey in the Rock		\$24	\$29
	Laurel Lodge		\$25	\$28
	Pagoda		\$22	\$24
	Ramada Inn		\$32	\$36
Bluefield	Twin Falls Lodge		\$26	\$30
	Brier Red Carpet Inn	304	\$25	\$28
	Econo-Travel		\$21	\$25
	Holiday Inn		\$40	\$44
Buckhannon	Bicentennial	304	\$22	\$28
	Centennial		\$23	\$34
	El Rancho	304	\$21	\$29
Charleston	Executive Inn		\$34	\$38
	Holiday Inn Broad St		\$40	\$44
	Holiday Inn Civic Ctr		\$38	\$42
	Holiday Inn Downtown		\$44	\$48
	Holiday Inn Elk River		\$36	\$42

Figure 1. Screen E--left-justified

West Virginia

City	Motel/Hotel	Area Code	Rates	
			Phone	Single Double
Beckley	Best Western Lodge	304	252-0671	\$26 \$27
	Holiday Inn		255-1511	\$32 \$37
	Honey in the Rock		252-7391	\$24 \$29
	Laurel Lodge		255-2161	\$25 \$28
	Pagoda		253-7373	\$22 \$24
	Ramada Inn		252-8661	\$32 \$36
	Ramada Inn N		252-6131	\$42 \$54
	Twin Falls Lodge		294-4000	\$26 \$30
Bluefield	Brier Red Carpet Inn	304	325-9111	\$25 \$28
	Econo-Travel		327-8171	\$21 \$25
	Holiday Inn		325-5421	\$40 \$44
Buckhannon	Bicentennial	304	472-5000	\$22 \$28
	Centennial		472-4100	\$23 \$34
Charleston	El Rancho	304	727-2201	\$21 \$29
	Executive Inn		345-8820	\$34 \$38
	Holiday Inn Broad St		343-4661	\$40 \$44
	Holiday Inn Civic Ctr		345-0600	\$38 \$42
	Holiday Inn Downtown		344-4092	\$44 \$48
	Holiday Inn Elk River		323-4521	\$36 \$42

Figure 1. Screen E--random

South Carolina

Charleston

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	747-0961	\$25	\$30
Days Inn	881-1800	\$18	\$24
Holiday Inn N	744-1621	\$36	\$46
Holiday Inn SW	556-7100	\$33	\$47
Howard Johnsons	524-4140	\$31	\$36
Ramada Inn	774-8281	\$33	\$40
Sheraton Inn	744-2401	\$34	\$42

Columbia

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	796-9400	\$29	\$34
Carolina Inn	799-8200	\$42	\$48

Figure 1. Screen F--predictable

South Carolina

Charleston

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	747-0961	\$26	\$30
Days Inn	881-1800	\$18	\$24
Holiday Inn N	744-1621	\$36	\$46
Holiday Inn SW	556-7100	\$33	\$47
Howard Johnsons	524-4140	\$31	\$36
Ramada Inn	774-8281	\$33	\$40
Sheraton Inn	744-2401	\$34	\$42

Columbia

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	796-9400	\$29	\$34
Carolina Inn	799-8200	\$42	\$48

Figure 1. Screen F--left-justified

South Carolina

Charleston

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	747-0961	\$26	\$30
Days Inn	881-1800	\$18	\$24
Holiday Inn N	744-1621	\$36	\$46
Holiday Inn SW	556-7100	\$33	\$47
Howard Johnsons	524-4140	\$31	\$36
Ramada Inn	774-8281	\$33	\$40
Sheraton Inn	744-2401	\$34	\$42

Columbia

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	796-9400	\$29	\$34
Carolina Inn	799-8200	\$42	\$48

Figure 1. Screen F--left-justified

South Carolina		
Charleston Best Western	Phone: (803) 747-0961	Sngl: \$26 Dbl: \$30
Charleston Days Inn	Phone: (803) 881-1800	Sngl: \$18 Dbl: \$24
Charleston Holiday Inn N	Phone: (803) 744-1621	Sngl: \$36 Dbl:
\$46		
Charleston Holiday Inn SW	Phone: (803) 556-7100	Sngl: \$33 Dbl:
\$47		
Charleston Howard Johnsons	Phone: (803) 524-4140	Sngl: \$31 Dbl:
\$36		
Charleston Ramada Inn	Phone: (803) 774-8281	Sngl: \$33 Dbl: \$40
Charleston Sheraton Inn	Phone: (803) 744-2401	Sngl: \$34 Dbl: \$42

Figure 1. Screen G--predictable

South Carolina						
Charleston	Best Western	Phone: (803)	747-0961	Sngl: \$26	Dbl: \$30	
Charleston	Days Inn	Phone: (803)	881-1800	Sngl: \$18	Dbl: \$24	
Charleston	Holiday Inn N	Phone: (803)	744-1621	Sngl: \$36	Dbl: \$46	
Charleston	Holiday Inn SW	Phone: (803)	556-7100	Sngl: \$33	Dbl: \$47	
Charleston	Howard Johnsons	Phone: (803)	524-4140	Sngl: \$31	Dbl: \$36	
Charleston	Ramada Inn	Phone: (803)	774-8281	Sngl: \$33	Dbl: \$40	
Charleston	Sheraton Inn	Phone: (803)	744-2401	Sngl: \$34	Dbl: \$42	

Figure 1. Screen G--left-justified

South Carolina

Charleston Best Western	Phone: (803) 747-0961	Sngl: \$26	Db1: \$30
Charleston Days Inn	Phone: (803) 881-1800	Sngl: \$18	Db1: \$24
Charleston Holiday Inn SW	Phone: (803) 556-7100	Sngl: \$33	Db1: \$47
Charleston Holiday Inn Nw	Phone: (803) 667-6131	Sngl: \$23	Db1: \$42
Charleston Howard Johnsons	Phone: (803) 524-4140	Sngl: \$31	Db1: \$36
Charleston Ramada Inn	Phone: (803) 774-8281	Sngl: \$33	Db1: \$40
Charleston Sheraton Inn	Phone: (803) 744-2401	Sngl: \$34	Db1: \$42

Figure 1. Screen G--random

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Concord	Holiday Inn	704	786-5181	\$29	\$34
	Quality Inn Colonial		782-2146	\$18	\$28
Dunn	Cricket Inn	919	892-8711	\$16	\$20
	Days Inn		894-2031	\$18	\$22
	Holiday Inn		892-2162	\$24	\$31
	Howard Johnsons		892-6181	\$21	\$26

Figure 1. Screen H--predictable

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Concord	Holiday Inn	704	786-5181	\$29	\$34
	Quality Inn Colonial		782-2146	\$18	\$28
Dunn	Cricket Inn	919	892-8711	\$16	\$20
	Days Inn		894-2031	\$18	\$22
	Holiday Inn		892-2162	\$24	\$31
	Howard Johnsons		892-6181	\$21	\$26

Figure 1. Screen H---left-justified

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Concord	Holiday Inn	704	786-5181	\$29	\$34
	Quality Inn Colonial		782-2146	\$18	\$28
Dunn	Cricket Inn	919	892-8711	\$16	\$20
	Days Inn		894-2031	\$18	\$22
	Holiday Inn		892-2162	\$24	\$31
	Howard Johnsons		892-6181	\$21	\$26

Figure 1. Screen H--random

Appendix B
Analysis of Variance Tables

Table A

Analysis of Variance Table for Median Reaction Time in Experiment 1

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>	<u>p</u>	<u>Epsilon Correction</u>
Subjects	7	292.000	41.714			
Practice Session Error	3 21	171.520 313.042	57.173 14.907	3.835	.0246	.35
Type of Task Error	3 21	646.072 523.596	215.357 24.933	8.637	.0006	.38
Practice by Task Error	9 63	147.076 1033.480	16.342 16.404	.996	.4525	.12
Type of Screen Error	1 7	.328 10.822	.328 1.546	.212	.6589	1.00
Session by Screen Error	3 21	3.417 31.091	1.139 1.481	.769	.5240	.71
Task by Screen Error	3 21	.521 23.323	.174 1.111	.156	.9245	.45
Session by Task by Screen Error	9 63	8.907 67.736	.990 1.075	.920	.5136	.24

Table B

Analysis of Variance Table for Median Reaction Time in Experiment 2

<u>Source of Variation</u>	<u>df</u>	<u>Sum of Squares</u>	<u>Mean Square</u>	<u>F</u>	<u>p</u>	<u>Epsilon Correction</u>
Subjects	15	260.678	17.379			
Type of Screen Error	2 30	9.058 67.118	4.529 2.237	2.024	.1497	.97
Type of Task Error	3 45	578.687 257.909	192.896 5.731	33.656	.0000	.40
Screen by Task Error	6 90	19.670 42.598	3.278 .473	6.927	.0000	.36
Practice Session Error	2 30	9.468 16.071	4.734 .536	8.837	.0010	.75
Screen by Session Error	4 60	1.241 20.846	.310 .347	.893	.4737	.68
Task by Session Error	6 90	3.304 23.524	.551 .261	2.107	.0601	.55
Screen by Task by Session Error	12 180	.593 39.216	.049 .218	.227	.9969	.28

Appendix C
Subjective Rating Forms

North Carolina

Motel/Hotel	Telephone	Single	Double
Asheboro Area Code: 919			
Holiday Inn	629-2101	\$32	\$39
Sir Robert	625-5171	\$18	\$24
Asheville Area Code: 704			
American Court	253-4427	\$20	\$24
Best Western	298-5562	\$32	\$32
Buena Vista	274-1646	\$22	\$23
Days Inn Airport	684-2281	\$27	\$31
Days Inn East	298-5140	\$27	\$31
Econo Lodge	298-5519	\$23	\$29
Forest Manor Lodge	274-3531	\$40	\$55
Holiday Inn I-40 East	298-5611	\$40	\$50
Holiday Inn West	667-4501	\$40	\$48
Howard Johnsons South	274-2300	\$50	\$60
Nakon	667-9103	\$20	\$30
Ramada Inn	254-7451	\$38	\$42
Sheraton Motor Inn	253-1851	\$36	\$47

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

North Carolina

Motel/Hotel	Telephone	Single	Double
Asheboro Area Code: 919			
Holiday Inn	629-2101	\$32	\$39
Sir Robert	625-5171	\$18	\$24
Asheville Area Code: 704			
American Court	253-4427	\$20	\$24
Best Western	298-5562	\$32	\$32
Buena Vista	274-1646	\$22	\$23
Days Inn Airport	684-2281	\$27	\$31
Days Inn East	298-5140	\$27	\$31
Econo Lodge	298-5519	\$23	\$29
Forest Manor Lodge	274-3531	\$40	\$55
Holiday Inn I-40 East	298-5611	\$40	\$50
Holiday Inn West	667-4501	\$40	\$48
Howard Johnsons South	274-2300	\$50	\$60
Nakon	667-9103	\$20	\$30
Ramada Inn	254-7451	\$38	\$42
Sheraton Motor Inn	253-1851	\$36	\$47

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

North Carolina

Motel/Hotel	Telephone	Single	Double
Asheboro Area Code: 919			
Holiday Inn	629-2101	\$32	\$39
Sir Robert	625-5171	\$18	\$24
Asheville Area Code: 704			
American Court	253-4427	\$20	\$24
Best Western	298-5562	\$32	\$32
Buena Vista	274-1646	\$22	\$23
Days Inn Airport	684-2281	\$27	\$31
Days Inn East	298-5140	\$27	\$31
Econo Lodge	298-5519	\$23	\$29
Forest Manor Lodge	274-3531	\$40	\$55
Holiday Inn I-40 East	298-5611	\$40	\$50
Holiday Inn West	667-4501	\$40	\$48
Howard Johnsons South	274-2300	\$50	\$60
Nakon	667-9103	\$20	\$30
Ramada Inn	254-7451	\$38	\$42
Sheraton Motor Inn	253-1851	\$36	\$47

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

Maryland

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Laurel	Holiday Inn	301	498-0900	\$40	\$48
	Quality Inn Colony 7		725-5100	\$34	\$40
Oakland	Alpine Village	301	387-5534	\$36	\$39
	Lake Breez		387-5564	\$36	\$40
	Lake Side Motor Court		387-5566	\$38	\$42
	Oak-Land		334-2171	\$21	\$26
	Town		334-3955	\$25	\$30
	Will o' the Wisp		387-5503	\$47	\$59
Ocean City	Carousel Hotel	301	524-3464	\$79	\$95
	Quality Inn Boardwalk		289-4401	\$74	\$90
	Quality Inn Stardust		289-6444	\$75	\$75
	Sahara		289-8101	\$52	\$62
	Sheraton Fontainbleau		524-3535	\$35	\$45
	Surf and Sands		289-7161	\$28	\$60

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislik			Neither Dislike or Like			Strongly Like

Maryland

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Laurel	Holiday Inn	301	498-0900	\$40	\$48
	Quality Inn Colony 7		725-5100	\$34	\$40
Oakland	Alpine Village	301	387-5534	\$36	\$39
	Lake Breez		387-5564	\$36	\$40
	Lake Side Motor Court		387-5566	\$38	\$42
	Oak-Land		334-2171	\$21	\$26
	Town		334-3955	\$25	\$30
	Will o' the Wisp		387-5503	\$47	\$59
Ocean City	Carousel Hotel	301	524-3464	\$79	\$95
	Quality Inn Boardwalk		289-4401	\$74	\$90
	Quality Inn Stardust		289-6444	\$75	\$75
	Sahara		289-8101	\$52	\$62
	Sheraton Fontainebleau		524-3535	\$35	\$45
	Surf and Sands		289-7161	\$28	\$60

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

Maryland

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Laurel	Holiday Inn	301	498-0900	\$40	\$48
	Quality Inn Colony 7		725-5100	\$34	\$40
Oakland	Alpine Village	301	387-5534	\$36	\$39
	Lake Breez		387-5564	\$36	\$40
	Lake Side Motor Court		387-5566	\$38	\$42
	Oak-Land		334-2171	\$21	\$26
	Town		334-3955	\$25	\$30
	Will o' the Wisp		387-5503	\$47	\$59
Ocean City	Carousel Hotel	301	524-3464	\$79	\$95
	Quality Inn Boardwalk		289-4401	\$74	\$90
	Quality Inn Stardust		289-6444	\$75	\$75
	Sahara		289-8101	\$52	\$62
	Sheraton Fontainbleau		524-3535	\$35	\$45
	Surf and Sands		289-7161	\$28	\$60

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

Pennsylvania

City	Motel/Hotel	Area		Rates	
		Code	Phone	Single	Double
Lancaster	Best Western	717	768-3636	\$31	\$40
	Cherry Lane Motor Inn		687-7646	\$42	\$44
	Historic Strasburg Inn		687-7691	\$53	\$69
	Holiday Inn East		299-2551	\$35	\$62
	Holiday Inn North		393-0771	\$28	\$34
	Howard Johnsons		397-7781	\$46	\$50
	Quality Inn		898-2431	\$29	\$44
	Quiet Haven		397-6231	\$32	\$34
	Red Caboose Lodge		687-6646	\$27	\$33
	Sheraton Lancaster		393-0731	\$42	\$48
	Willow Valley Farm		464-2711	\$50	\$53
Lebanon	Lantern Lodge	717	866-6536	\$30	\$42
	Lebanon Treadway Inn		273-6771	\$34	\$40
	Tulpehocken Manor		866-4926	\$30	\$35
Lewisburg	Colonial Crest	717	523-1106	\$22	\$26

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

Pennsylvania

City	Motel/Hotel	Area		Rates	
		Code	Phone	Single	Double
Lancaster	Best Western	717	768-3636	\$31	\$40
	Cherry Lane Motor Inn		687-7646	\$42	\$44
	Historic Strasburg Inn		687-7691	\$53	\$69
	Holiday Inn East		299-2551	\$35	\$62
	Holiday Inn North		393-0771	\$28	\$34
	Howard Johnsons		397-7781	\$46	\$50
	Quality Inn		898-2431	\$29	\$44
	Quiet Haven		397-6231	\$32	\$34
	Red Caboose Lodge		687-6646	\$27	\$33
	Sheraton Lancaster		393-0731	\$42	\$48
	Willow Valley Farm		464-2711	\$50	\$53
Lebanon	Lantern Lodge	717	866-6536	\$30	\$42
	Lebanon Treadway Inn		273-6771	\$34	\$40
	Tulpehocken Manor		866-4926	\$30	\$35
Lewisburg	Colonial Crest	717	523-1106	\$22	\$26

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

Pennsylvania

City	Motel/Hotel	Area		Rates	
		Code	Phone	Single	Double
Lancaster	Best Western	717	768-3636	\$31	\$40
	Cherry Lane Motor Inn		687-7646	\$42	\$44
	Historic Strasburg Inn		687-7691	\$53	\$69
	Holiday Inn East		299-2551	\$35	\$62
	Holiday Inn North		393-0771	\$28	\$34
	Howard Johnsons		397-7781	\$46	\$50
	Quality Inn		898-2431	\$29	\$44
	Quiet Haven		397-6231	\$32	\$34
	Red Caboose Lodge		687-6646	\$27	\$33
	Sheraton Lancaster		393-0731	\$42	\$48
	Willow Valley Farm		464-2711	\$50	\$53
Lebanon	Lantern Lodge	717	866-6536	\$30	\$42
	Lebanon Treadway Inn		273-6771	\$34	\$40
	Tulpehocken Manor		866-4926	\$30	\$35
Lewisburg	Colonial Crest	717	523-1106	\$22	\$26

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Asheboro	Holiday Inn	919	629-2101	\$32	\$39
	Sir Robert		625-5171	\$18	\$24
Asheville	American Court	704	253-4427	\$20	\$24
	Best Western		298-5562	\$32	\$32
	Buena Vista		274-1646	\$22	\$23
	Days Inn Airport		684-2281	\$27	\$31
	Days Inn East		298-5140	\$27	\$31
	Econo Lodge		298-5519	\$23	\$29
	Forest Manor Lodge		274-3531	\$40	\$55
	Holiday Inn I-40 East		298-5611	\$40	\$50
	Holiday Inn West		667-4501	\$40	\$48
	Howard Johnsons South		274-2300	\$50	\$60
	Nakon		667-9103	\$20	\$30
	Ramada Inn		254-7451	\$38	\$42
	Sheraton Motor Inn		253-1851	\$36	\$47

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Asheboro	Holiday Inn	919	629-2101	\$32	\$39
	Sir Robert		625-5171	\$18	\$24
Asheville	American Court	704	253-4427	\$20	\$24
	Best Western		298-5562	\$32	\$32
	Buena Vista		274-1646	\$22	\$23
	Days Inn Airport		684-2281	\$27	\$31
	Days Inn East		298-5140	\$27	\$31
	Econo Lodge		298-5519	\$23	\$29
	Forest Manor Lodge		274-3531	\$40	\$55
	Holiday Inn I-40 East		298-5611	\$40	\$50
	Holiday Inn West		667-4501	\$40	\$48
	Howard Johnsons South		274-2300	\$50	\$60
	Nakon		667-9103	\$20	\$30
	Ramada Inn		254-7451	\$38	\$42
	Sheraton Motor Inn		253-1851	\$36	\$47

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Asheboro	Holiday Inn	919	629-2101	\$32	\$39
	Sir Robert		625-5171	\$18	\$24
Asheville	American Court	704	253-4427	\$20	\$24
	Best Western		298-5562	\$32	\$32
	Buena Vista		274-1646	\$22	\$23
	Days Inn Airport		684-2281	\$27	\$31
	Days Inn East		298-5140	\$27	\$31
	Econo Lodge		298-5519	\$23	\$29
	Forest Manor Lodge		274-3531	\$40	\$55
	Holiday Inn I-40 East		298-5611	\$40	\$50
	Holiday Inn West		667-4501	\$40	\$48
	Howard Johnsons South		274-2300	\$50	\$60
	Nakon		667-9103	\$20	\$30
	Ramada Inn		254-7451	\$38	\$42
	Sheraton Motor Inn		253-1851	\$36	\$47

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

West Virginia

		Area	Rates		
City	Motel/Hotel	Code	Phone	Single	Double
Beckley	Best Western Lodge	304	252-0671	\$26	\$27
	Holiday Inn		255-1511	\$32	\$37
	Honey in the Rock		252-7391	\$24	\$29
	Laurel Lodge		255-2161	\$25	\$28
	Pagoda		253-7373	\$22	\$24
	Ramada Inn		252-8661	\$32	\$36
	Twin Falls Lodge		294-4000	\$26	\$30
Bluefield	Brier Red Carpet Inn	304	325-9111	\$25	\$28
	Econo-Travel		327-8171	\$21	\$25
	Holiday Inn		325-5421	\$40	\$44
Buckhannon	Bicentennial	304	472-5000	\$22	\$28
	Centennial		472-4100	\$23	\$34
Charleston	El Rancho	304	727-2201	\$21	\$29
	Executive Inn		345-8820	\$34	\$38
	Holiday Inn Broad St		343-4661	\$40	\$44
	Holiday Inn Civic Ctr		345-0600	\$38	\$42
	Holiday Inn Downtown		344-4092	\$44	\$48
	Holiday Inn Elk River		323-4521	\$36	\$42

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

West Virginia

		Area		Rates	
City	Motel/Hotel	Code	Phone	Single	Double
Beckley	Best Western Lodge	304	252-0671	\$26	\$27
	Holiday Inn		255-1511	\$32	\$37
	Honey in the Rock		252-7391	\$24	\$29
	Laurel Lodge		255-2161	\$25	\$28
	Pagoda		253-7373	\$22	\$24
	Ramada Inn		252-8661	\$32	\$36
Bluefield	Twin Falls Lodge		294-4000	\$26	\$30
	Brier Red Carpet Inn	304	325-9111	\$25	\$28
	Econo-Travel		327-8171	\$21	\$25
	Holiday Inn		325-5421	\$40	\$44
Buckhannon	Bicentennial	304	472-5000	\$22	\$28
	Centennial		472-4100	\$23	\$34
Charleston	El Rancho	304	727-2201	\$21	\$29
	Executive Inn		345-8820	\$34	\$38
	Holiday Inn Broad St		343-4661	\$40	\$44
	Holiday Inn Civic Ctr		345-0600	\$38	\$42
	Holiday Inn Downtown		344-4092	\$44	\$48
	Holiday Inn Elk River		323-4521	\$36	\$42

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

West Virginia

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Beckley	Best Western Lodge	304	252-0671	\$26	\$27
	Holiday Inn		255-1511	\$32	\$37
	Honey in the Rock		252-7391	\$24	\$29
	Laurel Lodge		255-2161	\$25	\$28
	Pagoda		253-7373	\$22	\$24
	Ramada Inn		252-8661	\$32	\$36
	Ramada Inn N		252-6131	\$42	\$54
	Twin Falls Lodge		294-4000	\$26	\$30
Bluefield	Brier Red Carpet Inn	304	325-9111	\$25	\$28
	Econo-Travel		327-8171	\$21	\$25
	Holiday Inn		325-5421	\$40	\$44
Buckhannon	Bicentennial	304	472-5000	\$22	\$28
	Centennial		472-4100	\$23	\$34
Charleston	El Rancho	304	727-2201	\$21	\$29
	Executive Inn		345-8820	\$34	\$38
	Holiday Inn Broad St		343-4661	\$40	\$44
	Holiday Inn Civic Ctr		345-0600	\$38	\$42
	Holiday Inn Downtown		344-4092	\$44	\$48
	Holiday Inn Elk River		323-4521	\$36	\$42

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

South Carolina

Charleston

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	747-0961	\$26	\$30
Days Inn	881-1800	\$18	\$24
Holiday Inn N	744-1621	\$36	\$46
Holiday Inn SW	556-7100	\$33	\$47
Howard Johnsons	524-4140	\$31	\$36
Ramada Inn	774-8281	\$33	\$40
Sheraton Inn	744-2401	\$34	\$42

Columbia

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	796-9400	\$29	\$34
Carolina Inn	799-8200	\$42	\$48

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

South Carolina

Charleston

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	747-0961	\$26	\$30
Days Inn	881-1800	\$18	\$24
Holiday Inn N	744-1621	\$36	\$46
Holiday Inn SW	556-7100	\$33	\$47
Howard Johnsons	524-4140	\$31	\$36
Ramada Inn	774-8281	\$33	\$40
Sheraton Inn	744-2401	\$34	\$42

Columbia

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	796-9400	\$29	\$34
Carolina Inn	799-8200	\$42	\$48

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

South Carolina

Charleston

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	747-0961	\$26	\$30
Days Inn	881-1800	\$18	\$24
Holiday Inn N	744-1621	\$36	\$46
Holiday Inn SW	556-7100	\$33	\$47
Howard Johnsons	524-4140	\$31	\$36
Ramada Inn	774-8281	\$33	\$40
Sheraton Inn	744-2401	\$34	\$42

Columbia

Area Code: 803

Motel/Hotel	Phone	Single	Double
Best Western	796-9400	\$29	\$34
Carolina Inn	799-8200	\$42	\$48

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

South Carolina

Charleston	Best Western	Phone: (803) 747-0961	Sngl: \$26	Dbl: \$30
Charleston	Days Inn	Phone: (803) 881-1800	Sngl: \$18	Dbl: \$24
Charleston	Holiday Inn N	Phone: (803) 744-1621	Sngl: \$36	Dbl: \$46
Charleston	Holiday Inn SW	Phone: (803) 556-7100	Sngl: \$33	Dbl: \$47
Charleston	Howard Johnsons	Phone: (803) 524-4140	Sngl: \$31	Dbl: \$36
Charleston	Ramada Inn	Phone: (803) 774-8281	Sngl: \$33	Dbl: \$40
Charleston	Sheraton Inn	Phone: (803) 744-2401	Sngl: \$34	Dbl: \$42

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

South Carolina

Charleston Best Western

Phone: (803) 747-0961 Sngl: \$26 Db1: \$30

Charleston Days Inn

Phone: (803) 881-1800 Sngl: \$18 Db1: \$24

Charleston Holiday Inn N

Phone: (803) 744-1621 Sngl: \$36 Db1: \$46

Charleston Holiday Inn SW

Phone: (803) 556-7100 Sngl: \$33 Db1: \$47

Charleston Howard Johnsons

Phone: (803) 524-4140 Sngl: \$31 Db1: \$36

Charleston Ramada Inn

Phone: (803) 774-8281 Sngl: \$33 Db1: \$40

Charleston Sheraton Inn

Phone: (803) 744-2401 Sngl: \$34 Db1: \$42

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

South Carolina

Charleston Best Western Phone: (803) 747-0961 Sngl: \$26 Db1: \$30
 Charleston Days Inn Phone: (803) 881-1800 Sngl: \$18 Db1: \$24
 Charleston Holiday Inn SW Phone: (803) 556-7100 Sngl: \$33 Db1: \$47
 Charleston Holiday Inn NW Phone: (803) 667-6131 Sngl: \$23 Db1: \$42
 Charleston Howard Johnsons Phone: (803) 524-4140 Sngl: \$31 Db1: \$36
 Charleston Ramada Inn Phone: (803) 774-8281 Sngl: \$33 Db1: \$40
 Charleston Sheraton Inn Phone: (803) 744-2401 Sngl: \$34 Db1: \$42

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very			Neither			Very
Slowly			Fast nor Slow			Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly			Neither			Strongly
Dislike			Dislike or Like			Like

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Concord	Holiday Inn	704	786-5181	\$29	\$34
	Quality Inn Colonial		782-2146	\$18	\$28
Dunn	Cricket Inn	919	892-8711	\$16	\$20
	Days Inn		894-2031	\$18	\$22
	Holiday Inn		892-2162	\$24	\$31
	Howard Johnsons		892-6181	\$21	\$26

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Concord	Holiday Inn	704	786-5181	\$29	\$34
	Quality Inn Colonial		782-2146	\$18	\$28
Dunn	Cricket Inn	919	892-8711	\$16	\$20
	Days Inn		894-2031	\$18	\$22
	Holiday Inn		892-2162	\$24	\$31
	Howard Johnsons		892-6181	\$21	\$26

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

North Carolina

City	Motel/Hotel	Area Code	Phone	Rates	
				Single	Double
Concord	Holiday Inn	704	786-5181	\$29	\$34
	Quality Inn Colonial		782-2146	\$18	\$28
Dunn	Cricket Inn	919	892-8711	\$16	\$20
	Days Inn		894-2031	\$18	\$22
	Holiday Inn		892-2162	\$24	\$31
	Howard Johnsons		892-6181	\$21	\$26

How quickly could you perform the described task using this display?

1	2	3	4	5	6	7
Very Slowly			Neither Fast nor Slow			Very Quickly

How much do you like the way this screen is arranged?

1	2	3	4	5	6	7
Strongly Dislike			Neither Dislike or Like			Strongly Like

Appendix D
Raw Error Data for Experiment I & II

EXPERIMENT I

SESSION	TASK	DISPLAY	NO. OF ERRORS	% OF ERROR(n=128)
1	1	1	0	.00
1	1	2	0	.00
1	2	1	0	.00
1	2	2	0	.00
1	3	1	8	.06
1	3	2	3	.02
1	4	1	5	.04
1	4	2	3	.02
2	1	1	1	.01
2	1	2	1	.01
2	2	1	0	.00
2	2	2	1	.01
2	3	1	2	.02
2	3	2	1	.01
2	4	1	3	.02
2	4	2	3	.02
3	1	1	1	.01
3	1	2	0	.00
3	2	1	0	.00
3	2	2	1	.01
3	3	1	4	.03
3	3	2	1	.01
3	4	1	2	.02
3	4	2	1	.01
4	1	1	1	.01
4	1	2	0	.00
4	2	1	1	.01
4	2	2	1	.01
4	3	1	2	.02
4	3	2	5	.04
4	4	1	1	.01
4	4	2	2	.02

EXPERIMENT II

DISPLAY	TASK	NO. OF ERRORS	% OF ERROR (n=768)
1	1	5	.007
1	2	6	.008
1	3	13	.017
1	4	15	.020
2	1	3	.004
2	2	2	.003
2	3	12	.016
2	4	17	.002
3	1	3	.004
3	2	1	.001
3	3	21	.027
3	4	27	.035