INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION, *18*(2), 167–182
Copyright © 2005, Lawrence Erlbaum Associates, Inc.

Usability of Graphical Icons in the Design of Human–Computer Interfaces

Jerzy Grobelny
Institute of Industrial Engineering & Management
Technical University of Wroclaw

Waldemar Karwowski
Center for Industrial Ergonomics, Department of Industrial Engineering
University of Louisville

Colin Drury
Department of Industrial Engineering
University at Buffalo

This study investigated usability of graphical icons in the design of human–computer interfaces. A simple additive model of icon acquisition time, including terms for Index of Difficulty, mode of icon array, and $\log_2$ (number of icons) explained 95% of the variance in mean times covering all 36 conditions over 4 laboratory experiments conducted in the study. The acquisition of icons in graphical menus was subject to Fitts' Law. Where the dialog box only appeared on the screen when an action was initiated, the acquisition time included a choice time that was additive with the movement time. When a choice was required during the task in addition to a movement, both the number of icons and their configuration affected the choice time. Design implications are that the number of icons should be minimized in pop-up menus and that the icons should be arranged in a manner reflective of the shape of the useful field of view. Frequently used menus or icon arrays should be permanently visible to minimize performance time (e.g., by using task bars to hold common icons). Dialogue windows, especially with a larger number of icons, should be built in compact, perhaps square, configurations. If it is not possible to use the square icon configuration, it is desirable to use the horizontal configuration, which can be utilized more efficiently than the vertical configuration

1. INTRODUCTION

Contemporary computer systems utilize sophisticated graphics enabling development and use of icons, pictograms, or symbols for functional keys, which are meta-

Requests for reprints should be sent to Dr. W. Karwowski, Center for Industrial Ergonomics, Lutz Hall, Room 445, University of Louisville, Louisville, KY 40292. E-mail: karwowski@louisville.edu

phors of real icons, commands, and processes modeled by designers. Pointing and clicking on icons or windows using the computer mouse are considered standard elements of communication between human operator and computer software. The basic icons of human–computer dialogue, including graphical icons, has become more and more uniform and homomorphic regardless of the computer operating system. Human–computer communication using the method of pointing and confirming (clicking) was introduced into computer systems developed under the mode of direct manipulation (DM; Ziegler & Fahnrich, 1988). The dominant tool for the purpose of such communication is a computer mouse.

The basic studies of different methods and techniques for selection of graphical icons using the "point and confirm" method originated in the 1960s. In particular, of special interest were text editors. For example, English et al. (1967) compared the times of positioning a cursor on a computer screen using a mouse, light pen, and tablet, and concluded that the mouse was the fastest tool. During the 1970s, attempts were made to systematically describe the process of selecting icons on the computer screen using Fitts' Law. Card, English, and Burr (1978) described the results comparing the quality of a computer mouse, joystick, and keyboard as tools allowing selection of a specific piece of text. The analysis showed that the computer mouse was the most effective tool with respect to positioning time, defined as the time from beginning of cursor movement to making a confirming click on the selected object. The computer mouse also showed the greatest resistance to human operator's errors. Fitts' Law became a basis for studies on selecting graphical icons reported by Epps (1986). The experiment described in his work required pointing a cursor, which was placed in a randomly selected position of a computer screen, on a square-shaped target. The developed equation for a computer mouse was quite different than the one reported by Card et al. (1978) with respect to the equation coefficients. This indicates the complexity of human cognitive processes and presence of many unidentified factors that affect coordination of the hand–eye system. A review of applications of Fitts' Law to study different pointing tools was reported by MacKenzie (1992).

2. STUDY OBJECTIVES

This study investigated the process of using graphical icons in the design of human–computer interfaces. The acquisition of icons in graphical menus was studied using Fitts' Law. The icon acquisition time was studied in terms of Index of Difficulty, mode of icon array (configurations), and number of icons used. Fitts' Law describes eye-controlled motor activity. However, the problem of selecting icons on a computer screen is more complex. Typically, not all the options of a given menu system can be seen simultaneously on the computer screen. Furthermore, before graphical icons can be selected, they must first be identified. In such a case, the selection process must consider not only the human motor activity, but also human cognition (Card, Moran, & Newell, 1983). The original modeling work by Card et al. (1983) was especially important in this respect. In particular, Card et al. investigated the relations that describe time of simple motor and cognitive tasks. Gillan, Holden, Adam, Rudisill, and Magee (1992) investigated point–click and point–drag movements and showed that the structure of such movements depends on the specificity of the performed tasks, including knowledge of the users' sequence of movements, direction of movement, typical movement amplitudes, and target sizes. Our research goes beyond the studies reported by Gillan et al. (1992), investigating the aspects of timing of target presentation and the way target grouping influences the degree to which other processes than those accounted by Fitts' Law need to be taken into account. In particular, possible components of icon acquisition were examined, where cognitive and motor functions may overlap or interfere under some circumstances.

Despite their prevalence in industry and computer operation, studies of combined cognitive and motor control tasks have not been common. Tasks such as sorting letters into mail slots, visual inspection, and icon acquisition require the operator to make a destination decision and then to execute that decision by making the appropriate movement. A very early study by Crossman (1953) of a card-sorting task had participants sort cards by red/black, suits, numbers, and so forth into different numbers of piles, where the number of piles and their probabilities represented different amounts of choice information. He used Hicks' Law to show that the time to sort each card was the usual linear function of the amount of information in bits/card. He noted, however, that when the deck of cards was held face-up instead of face-down, times were generally shorter, and by approximately the magnitude of the movement time. The data could be interpreted as the total time being either the sum of choice time and movement time (face-down) or the maximum of choice time and movement time (face-up). Does such an effect occur in icon acquisition, where under some circumstances the final target icon is available before the movement, but in other cases it may only appear when the movement is required to begin?

A comprehensive review of combined cognitive and movement tasks was presented by Hoffmann and Lim (1997). They reviewed the literature on both sequential and concurrent tasks, showing that for almost all sequential tasks, the best fit to the data was a strictly additive form:

$$\begin{aligned} &\textit{Total Time} = \textit{Choice Time} + \textit{Movement Time} \\ &\textit{or Total Time} = a + b(H) + c(ID) \end{aligned} \tag{1}$$

where H is the information processed and ID is the Fitts' Index of Difficulty (see next section).

In contrast, concurrent tasks, including their own experiment, showed interference between choice and movement control. In their own task, the exact location of a target was not revealed until 30 msec after the start of the movement. Hoffmann and Lim's interpretation was that interference occurred because of common processing limitations, which could arise either from the Wickens model or Multiple Resource Theory from the use of the same brain hemisphere for processing. Similar resources may need to be shared as both the choice and movement are tasks with manual outputs. Because the left and right hemispheres of the brain control contralateral movements, then the right hand movement (most participants are right-handed) must use the left hemisphere, which is also required for hemispheric overflow of verbal reasoning (Hellige & Cox, 1976), resulting in interference. There is evidence that combined choice and movement control tasks may be sequential (additive model), partially overlapped (max[choice time, movement time]), or interfering (Wickens' model). For an icon acquisition task typical of modern interfaces, there is a natural question of which model will be appropriate.

3. APPLICATION OF FITTS' LAW FOR POINTING AND SELECTING ICONS

An important part in the selection process for a graphical object (icon) is the movement to the target. The model that has been typically used to describe the time of such motion is Fitts' Law, which in its original form (Fitts & Peterson, 1964) relates movement time to Index of Difficulty (ID):

$$MT = a + b(ID) \tag{2}$$

ID was originally defined as:

$$ID = \log_2\left(\frac{2A}{W}\right) \tag{3}$$

where A denotes motion amplitude and W denotes target width (tolerance). The fit to data can be improved especially when ID is small, using the Index of Difficulty proposed by (Welford, 1968):

$$ID = \log_2\left(\frac{A}{W} = .5\right) \tag{4}$$

In applications to computer tasks, Mackenzie (1992) found the movement time in msec as:

$$MT = 12.8 + 94.7ID \tag{5}$$

using the original formulation of ID in Equation 2.

The task in Fitts' experiment was alternate tapping between two vertical ribbons of different widths. A similar relation for MT was reported by Fitts and Peterson (1964) for the task of discrete target reaching, with no return motion possible or allowed, although a smaller coefficient of ID of 74 was found.

The reciprocal of ID has been traditionally interpreted as an index of performance (IP) describing the efficiency (information capacity) of the eye-hand syste–while performing specific activity, measured in bits per second (Mackenzie, 1992). Greater coefficients for ID indicate poorer performance of the motor system on a given task. In Equation 5, the IP was equal to 10.6, while in the discrete task described by Fitts and Peterson (1964), the IP was 13.5, making this an easier task than reciprocal tapping.

There were also differences (also pointed out by MacKenzie, 1992) in the results of studies on pointing to icons on a computer screen using the framework of Fitts' Law. The task of interest in the study (i.e., pointing with the computer mouse) was first investigated by Card et al. (1978). Using Equation 2 with the Welford Index (4), they reported the results of their experiments, using the task of marking of rectangular fragments from typed text:

$$MT = 1030 + 96ID; \quad \textit{with an } IP = 10.4 \tag{6}$$

Epps (1986) reported similar studies, which investigated the process of selecting a square on the computer screen (with differing dimensions and location), leading to the following equation:

$$MT = 108 + 392ID \tag{7}$$

With an IP = 2.6, this value differs from the previous IP (see Equation 6) almost fourfold. Comparing the aforementioned results, MacKenzie (1992) asked the question "What then is the equivalent of Fitts' Law for a computer mouse?" In his analysis he postulated the following potential sources of variance:

1. Differences in mouse construction (the parameters of different mice are not precisely described and the studies were performed 8 years apart).
2. Methods of measurement of an Index of Difficulty; while in both cases the Welford Index was used (3), Card et al. (1978) determined the target width (W) based on width of the word being targeted. Such a method was not always rational because the actual width of a target was dependent on the approach angle to the target. On the other hand, Epps (1986) utilized square targets and, therefore, the effect of an approach angle in his experiments was minimal.
3. Methods of error treatment; while Card et al. (1978) did not analyze the trials with errors, Epps (1986) requested participants to repeat clicking in case of activating the mouse outside the target space.
4. To these it can also be added that the tasks differed considerably and Fitts' Law has typically been sensitive to task details.

Large differences in the equations for MT discussed earlier, which refer to simple pointing actions for single icons shown on a computer screen, prompted us to investigate the possible Fitts' relation for selecting icons by clicking on relevant objects from computer menus. Such selection tasks are universally used options in many contemporary computer programs. In the first phase of this study, the investigations were conducted for different ID values generated through combination of an object (graphical icon) size and its location on the computer screen. In the second phase of the study, the experimental task was enlarged. Selecting an object involved the choice from a set of icons, with different sets of menus displayed as dialogue windows, which could only be seen during the icon selection process. Martin (1988) showed that square keys were the best among many possible configurations that could be used for design of touch computer screens and square icons have simpler forms of Fitts' Law (Hoffmann & Sheikh, 1994).

4. METHOD

Four separate experiments were conducted to test progressively various aspects of icon configuration. They all used the same computer programs for presenting stimuli but used different groups of participants. The four experiments were as follows:

- Experiment 1: A baseline experiment to determine how Fitts' Law applied to the task of reaching to rectangular icons, using three different icon sizes and three different configurations.
- Experiment 2: An extension of the icon selection task to larger sets of icons (i.e., from 8 to 16). This changed the information in each choice, H.
- Experiment 3: Extension to the case where icons only appear after a control action (i.e., the mode of presentation changed). This removed the possibility of overlap processing between choice and movement.
- Experiment 4: A repeat of Experiment 2 using the nonoverlap condition of Experiment 3.

4.1. Methods and Procedures

In the MS Windows™ environment, the most often used configurations of graphical icons appear to be horizontal, vertical, and square layouts, so these were compared using icons of three sizes, with a large size of 16mm, a medium size of 12mm, and a small size of 8mm. The horizontal icon configuration contained all elements in one row; the square icon configuration contained all elements arranged in a square configuration whereas the vertical icon configuration contained all elements in one column. In each screen, there were either 8 (Experiments 1 and 3) or 16 (Experiments 2 and 4) basic graphical elements (icons), with one additional element "OK." To test different indexes of difficulty, a computer program was written based on Excel™ 4.0 to present the different window layouts and icon configurations. A Pentium 100 PC with the 15-inch monitor (800 × 600 pixels) was used in all experimental trials.

Figure 1 illustrates examples of the dialogue windows used in Experiment 1. A total of nine combinations (3 window structures × 3 window sizes) were used as experimental trials. The program could place the predesigned graphical structures at any field of the computer screen. This program was used to randomly generate instructions for participants to search for a given icon and recorded the icon selection times and errors. The errors were defined as those actions that resulted in clicking with a mouse outside the field of the target icon.

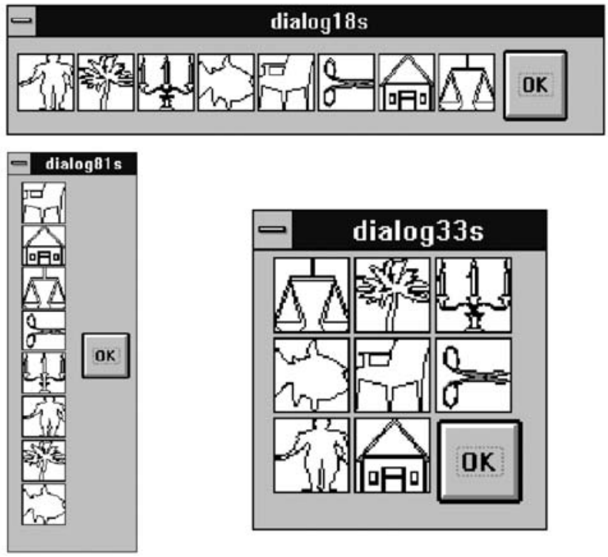

FIGURE 1 Examples of tested panels (icons and icon layout configurations).

The icon selection time was measured from the moment of pressing the "OK" key (always located in the middle of the screen) to the moment of confirming the icon selection, by clicking with a mouse when the cursor was placed within the desired icon. To eliminate potential learning effects due to the shape or meaning of graphical objects, icons used in the experiments (see Figure 1) were simple pictograms that were known to the participants and were uniquely identifiable (Paap & Roske-Hofstrand, 1988; Ziegler & Fahnrich, 1988).

The icon window was also placed on the computer screen so that the horizontal layout was placed in the upper part of the dialogue window, whereas the vertical layout was placed near the left edge of the window. In the case of a square layout configuration, the left upper corner of the window was arbitrarily preselected as the location for this mode of icon grouping. For each of the nine experimental trials, the ID was established based on direct measurement of the magnitude of cursor motion on the computer screen and icon size (8mm, 12mm, and 16mm). The ID values for each configuration were calculated using Welford's (1968) equation, with the value of A equal to the mean movement distance of the cursor on a given task.

In Experiments 2 and 4, the interface was identical except that the number of icons was increased from 8 to 16. The geometrical icon configuration was such that vertical windows contained two columns of icons, whereas the horizontal windows contained two rows of icons, using a total of 16 icons on each screen (plus the icon "OK").

In Experiments 3 and 4, the Mode of presentation of the icons changed. Windows with icons were displayed on the computer screen after confirmation of a given task and simultaneous activation of a timer, and were removed from the screen after the specific icon was selected with a computer mouse. The original condition (Experiments 1 and 2) was known as the Menu Mode, whereas this new condition (Experiments 3 and 4) was called the Dialog Box Mode.

4.2. Participants

In each experiment, 10 participants were selected from students and technicians who used computers every day with the MS Windows interface environment. The age ranges for the experiments were:

- Experiment 1: 16–53 years
- Experiment 2: 25–59 years
- Experiment 3: 16–41 years
- Experiment 4: 25–59 years

4.3. Experimental Procedures

The experiment simulated casual work of advanced computer users. As all participants were very experienced computer and mouse interface users, no special training was used before the experiment and the learning process was not analyzed. An initial session taught identification of the selected icons and structures. Each participant undertook a short training exercise to learn the experimental computer program and was briefed about the purpose of the study and methods for experimental trials. A series of pre-experimental trials were conducted, during which each participant experienced different icon dimensions and configurations in a random order. During this practice session, the participant was asked to place the dialogue window in a predefined location of the computer screen and perform nine icon selections in each location. The aforementioned conditions were randomly generated and presented to each participant in the same order for any given experimental set.

In each experiment, each participant performed the experimental trials in a random order, and the times associated with each of the nine trials were recorded. In Experiments 1 and 2, the dialogue window was shown on the computer screen all the time so that time values reflected only the movement time measured from the middle of the screen to the given object, plus the time required to perform a left mouse click on the icon. In these experiments, the search for graphical icons could be conducted before activation of the timer with the "OK" key placed in the middle of the computer screen, simulating the process of selecting graphical icons where the location was well known to participants. This Menu Mode of human–computer interaction is directly related to utilization of the menu with graphical icons and a set of simulated software keys shown on the computer screen, which the user may call up to remain on the screen. Such a human–computer interface where icons are located in constant positions is used widely in computer-aided graphics programs (i.e., various CAD systems). In Experiments 3 and 4, a Dialog Box Mode was used in which the icon was only shown when an action was taken by the participant. This was typical of pop-up dialog boxes, where the next choice is only revealed after a specific action.

4.4. Results

It should be noted that the number of observed errors was small. The only errors occurred for the small- and middle-size icon sets, and these accounted for less than 2% for all experiments.

Initially, each experiment was analyzed separately, but these results are not reported here to ensure clarity of presentation of the main issues in the study (i.e., the effects of Index of Difficulty, Information per choice, and mode of operation [dialog vs. menu] on task performance times). For reference, the actual mean selection times are presented in Table 1.

There was a significant effect of icon size in all experiments, as analyzed using Fitts' Law for each experiment. These regressions were all significant except for Experiment 4, where the significance level only reached $p = .065$. All are shown in Table 2.

An analysis of variance (ANOVA) was conducted on the mean times for each combination of icon size and icon configuration across the four experiments. The four experiments were further classified as two factors: Number of Icons (8 or 16) and stimulus Mode (Menu Mode or Dialog Box Mode). This analysis was then used to guide the derivation of regression models to fit the overall data so that interpre-

Table 1: Summary of Average Selection Times (in msec) for All Experiments

		Icon Configuration		
Experiment	*Icon Size*	*Horizontal*	*Square*	*Vertical*
Experiment 1	Small (8 mm)	929	973	972
(8 icons, menu mode)	Medium (12 mm)	870	885	955
	Big (16 mm)	779	756	852
Experiment 2	Small (8 mm)	907	973	927
(16 icons, menu mode)	Medium (12 mm)	753	816	788
	Big (16 mm)	726	671	724
Experiment 3	Small (8 mm)	1823	1531	1705
(8 icons, dialog box mode)	Medium (12 mm)	1572	1446	1830
	Big (16 mm)	1459	1342	1583
Experiment 4	Small (8 mm)	1840	1779	1930
(16 icons, dialog box mode)	Medium (12 mm)	1650	1635	1806
	Big (16 mm)	1785	1694	1860

Table 2: Summary of Fitts' Law Regressions Across All Experiments

Number of Icons	*Menu Mode*	*Dialog Box Mode*
8	Experiment 1	Experiment 3
	Time = 239 + 201 ID	Time = 589 + 310 ID
	$R^2 = .916$, $p = .001$	$R^2 = .490$, $p = .036$
16	Experiment 2	Experiment 4
	Time = 210 + 201 ID	Time = 1358 + 140 ID
	$R^2 = .743$, $p = .003$	$R^2 = .0.406$, $p = .065$

tation could be made of the applicability of an additive model of the components of the acquisition times. There were significant main effects of all the variables: Mode, $F(1, 35) = 4488$, $p < .001$; Number, $F(1, 35) = 20$, $p = .011$; Configuration, $F(2, 35) = 31$, $p = .004$; and Size, $F(2, 35) = 65$, $p = .001$. There were three significant two-way interactions: Number × Size, $F(2, 35) = 9$, $p = .028$; Mode × Configuration, $F(, 35) = 22$, $p = .008$; and Mode × Number, $F(1, 35) = 112$, $p < .001$; and one significant three-way interaction, Number × Mode × Size, $F(2, 35) = 8$, $p = .038$, that was not pursued further as the subsequent regression analysis better captured this effect. The effects of Size and Configuration are also combined in the Fitts' Law analysis as both contribute to the ID value.

Figure 2 presents the Mode × Number interaction, where the main effect of Mode is readily apparent as well as the difference between the 8-icon and 16-icon times for Experiments 3 and 4 where the icon array only appeared after the timing started. One interpretation of Figure 2 is that there is no increase in time with number of icons where the icon array is always visible because the choice has already been made when the timing starts. Thus factors affecting choice between different numbers of alternative times (e.g., Hicks' Law) will be irrelevant to the measured time. In contrast, where the choice time must be added to the movement time, any increase in number of choices will indeed affect the measured time.

Figure 3 gives Mode × Configuration interaction. The significant effect of configuration for the Menu Mode case (Experiments 3 and 4), compared with a minimal effect for the Dialog Box Mode condition, is evident. Tukey post hoc tests showed all differences between configurations to be significant for the Appears experiments. The major change is between the two linear arrangements and the square configuration, with the square being over 100 msec faster than either linear alternative.

To explore the Fitts' Law effects, note from Table 2 that all but one experiment produced an individually significant effect of ID on time. There were also signifi-

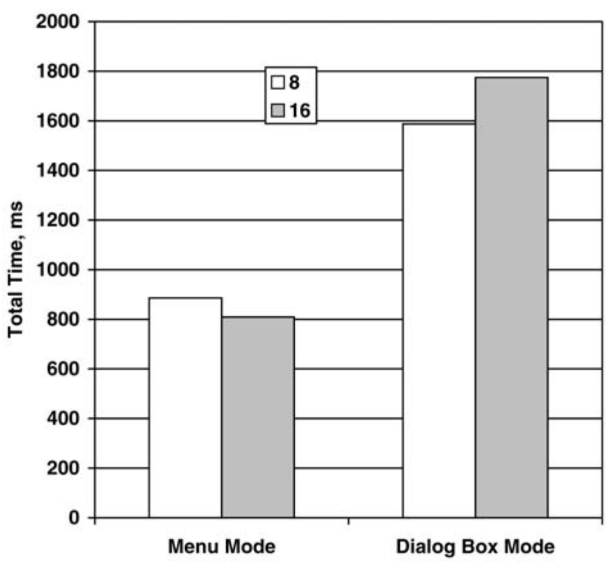

FIGURE 2 Differences in performance between the four experiments for small (8 mm) and big (16 mm) graphical icons.

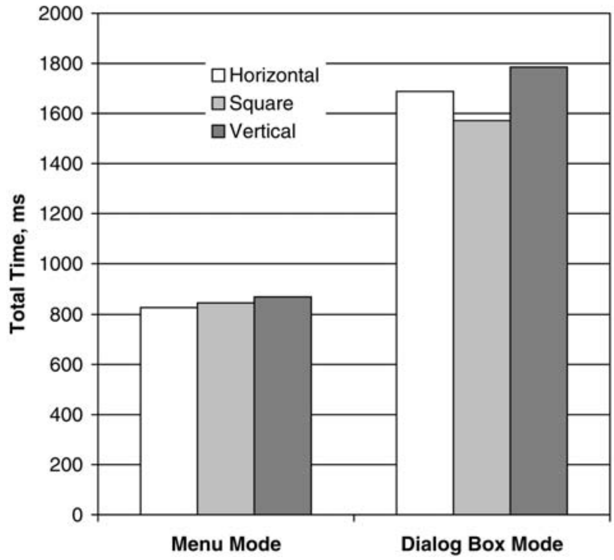

FIGURE 3 Effects of Mode × Configuration interaction in performance times.

cant main effects for Icon Size and Configuration in the overall ANOVA, and so one would expect that the overall Fitts' Law model would be a useful summary of the total data set. In fact, if the overall ANOVA is changed to include only Experiment (or equivalently Mode and Number) with ID as a covariate, the model accounts for 97% of the total sum of squares. In contrast, the earlier full four-factor model with no covariate accounted for 99.9% of total sum of squares. This Fitts' Law formulation may not be perfect, but it does account for a very high fraction of the variability in the complete data set. A regression approach was therefore used to fit a single equation to all of the mean data in the four experiments by using the factors of ID, Mode (1 or 2) and the information in the number of icons (H = 3 or 4 bits). The result was:

$$Total\ Time = -999 + 208ID + 834Mode + 105H \tag{12}$$

With $r^2 = .947$, $p < .001$. All four coefficients were significant individually at $p < .01$. Figure 4 shows these results with the fitted lines. Note that in an additive regression equation, all lines are restricted to the same slope (208 msec/unit ID), even though there was a range of different slopes in the four experiments (Table 2).

5. DISCUSSION: DESIGN OF DIALOGUE WINDOWS

The questions posed in the Introduction concerning the applicability of Fitts' Law to icon acquisition, the configuration of dialog boxes, and the way in which dialog boxes appear onscreen can now be addressed in light of the findings of these four experiments. This discussion can be approached through the overall analyses and regressions to find where common effects arise in the studies.

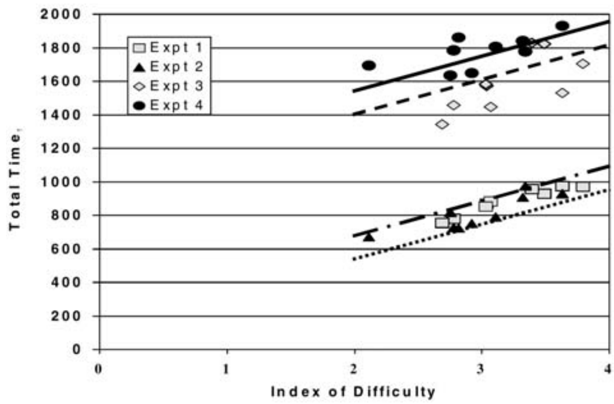

FIGURE 4 Fit of additive model to data from all experiments.

The individual experiment regressions and the overall regression equation showed once again that Fitts' Law explains a useful fraction of the variance in movement times for mouse movements to locations on a computer screen. As Fitts' Law combines the position information and icon size information, it continues to provide a useful metric for designers to make best use of screen real estate in new interfaces. As several authors, such as Gillan et al. (1992), Epps (1986), and McKenzie (1992), pointed out, Fitts' Law should and does apply in one form or another, as it predicts a number of movement actions on a computer interface. With the slope coefficients well within the range reported in other experiments, this was confirmed. The individual slopes ranged from 140 to 310 msec/bit and the common slope in Equation 15 was 208 msec/bit. Quoted values in the literature range from about 90 (Card et al., 1978; Gillan et al., 1992; Mackenzie, 1992) to 392 (Epps, 1986). Different tasks (e.g., point and click vs. drag and drop) may well give rise to detailed differences in movement times, but all appear to be manifestations of Fitts' Law, especially if the definition of target size is made specific to the task (Gillan et al., 1992). Interestingly, in the two experiments (1 and 2), which found the best fit in terms of r^2, the coefficients were almost identical at about 200 msec/bit. In Experiments 3 and 4, where the times were much longer and there was an additional task of finding the correct icon before movement could commence, the r^2 values were much lower and the coefficients more variable (about 150 and 300 msec/bit). It would be expected that the more the overall performance is dominated by the movement alone, the better Fitts' Law will be able to predict the outcomes.

In fact, the outcomes could be predicted quite well with a simplified overall regression model based on Fitts' Law for the movement component and different factors for other components. Equation 12 had an r^2 value of about .95, despite neglecting some subtleties found in other analyses. First, it used an additive model where the Fitts' Law component was identical for all four experiments, despite reporting of different slopes for the four cases (Table 2). Second, the model treated the two between-experiment factors as strictly additive, despite their significant interaction (Figure 2). These effects can be seen in Figure 4, where the slopes of the lines are clearly a better fit for Experiments 1 and 2 than for 3 and 4. Also, Figure 4 shows that the two regression lines for Experiments 1 and 2 are the same distance apart as are those for Experiments 3 and 4, although the two data sets do not reflect this model at this level of detail. The fitting of a regression model including the interaction between Mode and Number of Icons was explored, and that model had an improved r^2 of .97. However, the model coefficients made little physical sense, being positive for the interaction term, but negative for the separate Mode and Number terms. The predilection was for a model that was closer to cognitive reality than to empirical performance, so that the extra predictive power was sacrificed for explanatory efficacy. As always, treating the experiments individually would allow better data-fitting at the expense of economy of modeling. For many design purposes, a model explaining 95% of the variance will be adequate for design use.

The overall ANOVA results and regression model both show the additive nature of the components of the mouse movements required for icon acquisition. The Mode factor was highly significant and indeed made a difference of 834 msec in Equation 12. The total time for the pop-up dialog box that appeared on a click was much longer, representing the addition of a choice time to the movement time, as predicted by Hoffmann and Lim (1997). There was no evidence that the choice time was anything but additive, in the sense that a model such as Equation 12 was an adequate fit to the data set.

Further evidence of additivity is that the factor expected to affect only choice time, Number of Icons in each dialog box, indeed affected only the conditions when the choice was part of the task (i.e., Experiments 3 and 4). For these, the number of icons was significant, as seen in Figure 2, with choice between 8 icons being about 200 msec shorter than choice between 16 icons. This was modeled as a straightforward information theory effect, essentially Hicks' Law (Hoffmann & Lim, 1997). This produced the H term in Equation 12, where H is the information per choice in bits, equal to $\log_2$ (Number of Icons). Others have suggested similar formulations of choice time in human–computer interaction. For example, Landauer and Nachbar (1985; described by Paap & Roske-Hofstrand, 1988) suggested a relation between icon selection time and number of icons. This relation describes a process of searching for numerical digits shown on the screen menu (i.e., a search with fitting of categories) also called the *class inclusion matching*. This type of search is more cost effective than the classical method of searching for known labels (identity matching), which exhibits linear dependency from the number of icons on the dialogue window (Paap & Roske-Hofstrand, 1988). With only two levels of number of icons, the form of the relation between choice time and number of icons could not be tested. However, an information theory interpretation would give rise to a logarithmic function, whereas a visual search model would suggest a linear dependency.

The visual search process, particularly where multiple fixations are required, is typically a self-terminating serial search, unless special conditions are met where parallel processing may be possible. The time to search a field is linear with field size or with number of items on the field (Drury & Clements, 1978). If stimulus (target and background) conditions are appropriate, and if the participant is highly experienced, then search becomes pre-attentive and does not require sequential search, at least within a single fixation (Wolfe, 1994). When a dialog box suddenly appears, there is no a priori reason why search should be pre-attentive, so that a linear relation between search time and number of icons would be expected. Only an experiment with more than two levels of number of icons could answer this question of relation form in the icon acquisition setting. In the absence of such evidence, the amount of information rather than number of icons in Equation 12 was used, because the two values (3 and 4) only differed by 33% rather than by 100% for number of icons. The difference between the times as a function of number of icons was smaller than the effect of Mode, so that this compressed function was deemed more appropriate. In addition, it fits well with the original Card et al. (1983) formulation. From a design viewpoint, obviously the fewer icons, the more rapid the choice, but only if the choice time needs to be an explicit part of task time. A second design implication would be that dialog boxes of icon arrays that are always present, at least within an application, will improve performance time over pop-up boxes. Thus, as one would expect, the Task Bar present in most graphical user interfaces is indeed a useful innovation.

The visual search process may also provide an explanation of the effects of icon array configuration. This factor was not included in the overall regression model as to some extent it is implicit in the Index of Difficulty calculations. Thus, with linear arrays the movement distances are different from those with a square array, as shown by the fact that in Figure 4 there are nine distinct ID values for each experiment rather than only three if all movement distances had been identical. However, configuration did make a significant difference to performance times, particularly where the dialog box appeared on a click rather than being present before the trial (Figure 3). Again, the absence of a similar affect for the Always On conditions (Experiments 1 and 2, see Table 2) argues that the locus of this effect lies in the choice component of the task rather than the movement part.

It is worth noting that not all movements in human–computer interaction can be described by Fitts' Law, because not all are limited only by end accuracy. For example, Accot and Zhai (1997) considered a class of movements where the path of the movement rather than its end-point accuracy is the limitation. These have been termed *Path Control* tasks in contrast to *Terminal Aiming* tasks by Montazer and Drury (1989) and for these the movement speed is linearly related to the lateral accuracy requirement. DeFazio, Wittman, and Drury (1992) studied path control tasks using a computer–mouse interface and found the same linear speed–accuracy relation as Accot and Zhai (1997).

6. CONCLUSION

The results of experimental studies discussed earlier allow for the following explanation of the process of using graphical icons in human–computer interfaces by experienced users to be proposed, and for its design implications.

1. The acquisition of icons in dialogue windows and, in particular, graphical menus, is subject to Fitts' Law. The slope of the line relating acquisition time to Index of Difficulty may vary between conditions, but as a first approximation, a similar functional form can apply to all the conditions tested, with a slope well within the range normally seen in mouse movement tasks. The fit of Fitts' Law implies

that minimizing movement distance and maximizing icon size will minimize movement time. In these studies, the performance criterion for using the mouse was to perform the task as quickly as possible and with no errors. The fact that there was a small errors rate (below 2% and smaller than those reported by Card et al., 1978) suggests that equal attention was given to both of these criteria.

2. Where the dialog box only appears on the screen when an action is initiated, the acquisition time includes a choice time that is additive with the movement time. For design of computer interfaces, the implication is that frequently used menus or icon arrays should be permanently visible to minimize performance time (e.g., by using task bars to hold common icons).

3. When a choice is required during a task in addition to a movement, a choice time will be affected by both the number of icons and their configuration. These effects can be explained in terms of visual search where multiple fixations are required. Design implications are that the number of icons should be minimized in pop-up menus and that the icons should be arranged in a manner reflective of the shape of the Useful Field of View. Dialogue windows, especially with a larger number of icons, should be built in compact, perhaps square, configurations. If it is not possible to use the square icon configuration, it is desirable to use the horizontal configuration, which can be utilized more efficiently than the vertical configuration.

4. A simple additive model of acquisition time, including terms for Index of Difficulty, Mode of icon array, and $\log_2$ (number of icons) was adequate to explain 95% of the variance in mean times covering all 36 conditions over the four experiments reported here.

REFERENCES

Accot, J., & Zhai, S. (1997). Beyond Fitts' Law: Models for trajectory-based HCI tasks. *Proceedings of ACM CHI '97, Conference on Human Factors in Computing Systems*, 295–302.

Card S. K., English W. K., & Burr B. J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection of CRT. *Ergonomics, 21,* 601–613.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human–computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Crossman, E. W. R. F. (1953). Entropy and choice time: The effect of frequency unbalance on choice-response. *Quarterly Journal of Experimental Psychology, 5*, 41–51.

Drury, C. G., & Clement, M. R. (1978). The effect of area, density, and number of background characters on visual search, *Human Factors, 20,* 597–602.

Epps, B. (1986). Comparison of six cursor control devices based on Fitts' law models. *Proceedings of the 30th Annual Meeting of the Human Factors Society*, 327–331.

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology*, *67*, 103–12.

Gillan, D. J., Holden, K., Adam, S., Rudisill, M., & Magee, L. (1992). How should Fitts' Law be applied to human–computer interaction? *Interacting With Computers*, *4*, 291–313.

Hellige, J. B., & Cox, P. J. (1976). Effects of concurrent verbal memory on recognition of stimuli from the left and right visual fields. *Journal of Experimental Psychology: Human Performance and Perception, 89,* 50–60.

Hoffmann, E. R., & Lim, J. T. A. (1997). Concurrent manual-decision tasks. *Ergonomics, 40,* 93–318.

Hoffmann, E. R., & Sheikh, I. (1994). Effect of varying target height in a Fitts' movement task. *Ergonomics*, *35*, 1071–1088.

Landauer, T. K., & Nachbar, D. W. (1985). Selection from alphabetic and numeric menu trees using a touch screen: Breadth, depth and width. In *CHI 1985 Proceedings (Special Issue Of SIGCHI)*. New York: Association of Computing Machinery.

MacKenzie, I. S. (1992). Fitts' Law as a research and design tool in human–computer interaction. *Human–Computer Interaction*, *7*, 91–139.

Martin, G. L. (1988). Configuring a numeric keypad for a touch screen. *Ergonomics, 31*, 945–953.

Montazer, M. A., & Drury, C. G. (1989). A Test of Begg's model for self-paced movements. *Ergonomics,* 32, 497–512.

Paap, K. R., & Roske-Hofstrand, R. J. (1988). Design of menus. In M. Helander (Ed.), *Handbook of human–computer interaction*. Amsterdam: Elsevier.

Welford, A. T. (1968). *Fundamentals of skill.* New York: Methuen.

Wickens, C. D., & Hollands, J. G. (2000), *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*, 202–238.

Ziegler, J. E., & Fahnrich, K. P. (1988). Direct manipulation. In M. Helander (Ed.), *Handbook of Human–Computer Interaction*. Amsterdam: Elsevier.