Issues Related to HCI Application of Fitts's Law

Charles E. Wright

University of California, Irvine

Francis Lee

University of California, Irvine

RUNNING HEAD: Fitts's Law Issues

Corresponding Author's Contact Information:

Charles E. Wright

Email: cewright@uci.edu

Phone: 949-824-7589

Fax: 949-824-2307

Brief Authors' Biographies:
Charles Wright is a cognitive psychologist with an interest in motor control, mechanisms for response selection, and visual attention; he is an Associate Professor in the Department of Cognitive Sciences of the University of California, Irvine. **Francis Lee** is a Psychology major at the University of California, Irvine.

ABSTRACT

Taking Fitts's law as a premise – i.e., movement time is a linear function of an appropriate index of difficulty – we explore three concerns related to the collection and reporting of these data from the perspective of application within HCI. The central issue is whether results obtained using blocked target conditions are representative of performance in situations in which, as is often the case, target conditions vary from movement to movement. Although varied target conditions lead to longer movement times, the effect is additive and surprisingly small, suggesting that evaluating devices or designs using blocked data may be acceptable. With Zhai (2004) we argue against the practice, advocated by the ISO9241-9 standard (ISO, 2000), of using throughput as a one-dimensional summary for comparisons of devices or designs. Also questioned is whether analyses using an accuracy-adjusted index of difficulty are appropriate in all design applications.

CONTENTS

1. INTRODUCTION

Fitts's law is a highly successful formulation that describes how the time to complete a movement depends on the distance to be covered and the spatial accuracy required. Although Fitts's law does not apply to all movements, the class of movements to which it does apply is large and of immense practical significance. This applicability has stimulated interest in this formulation beyond its basic research origins. This interest has both led to a large increase in the number of papers appearing in traditionally applied outlets that use Fitts's law and to an international standard that specifies how Fitts's law results should be used to characterize and compare input devices (ISO, 2000).

The purpose of this paper is neither to extend nor question Fitts's Law, which we will take as a given. Instead we wish to explore three issues that have arisen as Fitts's law has been applied in human-computer interaction (HCI). The theme that unites these three issues is a concern that ideas and practices, which emerged from the basic research that provided the underpinnings of Fitts's law, have been adopted by applied researchers without sufficient scrutiny. Recent standardization efforts have made these issues more salient. Although standardization will almost certainly produce research results that are more consistent, standardization could also have negative effects if the research produced either is reported incompletely or is misleading when generalized to the situations that are of practical interest.

The central issue we will explore is whether an experimental design choice, which may make sense in a basic-research setting and is often used in applied studies, appropriately reflects the real-world situations to which the results are to be generalized. The two remaining issues emerged as we reflected on recommendations in the literature for how to analyze and report our data related to this central issue.

1.1. Background

Fitts's law holds that the time, T, to complete a speeded movement to a target is a linear function of an index of difficulty, ID, characterizing the movement:

$$T = a + b\,ID\,. \tag{1}$$

The index of difficulty depends on the distance, D, from the starting point to the center of the target, and the width, W, of the target. The definition of ID has evolved since the initial, admittedly ad hoc formulation proposed by Fitts (1954): $ID = \log_2(2D/W)$. The version of the index of difficulty now typically used in HCI applications is (MacKenzie, 1992; Soukoreff & MacKenzie, 2004)

$$ID = \log_2\left(\frac{D}{W} + 1\right). \tag{2}$$

All the proposed modifications of Equation 2 have the more general form $T = f(D/W)$; where $f()$ is a simple (i.e., linear, logarithmic, or power) function of the dimensionless ratio D/W (Guiard & Beaudouin-Lafon, 2004).[1]

A second, important development related to Fitts's law has grown out of the observation that participants in these experiments often do not adjust their performance as much as might be expected when target width is changed. Specifically, changes in endpoint variability are typically much smaller than would be expected when target width is manipulated. To compensate for this, Welford (1960; 1968, pp. 147-148) suggested that *ID* be replaced with an *effective* index of difficulty, ID_e, in which an estimate of the effective target width, W_e, computed from the movement endpoints, replaces W, the nominal target width.

$$ID_e = \log_2\left(\frac{D}{W_e} + 1\right) \tag{3}$$

Although Fitts used an accuracy adjustment in his later work (Fitts & Radford, 1966)[2], it has not always been adopted (MacKenzie, 1992).

Over the last half century, Fitts's law has been well-studied and has proven to be highly successful. Data obtained using a large variety of input devices across a broad array of conditions are well fit by Equation 1 with R^2 values of .80 or higher (Plamondon & Alimi, 1997); the variance accounted is generally highest when Equation (2) is used to calculate the index of difficulty and when the accuracy adjustment is used.

Fitts's papers contained elements of both basic and applied research. His formulation grew out an effort to understand human performance from the theoretical perspective of information theory. However, from a more practical perspective, he also proposed an index of performance (Fitts, 1954, Eq. (2)),

$$IP = {}^{ID}\!/_{MT} \tag{4}$$

as a measure of *throughput* combining both speed and accuracy. This measure, which has units bits of information per unit of time, was adapted from information theory where it is used as a measure of the channel capacity. Fitts's expectation was that throughput would be a constant that could be used to characterize and compare operator performance with different devices and in different movement contexts.

---

[1] In our opinion, the "best" formulation for the index of difficulty is a power function: $(D/W)^p$, where p is a fractional exponent ($0 < \rho \le 1$). Kvålseth (1980) first made the case that this form generally fits movement time data better than one based on the logarithm, the shape of which approaches that of the power function as *p* gets close to zero. We prefer this form because of the theoretical and empirical justifications for it provided by Meyer, Smith, Kornblum, Abrams, and Wright (1990). However, for descriptive purposes, Equation (2) is fine, and, since it is widely used in this literature, we will use it here.

[2] Although Fitts and Radford used the accuracy adjustment, the adjustment they used was based on the proportion of movements that were errors – ended outside of the target region – not the recorded movement endpoints.

1.2.Fitts's Law in Basic Research and Applied Settings

Although the specific form for the index of performance has been a subject of debate, Fitts's idea of using these or similar results to characterize movement situations and input devices has become increasingly influential in HCI (MacKenzie, 1992), especially after Card and his colleagues used the results from an application of Equation 1 (Card, English & Burr, 1978) to justify commercialization of the mouse by Xerox. Of specific interest to HCI researchers, Fitts's law has been found to apply to pointing and dragging using a mouse, trackball, stylus, joystick, and touchscreen. The results have been used both to assess and compare throughput and as part of larger models to predict performance in new user interfaces (e.g., Card, Moran, & Newell, 1983). This has led to the promulgation of an international standard, ISO9241-9, that provides guidelines for such evaluations (ISO, 2000). Other, more detailed, recommendations for how these evaluations should be conducted have been proposed by Soukoreff and MacKenzie (2004). These attempts at standardization are important if they help reduce the confusion in the literature due to conflicting results arising from methodological differences (MacKenzie, 1992). To evaluate this possibility for pointing movements made with the mouse, Soukoreff and MacKenzie (2004) compared 9 studies of that followed ISO9241-9 and 24 studies that did not. They found a dramatic increase in consistency of the results for the studies that followed ISO9241-9.

1.3.Issue 1: External Validity of Fitts's Law Studies

Inferences about cause-effect relationships based on specific scientific studies are said to possess external validity if they may be generalized from the unique and idiosyncratic settings, procedures and participants of those studies to other populations and conditions. This issue is often critical in design applications where the goal is to use information derived from research settings to make design decisions. Of course, the best way to settle external validity concerns is a replication using the settings, procedures, and participants of the intended application. However, for obvious practical reasons designers often prefer to generalize results from available, prior research when making design choices. Our central concern here is that the methodology of much of the research using Fitts's law, including studies adhering to the suggestions of ISO 9241-9 (ISO 2002) and Soukoreff and MacKenzie (2004), may generalize poorly to the situations typically encountered in the HCI applications of that research: i.e., the coefficients of Equation (1) derived from such research may deviate systematically from those obtained using procedures more like the situations encountered in the application environments.

The procedural aspect of particular concern here is the blocking of target conditions, where "target conditions" refers to combinations of D and W. Recommendation II of Soukoreff and MacKenzie (2004, p. 755) is in line with the practice followed in many studies based on Fitts's law. They suggest studying a variety of target conditions that include multiple levels of D and W chosen so that nominal ID values associated with the target conditions span a range between 2 and 8 bits. Each target condition should be presented enough times – they suggest between 15 and 25 – that an accurate estimate of the central tendency can be ascertained for each participant using each target condition. Although there is no specific recommendation to this effect, either by Soukoreff and

MacKenzie or in ISO 9241-9, a natural way to structure the repeated presentations for a particular target condition is to block them: i.e., present a sequence of trials all having the same target condition. Blocking of the target conditions certainly is not necessary. Pastel (2011) is an example of a study that randomizes the target conditions. However, of the nine studies cited in Soukoreff and MacKenzie (2004) as examples that have followed ISO 9241-9, the six that we could obtain and that included enough detail to determine how the conditions were ordered, all blocked target conditions. It is also suggestive that some studies report not only blocking target conditions but discarding the first several trials within each block so that the data analyzed would better reflect optimum performance possible in that target condition.

Blocking target conditions and discarding initial trials within blocks may make perfect sense in the context of basic research. However, it is far from clear that the results obtained using these procedures accurately describe performance in typical HCI applications. Making matters worse, we know of no studies that include or allow direct comparison of results obtained using blocked and fully varying target conditions.[3] The study reported here addresses that question.

In addition to manipulating the order of target conditions, this study also compares performance for discrete and continuous[4] movements. In discrete movements, the participant moves to a starting point and then, after a signal, initiates a movement to the target. After a pause, and possibly some feedback, this procedure is then repeated. In the continuous movement task, after completing a movement, the participant immediately initiates a subsequent movement, typically in the opposite direction; the process is repeated until the full sequence is done. Each of these procedures seems more or less like some HCI situations.

Unlike the manipulation of blocked versus varied target conditions, there is some data about the comparison of discrete and continuous movements. In his first paper, Fitts

[3] This statement is qualified because there have been several studies (e.g., Megaw, 1975) that looked at limited variations of the target conditions in a continuous, reciprocal, stylus-tapping task (like that studied by Fitts, 1954; see below). One experiment included conditions in which *D* was constant but the widths of the left and right targets were different. The analysis looked at MT as a function of the width of that target and the target of previous movement – since this was a continuous, reciprocal task, the endpoint of the previous movement is the starting point for the current movement. These were inversely related: MT time decreased substantially as the width of the previous target increased. Particularly striking was the observation that, when the previous target was larger than the current target, MT was faster than it was in a condition in which the width of the two targets was identical. Megaw also studied successive movements that had different *D*s, but constant *W*. In this condition, MT was about 35 ms slower and did not depend systematically on the target conditions of the previous movement.

[4] Our use of the term "continuous" for this condition follows the terminology introduced by Fitts and Peterson (1964). ISO 9241-9 (ISO 2002) and Soukoreff and MacKenzie (2004) use "serial" to describe the same condition. We are using the original terminology because it seems more accurately descriptive.

(1954) used a continuous, stylus-tapping task. A decade later, he (Fitts & Peterson, 1964) used the same apparatus in a discrete version of the task. Figure 2 of this second paper compares the data from these two experiments. In this figure, *MT* for the continuous task is longer than that for the discrete task across the full range of *ID* values studied and gets larger as *ID* increases: for an *ID* of 2, the difference was roughly 100 ms and it increased to roughly 210 ms for an *ID* of 7. Fitts and Peterson (1964) expected these results both because *MT* in their continuous task included between movement latencies that were excluded in their discrete task and, more fundamentally, because, in the discrete task, the participant starts each movement after having had time to program its parameters. However, the interpretation of their data is clouded because the average error rates were quite different in the two tasks: for the discrete task there were 10.5% target misses on average, almost ten times as high as the 1.2% for the cyclical task.

Guiard (1997) provides a direct comparison of discrete versus continuous movements using a linear positioning task. To make the conditions more similar, participants pushed a button on the manipulandum to signal the end of a movement and before initiating a discrete movement. Replicating the conclusion derived from Fitts's experiments, the slope of the function relating *MT* to ID_e was larger in the continuous condition (277 ms / bit) than in the discrete condition (205 ms / bit), and the direction of this difference was the same for all six participants. However, unlike the data from Fitts, Guiard found that, for all six participants, these functions crossed somewhere in the ID_e range between 2 and 6, and that this point of intersection was strongly correlated with a participant's overall movement time: the faster the participant the higher the ID_e at the point of intersection.

This manipulation of discrete versus continuous movements was included here because, for two reasons, we suspected that any differences between the blocked and varied target conditions might be larger with the continuous procedure than with the discrete procedure. From an information processing perspective and consistent with the expectation of Fitts and Peterson (1964), the implicit pressure to spend less time on movement planning in the continuous condition might impose a larger penalty when the target conditions are varied than when they remain constant. From the perspective of dynamical systems applied to the Fitts task (Guiard, 1993), the ability to recycle kinetic energy in the continuous task appears to depend on the harmonicity of the repetitive movements. Guiard (1993; 1997) has shown that, with target conditions blocked, this advantage is reduced for more difficult movements. It seems plausible that this advantage might also be reduced when the target conditions are varied.

1.4. Issue 2: Summaries Based on IP, the Index of Performance

Having collected the data to evaluate Issue 1, we encountered several issues about how to analyze and report it. As noted in the Background section, Fitts (1954) initially proposed the measure of throughput in Equation (4) as a single-valued "index of performance." However, if Equation 1 is correct that *MT* is a linear function of *ID*, then the intercept, *a*, would have to be zero or at least relatively small for this summary to be approximately constant. Although it might seem intuitively plausible for *MT* to approach zero as *ID* does and intercept values indistinguishable from zero have occasionally been observed, typically the intercept is found to be positive, and, in a few cases, including

Fitts and Peterson (1964, p. 107) negative intercepts have been reported. Because the zero intercept is outside the observable range of *ID* values,[5] its value is an extrapolation from the data. As with any extrapolation, estimates of the intercept should be interpreted cautiously. According to one widely cited interpretation of the intercept, it should be positive and reflect fixed perceptual and/or motor processes (e.g., target selection) that, although required, are not influenced by movement difficulty (Welford, 1968).

Whatever the source, as Zhai (2004) has elegantly pointed out, the presence of a non-zero intercept implies that *IP* must vary, sometimes substantially, across target conditions with different *ID*. This problem is hardly resolved by averaging across a set of target conditions as recommended by Soukoreff and MacKenzie (2004), because the resulting average must also necessarily depend on the distribution of the *ID* values of the conditions included in the average. This fact makes comparisons of *IP* across experiments problematic. As will be demonstrated in the Discussion, problems with *IP* constancy can arise in comparisons of conditions within an experiment.

Perhaps because of similar concerns, Fitts subsequently proposed using the inverse of the slope coefficient,

$$IP = {}^{1}\!/\!_{b}, \qquad (5)$$

as a measure of ''relatively constant information capacity over a range of movement conditions'' (Fitts and Radford, 1966, p. 476). Given the information theory approach that motivated Fitts's work in this area, this definition makes perfect sense. Equation 5 also has the advantage over Equation (4) that, for any range of ID values over which Equation (1) holds, the expected value of this estimate will be constant.[6] The disadvantage of this approach is that focusing solely on the inverse slope coefficient and ignoring a non-zero intercept discards important information. From an applied perspective, what typically matters is not a theoretical construct such as information capacity but rather the expected time required to complete pointing operations with different levels of difficulty and, as Equation (1) states, that time depends on both the

---

[5] Using Equation (3), the smallest plausible value of ID is 0.585. However, this describes conditions, in which $D = W/2$: i.e., the target region extends all the way back to the starting point.

[6] Soukoreff and MacKenzie (2004, p. 775) display an equation for 1/b that is related to an intermediate step in the standard derivation of the slope estimator for linear regression. They assert, incorrectly, that the form of this equation supports their claim that estimates of the inverse slope are sensitive to the values of the independent variable, in this case of *ID,* associated with the observations that are included in the estimate. However, if the function relating *MT* and *ID* is linear, as Equation (1) holds, then, unlike *IP*, the expected value of the inverse slope coefficient does not depend on the particular *ID* values observed. *ID* appears in the equation displayed by Soukoreff and MacKenzie to normalize the slope estimate. In effect this sets the scale, or units (e.g., bits per s), of the inverse slope coefficient. If the scale of the ID values were to change, for example if the logarithm were taken with a base of 10 rather than 2, this would simply change the units (i.e., to ms per digit).

intercept and the slope. Thus, we agree with Zhai (2004) that both coefficients should always be considered when characterizing perceptual-motor systems (Zhai, 2004).

Many reports using Fitts's law have included both the intercept and slope parameters. In fact, later in the paper cited above (Fitts & Radford, 1966, p. 480), Fitts talks about using both to characterize the human motor system. Also, the first paper to apply the Fitts Law approach to HCI, Card, English, and Burr (1978) also report the full equation for each device. However, a disturbing number of subsequent papers in applied areas (see Zhai, 2004, p. 795, for one listing) have based their assessments on either only Equation (4) or Equation (5).

Even more troubling is that Annex B to ISO 9241-9, which describe procedures for testing the efficiency and effectiveness of input devices, states that the goal of testing should be to "provide a measure of throughput" (ISO, 2000, p. 28) and goes on to define throughput as ID_e/MT (p. 30). Soukoreff and MacKenzie (2004) have provided useful elaborations and extensions of the procedural recommendations in ISO 0241-9. Their seven recommendations include detailed instructions for obtaining data and fitting it using the version of Equation (1) based on ID_e. These instructions implicitly acknowledge a role for both coefficients. However, their seventh recommendation is more in line with the position taken by ISO 9241-9. This recommendation is to be applied when the purpose of an analysis is to compare two or more conditions. Such comparisons are to be based on a variant of the *IP* measure based on ID_e, which they call throughput and label TP. They assert that the advantage of this approach is that "Calculated this way, TP is a complete measure encompassing both the speed and accuracy of the movement performance" (p. 760).

Although we can understand the appeal of being able to characterize and compare different operators, conditions, or devices using a one-dimensional metric, the inconvenient truth of Equation 1 is that this is not generally possible. The one special case in which this approach works is when the difference between the conditions being compared is effectively confined to a difference in intercepts. When the slopes differ, even if the intercepts are zero, the size of the *MT* difference between the conditions will depend on *ID*; when the functions cross, the sign of the difference will change. Any single-measure approach to comparing conditions ignores these differences.

As discussed in the previous section, Guiard (1997) found that the Fitts's law lines for discrete and continuous movements cross. Thus, whether the *MT* of discrete or continuous movements will be less depends on the movement difficulty. This is a subtlety that is completely lost if the comparison is based on only *IP* or the inverse slope. Because we expected similar issues in the data from our experiment, we will ignore the recommendation to base comparisons on either of these measures and, instead report both coefficients from the fit of Equation 1. We believe that this should become the general practice.

1.5. Issue 3: Use of ID_e and the Accuracy Adjustment

As outlined in the Background section, using the effective index of difficulty, ID_e, rather than *ID*, the value computed from the nominal conditions, makes sense both in terms of what participants actually do and because this step typically improves the fit of Equation (1). Given this, it is not surprising that both ISO 9241-9 (ISO 2002, p. 30) and Soukoreff and MacKenzie (2004, p. 755-757, Recommendation IV) endorse this accuracy adjustment.

While not questioning that the accuracy adjustment provides a better description of movement time, we believe that basing design decisions on accuracy-adjusted data is a mistake. The problem is simple: in design situations what is typically known are the nominal conditions not the effective conditions. It seems almost like common sense that it is a mistake to make predictions using a model based on ID_e as the input parameter when all that is known are *ID* values. If there were a way to predict how operators respond to the nominal target conditions to produce the effective target conditions, then, of course, one would wish to use the more accurate model of Fitts's law based on ID_e to predict movement times. However, to the best of our knowledge such a model does not exist. Given this concern, we decided to report the estimated coefficients for Equation 1 based on both *ID* and ID_e, since each is best suited for different purposes.

2. METHODS

2.1. Participants

There were 13 participants (7 males); all had vision correctible to 20/20 or better and were right-handed. One female participant dropped out after the second session, due to scheduling conflicts. Each of the remaining participants took part in three, one-hour sessions. Subjects were paid $10/hr. The protocol for this experiment was approved by the UCI Institutional Review Board.

2.2. Apparatus

A PC running a custom program written in MATLAB was used to present stimuli and record responses. Stimuli were presented on a 17" computer monitor running at a 60 Hz refresh rate with a resolution of 1280 x 1024 pixels. The screen was calibrated so that 1 pixel extended 0.25 mm in both the horizontal and vertical dimensions. Participants used a Logitech optical mouse (Model #M-98C) to make responses. All movement acceleration software was disabled so that a mouse movement of 1 mm produced a constant cursor movement of 5 mm (20 pixels).

2.3.Design

There were three factors all manipulated within subjects: the target condition factor and two running condition factors. The target condition factor had ten levels. As shown in Table 1, these were constructed from six levels each of *D* and *W* producing nine unique *ID* levels. The two running condition factors each had two levels producing four conditions.

Table 1. Nominal target conditions.

Target Distance, *D* (pixels)	Target Width, W (pixels)	Index of Difficulty, *ID* (bits)
50	9	1.92
100	10	2.58
200	14	3.03
200	10	3.46
100	4	3.75
400	16	3.75
800	15	4.79
600	9	5.10
400	4	5.67
600	4	6.25

The experiment was organized into blocks of 22 movements. The first two movements in each block were not included in the analyses. Each hour-long session consisted of 40 blocks, organized into four sets of ten blocks. One of the four running conditions was done in each of these four sets with their order balanced across groups of four participants using a different digram-balanced Latin Square for each of the three groups.

The first of the two running conditions determined whether the target conditions were varied or blocked. How this was implemented is explained further in the Procedures. In the varied condition, target conditions occurred in a pseudo-random order generated with two constraints. First, each of the ten target conditions occurred twice in movements 3 through 22 of each block; the target conditions for the first two movements were not included in this constraint. Second, a target condition could not be selected if it would result in a target closer than 100 pixels from the left or right edge of the screen. Across ten blocks each target condition occurred 20 times, ignoring the first two trials in each block. When target conditions were blocked, each of the ten target conditions defined all of the movements in one block. Across ten blocks each condition occurred once, producing data from 20 trials, ignoring the first two trials in each block. The order of the target conditions was approximately balanced across participants.

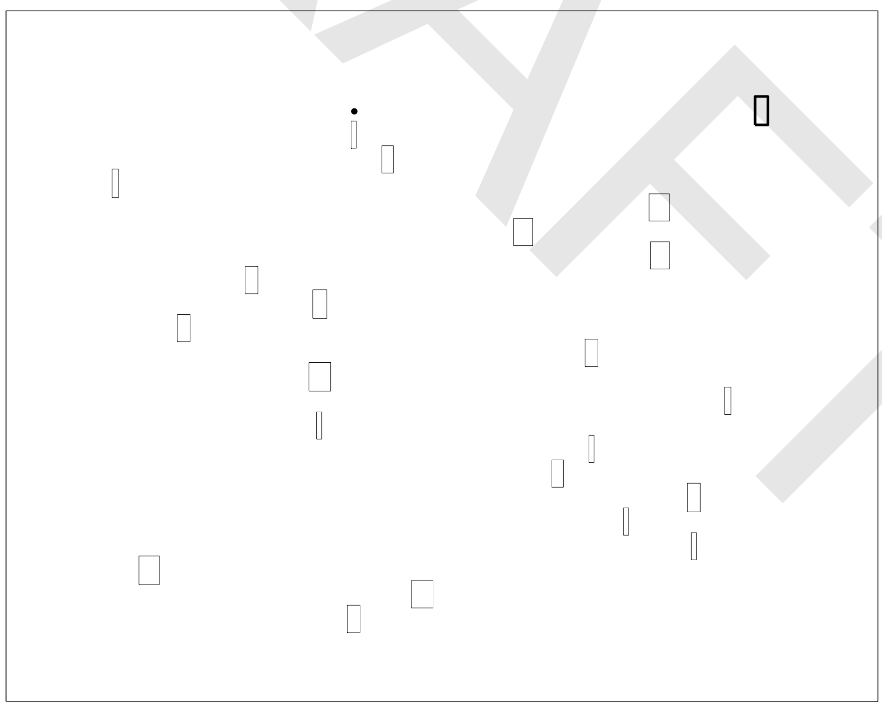
The second of the running condition factors determined whether the movements in a block were to be produced as a continuous sequence or discretely. How this manipulation was implemented is explained in the Procedures.

2.4.Procedures

Before each block a displayed message provided the participant a description of the running condition (blocked versus varied; discrete versus continuous). At the end of each block, a display told the participant the mean movement time and number of targets missed. The experimenter compared these with previous values in similar conditions and verbally encouraged the participant to move quickly while minimizing errors.

Figure 1 is a scale reproduction of an example stimulus display at the start of a block with varied target conditions. The small cross (5 pixels across) was the cursor and moved as the mouse moved. The dot was the starting point 6 pixels in diameter); it was displayed as a filled red circle when the cursor was too far from the starting point and a filled green circle when the cursor was within 3 pixels of the starting point. Once the cursor had remained in the target region for 0.5 s, a tone was presented. The tone indicated that the participant was free to start the first movement at any time. The onset of the tone also began the latency period which ended as soon as the mouse had been moved 5 pixels away from its initial position. Having moved the mouse to bring the cursor within the target rectangle, the participant pushed a mouse button indicating that the movement was complete. If the movement ended outside of the target rectangle, the error was noted but no immediate feedback was given.

Figure 1. Reproduction of a possible stimulus display at the start of a varied target condition block.



The 22 rectangles were the targets for the block. The target for the next movement was always highlighted by being displayed with a white outline on the gray background; all of the other targets were displayed with black outlines. In Figure 1, the highlighted target is drawn with a heavier line. The vertical center of the next target was always displayed at the same vertical position. This was possible because, as soon as the movement to a target has been completed, the box for that target disappeared from the screen, the vertical position of all of the remaining targets were moved up the screen, and the next target, which was now at the same vertical position as the previous target was highlighted.

Figure 1 is an example of the stimulus display when running the condition in which the target conditions were varied from one movement to the next. Note that the horizontal width of each rectangle, the target width, *W*, varied from target to target, but the vertical size is constant (35 pixels). The first movement was always to the right. In this case, it had a large target distance, *D*, and a moderate *W*. Each succeeding movement was always in the opposite direction. In this case, the second movement, to the left, had the same *D* and a smaller *W*, and the third movement, to the right, had a small *D*, and a somewhat larger *W*. When the target conditions were blocked, all of the rectangles had the same *W* and, ignoring the alternating directions, were the same *D* apart. Thus the target rectangles appeared in two vertical columns.

In the continuous running condition, the participant was free to start each successive movement as soon as the mouse button had been clicked to end the previous movement. The button click began the timing of the latency period that ended as soon as the mouse had been moved 5 pixels from the position recorded at the button click. In the discrete condition, as soon as the mouse button had been clicked a new starting point circle was displayed at the center of what had just been the movement target. At this point the participant had to move the cursor to the starting point and wait for the go signal just as with the first movement in the block.

3. RESULTS

3.1. Latency

For discrete movements, the latency was the time from the GO signal until the start of the movement was detected. For all but the first movement in a continuous movement block, there was no GO signal. For these blocks, what was recorded as the latency was the time between the mouse button press ending the previous movement and the detection of movement in the opposite direction, in essence the dwell time. Given these procedural differences, the mean latency for the discrete movements (525 ± 42 ms[7]) is not directly comparable to that for the continuous movements (72 ± 12 ms). There was no difference between the varied and blocked conditions [$t(11) = 0.589$]. Of potentially more interest is whether the *ID* of the upcoming movement influenced the latency. It did, but only in the discrete-varied condition where the slope relating latency to *ID* was 6.0 ± 5.7 ms per bit.

[7] This notation, $X \pm Y$, gives a mean value, the X, followed by the half width of the 95% confidence interval for that value, the Y.

The other three slope estimates were 1.6 ± 3.8 ms per bit for the discrete-blocked condition, -0.4 ± 1.5 ms per bit for the continuous-blocked condition, and -0.1 ± 1.5 ms per bit for the continuous-varied condition.

3.2. Practice Effects

Each participant produced data in each condition on each of three days. Not surprisingly, performance improved with practice when it was assessed either as the mean *MT* or as the slope relating *MT* to *ID*. The improvement was larger and statistically significant between Day 1 and Day 2 [the decrease in *MT* was 55.5 ms, *MSE* = 3078, $F(1, 11) = 47.889$, $p = .000$; the decrease in slope was 11.0 ms per bit, *MSE* = 461.1, $F(1, 11) = 12.530$, $p = .005$] and smaller and not quite significant between Days 2 and 3 [the decrease in *MT* was 13.5 ms, *MSE* = 2850, $F(1, 11) = 3.162$, $p = .103$; the decrease in slope was 4.2 ms per bit, *MSE* = 225.3, $F(1, 11) = 3.792$, $p = .077$]. Potentially more important is that the *MT* in the discrete condition was longer than that in the continuous condition for all three days: 98 ms, 69 ms, and 57 ms, respectively. The difference between Day 1 and Day 2 was almost statistically significant [*MSE* = 2589, $F(1, 11) = 4.004$, $p = .071$]; that between Days 2 and 3 did not approach significance [*MSE* = 1377, $F(1, 11) = 1.222$, $p = .293$]. The mean slope was smaller in the discrete condition for all three days: 2.8 ms per bit, 13.2 ms per bit, and 19.9 ms per bit, respectively. The difference between Day 1 and second encounter was statistically significant [*MSE* = 117.8, $F(1, 11) = 11.159$, $p = .007$]; that between Days 2 and 3 did not approach significance [*MSE* = 480.9, $F(1, 11) = 1.122$, $p = .312$]. Based on this pattern, the data from Day 1 were excluded from analyses that follow, and the data were collapsed across the second and third encounters within each condition. Including the Day 1 data does not qualitatively change any of the results reported below, but it does reduce the precision of some of the comparisons.

3.3. Movement Time versus Index of Difficulty

Figure 2 displays data averaged across participants showing the relationship between *MT* and *ID* for the four conditions of this experiment. Straight lines provide good fits to both the mean data (R^2 varied between .963 and .987) and to the data for each participant. Linear functions were fit separately to the data from each participant in each of the four running conditions. The resulting coefficients are summarized in Table 2. Fitts's law is typically parameterized using the slope and zero intercept. However, because these parameters must necessarily be highly correlated given the range of *ID* values, we prefer to also report and focus on the mean *MT* computed here at *ID* = 4, close to the average *ID*.

MT was larger for discrete than for continuous movements [$t(11) = 3.036$, $p = .011$]. There was not a reliable difference between the varied and blocked Conditions [$t(11) = 1.692$, $p = .119$]. However, both of these main effects were modified somewhat by their interaction [$t(11) = 2.832$, $p = .016$. As expected, the effect of this interaction was that the difference between varied and blocked sequences was larger in the continuous condition [$\Delta = 16 \pm 13$ ms, $t(11) = 2.688$, $p = 0.021$ (Bonferroni corrected $\alpha = .025$)], than in the discrete condition [$\Delta = 5 \pm 16$ ms, $t(11) = 0.755$]. The slope of the linear function relating

Figure 2. MT averaged over participants as a function of ID for each of the four conditions.

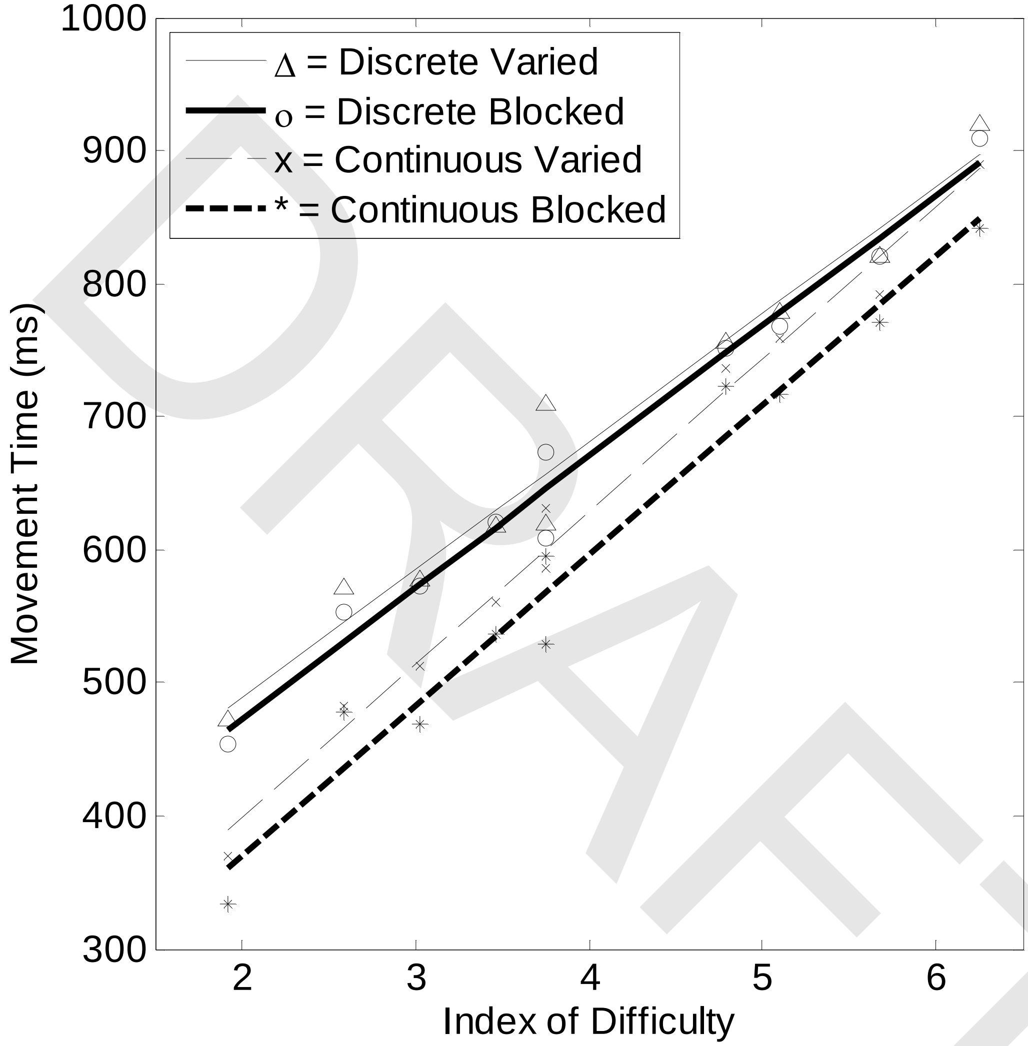


Table 2. Mean coefficients, averaged over participants, and the half width of the 95% confidence interval for the linear fit of MT versus ID for each of the four running conditions.

Mean MT (ms) at ID = 4

	Continuous	Discrete	Mean	Varied - Blocked
Varied	629 ± 42	681 ± 66	655 ± 50	
Blocked	596 ± 57	670 ± 84	633 ± 69	22 ± 29
Mean	612 ± 49	675 ± 74	644 ± 58	
Continuous - Discrete	-63 ± 45			Interaction 11 ± 8

MT versus ID Linear Slope (ms / bit)

	Continuous	Discrete	Mean	Varied - Blocked
Varied	114 ± 12	96 ± 11	105 ± 10	
Blocked	113 ± 12	99 ± 18	106 ± 14	-0.2 ± 6
Mean	114 ± 12	97 ± 14	106 ± 12	
Continuous - Discrete	17 ± 8			Interaction 2 ± 7

MT versus ID Zero Intercept (ms)

	Continuous	Discrete	Mean	Varied - Blocked
Varied	169 ± 44	297 ± 56	233 ± 38	
Blocked	144 ± 41	276 ± 61	210 ± 46	23 ± 19
Mean	157 ± 40	286 ± 56	221 ± 41	
Continuous - Discrete	-130 ± 53			Interaction 2 ± 25

MT to ID was larger for the continuous than for the discrete movements [t(11) = 4.607, p = .001]. For the slopes there was neither a reliable difference between the varied and blocked conditions [t(11) = 0.097] nor an interaction [t(11) = 0.676]. The effect of this combination of differences in the means and the slopes is that the MT of the continuous movements was smaller than that of the discrete movements; however, this difference was reduced as ID increased, and, for the continuous-varied condition, the advantage over the discrete conditions was eliminated at the highest IDs.

3.4. Missed-Target Errors

A simple interpretation of the *MT* versus *ID* relation is only possible if the proportion of missed target errors did not vary systematically across conditions. Although errors occurred on only 1.7% ± 1.2% of trials, this percentage did depend on both the running condition and the *W*. Errors occurred more often [$t(11) = 2.719$, $p = .020$] for continuous movements (3.0% ± 2.3%) than for discrete movements (0.4% ± 0.3%).[8] There was not a difference between the varied and blocked conditions [$t(11) = 0.859$] nor was there an interaction of these factors [$t(11) = 0.081$]. The percentage of errors also decreased significantly with increasing *W* [slope = -0.13 ± 0.11, $t(11) = -2.604$, $p = .025$] and this decrease was larger [$t(11) = 2.322$, $p = .027$] for continuous movements (-0.20 ± 0.18) than for discrete movements (-0.05 ± 0.04). This decrease did not vary reliably between the varied and blocked conditions [$t(11) = 1.682$, $p = .120$] nor was there an interaction of these factors [$t(11) = -0.542$].

3.5. Effective Target Width

The presence of systematic differences in the rate of missed target errors suggests that participants may have responded to changes in *W* differently across conditions. For this experiment, the standard deviation of the horizontal endpoint positions, sd_x, was used as the basis of the W_e estimate. A multiple regression of sd_x against *D* and *W* was performed for the data of each participant in each of the four running conditions. These three-parameter fits described the variation in sd_x well: the median of the R^2 for these fits was .86; the first and third quartiles were .82 and .88, respectively. We chose this model over one that simply used *ID* as a predictor, because for that model the median of $R^2 = .11$. The mean of sd_x, evaluated at $D = 300$ and $W = 20$ (i.e., $ID = 4$) was 5.4 pixel ± 0.6 pixel. sd_x increased slightly with *D* in three of the four running conditions as indicated by significant, positive slope estimates: discrete-varied, 0.082 ± 0.055 pixels per 100 pixels; discrete-blocked, 0.098 ± 0.064 pixels per 100 pixels; continuous-varied, -0.074 ± 0.074 pixels per 100 pixels; continuous-blocked, 0.084 ± 0.075 pixels per 100 pixels. Also, as would be expected, sd_x increased with *W* in all of running conditions [$t(11) = 18.566$, $p =$

[8] These values are representative of all but one participant whose error rate for the continuous movements was 12.9% but only 1.2% for the discrete movements. However, excluding this subject did not change the qualitative description of these data. The overall error rate dropped to 1.2% ± 0.7% of trials. Errors still occurred more often [$t(10) = 3.368$, $p = .007$] for continuous movements (2.1% ± 1.3%) than for discrete movements (0.4% ± 0.3%).

.000]. However, the slope describing this relationship was larger in the continuous movement conditions than in the discrete movement conditions [$t(11) = 40.321$, $p = .000$]: for discrete movements, the slope was 0.16 ± 0.04 pixels per pixel; for the continuous movements, it was 0.25 ± 0.02 pixels per pixel.

3.6. Effective Movement Distance

Just as it is dangerous to assume that participants respond to changes in *W* in a simple, consistent way, it is also possible that movement distances produced might not be simply related to *D*. To assess this possibility, a multiple regression of the average movement distance within each block versus *D* and *W* was performed for the data of each participant for each of the four running conditions. There was, however, almost no evidence for deviations of this sort. The intercept for this regression, which, to avoid extrapolation, was estimated at $D = 300$ and $W = 20$ was 299.96 ± 0.23 pixels, which does not differ reliably from the expected value. This intercept did not vary reliably across the four running conditions. The slope relating the actual distance to *D* was 1.01 ± 0.04 pixels per 100 pixels. It did not vary reliably across the running conditions. Similarly, the slope relating the actual distance to *W* was 0.006 ± 0.015 pixels per pixel, and this also did not vary reliably across the running conditions.

3.7. Movement Time versus Effective Index of Difficulty

Table 3 summarizes the coefficients for the linear relationship of *MT* and the effective index of difficulty, ID_e, which was computed as the $\log_2(D_e / W_e)$ where D_e is the mean of the measured movement amplitudes within a condition and $W_e = 4.133\ sd_x$. The constant in this equation is the one specified by ISO 9241-9 (ISO, 2000, Annex B, p. 29). Because this analysis produces different values of ID_e for each participant in each of the four running conditions, it is no longer possible to average data across participants as in Figure 2. However, straight lines summarize the data from individual participants well (across participants and the four running conditions, the median $R^2 = .86$). Compared with the fits based on *ID* in Table 2, these fits generally had smaller confidence intervals suggesting that using the estimated D_e and W_e, rather than their nominal values, produced results that were more consistent across participants.

The most striking change in the analysis based on ID_e is that the difference in mean *MT* between discrete and continuous movements is much smaller and no longer statistically reliable [$t(11) = 1.112$, $p = .290$]. The difference between the varied and blocked conditions is also smaller in this analysis; however, because of the increased precision this difference is reliable [$t(11) = 2.581$, $p = .026$]. The results in this analysis are also more straightforward because the interaction is smaller and no longer statistically reliable [$t(11) = 1.227$, $p = .245$]. As in the analysis based on *ID*, the slope of the linear function relating *MT* to *ID* was larger for the continuous than for the discrete movements [$t(11) = 3.198$, $p = .008$]. And again there was neither a reliable difference between the varied and blocked conditions [$t(11) = 1.585$, $p = .141$] nor an interaction [$t(11) = 0.102$].

Table 3. Mean coefficients, averaged over participants, and the half width of the 95% confidence interval for the linear fit of MT versus IDE for each of the four running conditions.

Mean MT (ms) at ID = 4

	Continuous	Discrete	Mean	Varied - Blocked
Varied	686 ± 34	698 ± 61	692 ± 47	24 ± 20
Blocked	655 ± 49	681 ± 74	668 ± 60	
Mean	671 ± 43	689 ± 66	680 ± 53	
Continuous - Discrete	-19 ± 37			Interaction 7 ± 13

MT versus ID Linear Slope (ms / bit)

	Continuous	Discrete	Mean	Varied - Blocked
Varied	120 ± 11	108 ± 13	114 ± 10	-4 ± 5
Blocked	124 ± 11	111 ± 15	118 ± 12	
Mean	122 ± 10	110 ± 13	116 ± 11	
Continuous - Discrete	13 ± 9			Interaction 0 ± 9

MT versus ID Zero Intercept (ms)

	Continuous	Discrete	Mean	Varied - Blocked
Varied	205 ± 33	265 ± 42	235 ± 27	38 ± 19
Blocked	158 ± 26	236 ± 54	197 ± 33	
Mean	181 ± 26	250 ± 44	216 ± 29	
Continuous - Discrete	-69 ± 43			Interaction 9 ± 30

3.8. Peak Velocity

The trajectories of movements in the Fitts task are usually observed to consist of an initial submovement, which ends close to or within the target, followed, if necessary, by one or more feedback-guided, corrective submovements (Meyer, Abrams, Kornblum, Wright, and Smith, 1988; Meyer, Smith, Kornblum, Abrams, Wright, 1990). One additional way in which the movements might differ systematically across the four running conditions is the speed of the initial submovement. Although a single submovement may vary in duration and average velocity, the velocity profile typically has a single peak and a fixed shape; this shape is multiplicatively scaled to produce movements with different average velocities and scaled in time to produce movements with different durations (Freund & Büdingen, 1978; Gordon & Ghez, 1987). Because of this invariance in the velocity profile shape, the peak velocity of a movement, which can be determined easily and reliably from movement trajectory data, is highly correlated with the average velocity of the initial submovement. Taking advantage of this regularity, a multiple regression of peak velocity against A and *W* was performed for the data of each participant in each of the four running conditions. The mean peak velocity was 3243 ± 519 pixels/s. Mean peak velocity did not depend reliably on the difference between discrete and continuous movements [$t(11) = 1.096$, $p = .296$], varied versus fixed conditions [$t(11) = 0.767$], or their interaction [$t(11) = 0.234$]. Peak velocity did increase with *D*. The slope of the linear fit was 9.1 ± 1.6 pixels per pixel; however, this slope did not differ across the four running conditions (all *p*-values < 1.0). Peak velocity also was not found to depend reliably on *W*.

4. DISCUSSION

The purpose of this study was to determine whether and, if so, how the coefficients characterizing the Fitts's law relationship for the mouse depend on how the task is organized. The primary factor explored was whether variation in distance and target width was blocked, as is often true in published studies, or whether the target conditions varied from movement to movement. This experiment also explored the effects of whether movements in a block were produced as separate, discrete movements or one, continuous movement sequence. This second factor has been shown previously to have effects on the Fitts's law slope; it was included here because we expected that its effects might interact with the possible effects of the first factor. Large differences due to the blocking factor would undermine the use for applied, design decisions of results obtained with blocked target conditions. In general, one might expect there to be differences across the four running conditions derived from these two factors based on two intuitively plausible ideas: that immediate experience with a specific movement will improve performance when that movement is repeated and that performance will be better on a movement when there is plenty of time to plan it.

4.1. Movement Time versus Index of Difficulty

The analysis for these data that does not include an accuracy adjustment is summarized in Figure 2 and Table 2. Because it is based on *ID* rather than ID_e, this is not the analysis that produces the best fit or the most easily interpreted results, nor is it the analysis recommended by either ISO 9241-9 (ISO 2002) or Soukoreff and MacKenzie (2004). However, as we have argued in the Introduction, the results of this analysis are important when predicting movement time in design applications where what would be known are the nominal rather than the effective target conditions.

As in previous research (Fitts & Peterson, 1964; Guiard, 1997; Megaw, 1975, Expt. 2) the slope relating *MT* and *ID* was larger (16%) for continuous movements than for discrete movements. There was, however, no slope difference comparing varied and blocked movement sequences.

The intercept of the discrete movements were 130 ms larger (59%) than that of the continuous movements. Because the zero intercept is an extrapolation to a condition (*ID* = 0) that is of no practical interest and the zero-intercept estimates can be strongly influenced by error in the estimate of the slope, our preference is to compare the mean *MT* estimated at an *ID* value in the middle of the range of practically interesting *ID* value studied: in this case *ID* = 4. However, even using this measure there is still a significant, 63 ms (9%) difference.

From the perspective of ANOVA, the slope difference between the discrete and the continuous movements amounts to an interaction between *ID* and this factor. As Figure 2 illustrates, continuous movements have lower MTs for the full range of *ID* values studied; however, the advantage of the continuous movements is reduced considerably for larger values of *ID*. Another way to characterize such an interaction is to look at simple-main effects. Within this framework, the significant result cited above for mean *MT*, can be considered a simple main effect for *ID* = 4. There is also a significant difference for *ID* = 2 ($\Delta = 96 \pm 47$ ms, $t(11) = 4.511$, $p = 0.001$), but not for *ID* = 6 ($\Delta = 30 \pm 50$ ms, $t(11) = 1.328$, $p = 0.211$), although the difference is still in the same direction.

Because there was no slope difference between the varied and blocked sequences, the mean *MT* and zero-intercept estimates were quite similar: the blocked sequences were about 22 ms faster than the varied sequences, a difference that is statistically reliable for the zero intercept but not for the mean *MT*. The direction of this difference is consistent with the expectation that repeating the same movement should lead to improved performance.

4.2. Issues Interpreting Movement Time versus Index of Difficulty

As would be expected, latency, the time to initiate a movement, was substantially longer for discrete than for continuous movements. Interestingly, there was not an overall latency difference between varied and blocked sequences. What might be a concern is that latency was found to depend on *ID* for discrete movements in varied sequences. Although this effect was only just barely reliable and relatively small 6 ms/bit (5% of the mean slope), it is worth considering because it is plausible that planning time, reflected in the latency, could be traded off for *MT*.

A larger and more direct concern is that a straightforward interpretation of the *MT* versus *ID* relationship only makes sense if error rates do not differ systematically across conditions. Unfortunately, although the average error rate was low in this experiment, 1.7%, it did vary across the running conditions: there was a substantially smaller error rate for discrete movements. The was also a decrease in error rate as *W* got larger, and, what is more troubling, this decrease was substantially larger for continuous movements than for discrete movements. These results are troubling because they are consistent with a speed-accuracy tradeoff: when *MT* was smaller, error rate was larger.

Consistent with the changes in error rate were the changes in sd_x and thus W_e, the effective target width. ISO 9241-9 (ISO, 2000, Annex B, p. 29) specifies that W_e be calculated using a formula that was originally derived from information theory (see Welford, 1968, pp. 147-148): $W_e = 4.133\ sd_x$. The multiplier here is equivalent to assuming that the movement endpoints have a Gaussian distribution and that the error rate is about 4%. A Gaussian endpoint distribution has often been observed for similar data (Woodworth, 1899; Crossman, 1960; Fitts & Radford, 1966) and is a reasonable description here. Ideally, one might expect W_e to be proportional to *W* with a slope of 1. However, applying this formula, suggests that the W_e versus *W* slope is about 0.66 for the discrete movements, which is significantly less than 1, and 1.03 for the continuous movements. Both of these numbers would be a little higher if the multiplier in the equation for W_e were increased to 4.773 to reflect the 1.7% average error rate observed here instead of the 4% rate implied by the formula. However, looking past the specific slope estimates, what is important here is that participants responded more strongly to variation in *W* when making continuous movements than when making discrete movements. MacKenzie and Isokoski (2008) obtained similar variations in the response to variations in *W* using instructions intended to induce different speed-accuracy tradeoff points.

There was also a systematic increase in sd_x associated with increasing *D*. However, even in the condition in which this slope was largest, the discrete-blocked condition, the associated changes in W_e were small from a practical perspective. For movements with $D = 50$, the expected reduction in W_e from the mean would be about 1 pixel, while for the movements with $D = 800$, the expected increase would be 2 pixels.

4.3. Movement Time versus Effective Index of Difficulty

Substituting W_e for *W* and thus using ID_e to predict *MT*, produces results that are more consistent with the expectations outlined in the Introduction. Because this approach leads to more consistency between participants, the intercept difference and the difference in mean movement time between varied and blocked sequences are statistically reliable: introducing variation of target characteristics increased *MT* by about 3.5%. At the same time, this analysis reduced the overall mean difference between continuous and discrete movements to a level that is no longer close to being reliable. The slope relating *MT* to ID_e for continuous movements is, however, still larger (11%) than that for discrete movements. As for the analysis based on *ID,* this slope difference can be interpreted as an interaction. For $ID = 2$, discrete movements are significantly slower ($\Delta = 44 \pm 36$ ms, $t(11) = 2.665$, $p = 0.022$). As shown in Table 3, the difference at

ID = 4 (19 ms) is no longer reliable. For *ID* = 6, the effect is in the other direction, but again it is not statistically reliable (Δ = -6 ± 21 ms, $t(11)$ = -0.313).

If it is correct that the difference between discrete and continuous movements has its primary effect on the slope, but varying target characteristics influences only mean movement time, then this suggests that these factors influence different processes. This is a theoretically interesting idea. An important goal for future research would be to identify where in the movement process these differences are taking place. As discussed in the previous section, continuous and discrete movements differ in the way that they reflect changes in target width. However, an analysis based on ID_e is intended to compensate for this difference, and yet the effects of both factors are still present. The stochastic optimized-submovement model (Meyer et al., 1988, 1990) suggests that another obvious place to look for the source of these differences is the velocity of the initial submovement. However, assuming that differences in initial submovement velocity are reflected in peak velocity, there was no evidence in these data for systematic differences across the running conditions.

4.4. Practical Implications

The Introduction raised three issues that are sources of potential concern for the way that results related to Fitts's law are applied in HCI. The first of these is whether there is an important effect of varying target conditions between trials rather than blocking them. A priori this is an issue because blocking target conditions simplifies data collection, but in most applied situations target conditions vary from movement to movement. However, for mouse movements such as those studied here, the 24 ms additive increment in *MT* switching from blocked to varied target conditions is sufficiently small relative to the overall *MT* (3.5%) that, for most applications it might be seen as inconsequential. Another perspective from which to view this difference is that across the 12 participants in this study, the mean *MT* values, assessed at *ID* = 4, ranged from 583 ms to 806 ms. Given the restricted population (college students) included in this study, the upper end of this range is undoubtedly very small compared with that of a less restricted population. Yet the effect of blocking is only 10% of the variation across participants observed here. Because the small size of this effect was unexpected, it would be prudent to replicate it, both with the mouse and with other devices before taking it too seriously.

The remaining two issues concern how data such as these should be analyzed and reported. The first of these involved the recommendation that *IP* be used to summarize results and compare conditions or devices. For a comparison of the varied versus blocked target conditions, this strategy works fine because there is little if any slope difference between these conditions. However, this approach is at best questionable for the comparison of the continuous versus discrete conditions. If *IP* is computed following the directions provided by Soukoreff and MacKenzie (2004), the means in each of the four running conditions are 5.76 ± 0.48 bits/ms in the discrete-varied condition, 6.02 ± 0.64 bits/ms in the discrete-blocked condition, 5.58 ± 0.31 bits/ms in the continuous-varied condition, and 5.95 ± 0.40 bits/ms in the continuous-blocked condition. However, these averages collapse across significant *IP* differences that depend systematically on *ID* ranging from 4.1 bits/ms to 7.9 bits/ms; almost a 2 to 1 ratio. A three-factor ANOVA, suggests that the effects of *ID* interact significantly with those of the two running conditions factors. These results make us uncomfortable with the recommendation that these comparisons be based on *IP* because it can obscure important differences in the results. Reporting only *IP* also strikes us as problematic, because for many applied design situations it is predicted movement time, not throughput, that is of interest.

We believe that a better approach is one that does not force a one-dimensional simplification of what are inherently at least two-dimensional data. The approach we have illustrated here involves both reporting the two coefficients from Equation 1 and, where there are slope differences, reporting simple main effects at representative levels of *ID*.

Comparing the results here with those previously reported comparing discrete versus continuous movements reveals a potentially interesting pattern. Using a mouse, we find a higher slope for continuous movements, but a lower intercept. The result is that the functions relating *MT* to ID_e crossed at an *ID* of 5.5. Guiard (1997) also studied movements with a mouse-like device and also found that the functions crossed, in that case with an *ID* of 4.2. In contrast, although both studies reporting reciprocal tapping movements also found the slope for discrete movements to be larger than that for continuous movements, these studies obtained results in which discrete movements were substantially faster than continuous movements across the full range of *ID* values sampled: 2 bits to 7.6 bits in Fitts & Peterson (1964), 4.3 bits to 7.1 bits in Megaw (1975). Although this may reflect a fundamental difference between tapping and mouse movements, it could also reflect measurement differences associated with apparatus. In both tapping studies, a metal stylus was used to make movements that ended by hitting metal targets. This may have reduced measured dwell time, increasing the measured *MT*.

The comparison of the discrete and continuous conditions may shed light on the final issue: when it is reasonable to base comparisons on analyses using *ID* versus ID_e that was also raised in the Introduction. The two analyses we have reported agree that the slope is higher for discrete movement than for continuous movements. However, the two analyses differ substantially in their predictions about the size of the MT difference across levels of difficulty. If one followed the recommendation to use the predictions of Equation 1 obtained using ID_e and then substituted *ID* values obtained from a to-be-evaluated design, this would result in prediction errors of up to 50 ms. Although the fit of Equation (1) is better when ID_e values are used, in this applied situation, more accurate predictions will be obtained using the fit of Equation 1 and *ID* values. In general, we believe that estimates from both fits should be reported.

4.5. Conclusions

This paper has explored three issues related to the application of research using Fitts's law in HCI. The importance of these issues has been underscored by the recent, important efforts to standardize that research. We have argued along with Zhai (2004) that ISO 9241-9 (ISO, 2000) needs to be amended to eliminate the recommendation that any one-dimensional measure of performance be used to evaluate devices or user interfaces. As Equation (1) states, *MT* in these tasks reflects two dimensions of variation and any attempt to collapse this onto one dimension will be misleading. We have also argued that, for some applications, fits of Equation (1) based on *ID* are more appropriate than those based on ID_e. Finally, we have demonstrated that although movement-to-movement variation in target conditions results in longer *MT*s, this effect is much smaller than might have been anticipated. This is important both because it raises interesting theoretical issues and because it suggests that the large body of data collected with blocked target conditions may still be usefully applied to situations that commonly involve varied target conditions.

NOTES

Authors' Present Addresses. 2554 Social & Behavioral Sciences, University of California, Irvine, CA, 92697-5100

HCI Editorial Record. (supplied by Editor)

REFERENCES

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 on a CRT. *Ergonomics, 21*(8), 601-613. doi:10.1080/00140137808931762

Card, S.K., Moran, T.P., Newell, A. (1983). *The Psychology of Human–computer Interaction*. Lawrence Erlbaum, Hillsdale, NJ.

Crossman, E.R.F.W. (1960). The information-capacity of the human motor-system in pursuit tracking. *Quarterly Journal of Experimental Psychology, 12*, 1–16.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), 381-391. doi:10.1037/h0055392

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology, 67*(2), 103-112. doi:10.1037/h0045689

Fitts, P. M., & Radford, B. K. (1966). Information capacity of discrete motor responses under different cognitive sets. *Journal of Experimental Psychology, 71*(4), 475-482. doi:10.1037/h0022970

Freund, H. J. & Büdingen, H. J. (1978). The relationship between speed and amplitude of the fastest voluntary contractions of human arm muscles. *Experimental Brain Research, 31*, 1-12.

Gordon, J. & Ghez, C. (1987). Trajectory control in targeted force pulses. II. Pulse height control. *Experimental Brain Research, 67*, 241-252.

Guiard, Y. (1993). On Fitts and Hooke's laws: Simple harmonic movement in upper-limb cyclical aiming. *Acta Psychologica, 82*, 139-159.

Guiard, Y. (1997). Fitts' Law in the discrete vs. cyclical paradigm. *Human Movement Science, 16*, 97–131.

Guiard, Y., & Beaudouin-Lafon, M. (2004). Fitts' law 50 years later: Applications and contributions from human-computer interaction. *International Journal of Human-Computer Studies*, 61(6), 747-750. doi:10.1016/j.ijhcs.2004.09.003

ISO, 2000. ISO 9241-9 International standard: Ergonomic requirements for office work with visual display terminals (VDTs)—Part 9: Requirements for non-keyboard input devices: International Organization for Standardization.

Kvålseth, T. O. (1980). An alternative to Fitts' law. *Bulletin of the Psychonomic Society, 16*, 171-373.

MacKenzie, I. S. (1992). Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human-Computer Interaction, 7,* 91-139.

MacKenzie, I. S., Isokoski, P. (2008). Fitts' Throughput and the Speed-Accuracy Tradeoff. *CHI 2008 Proceedings,* 1633-1636.

Megaw, E. D. (1975). Fitts tapping revisited. *Journal of Human Movement Studies, 1*(4), 163-171.

Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., and Smith, J. E. K. (1988). Optimality in Human Motor Performance: Ideal Control of Rapid Aimed Movements. *Psychological Review, 95*, 340-370.

Meyer, D. E., Smith, J. E. K., Kornblum, S., Abrams, R. A., Wright, C. E. (1990) Speed-accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. In M. Jeannerod (Ed.), *Attention and Performance XIII: Motor Representation and Control* (Chapter 6, pp. 173-226).Hillsdale, NJ: Lawrence Erlbaum.

Pastel, R. (2011). Positioning graphical objects on computer screens: A three-phase model. *Human Factors, 53*(1), 22-37. doi:10.1177/0018720810397353

Plamondon, R., Alimi, A.M. (1997). Speed/accuracy trade-offs in target-directed movements. *Behavioural and Brain Sciences, 20*, 279–349.

Welford, A.T. (1960). The measurement of sensory-motor performance: survey and reappraisal of twelve years' progress. *Ergonomics, 3*, 189–230.

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

Soukoreff, R. W., & MacKenzie, I. S. (2004). Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in HCI. International Journal of Human-Computer Studies, 61(6), 751-789. doi:10.1016/j.ijhcs.2004.09.001

Woodworth, R. S. (1899). The accuracy of voluntary movements. *Psychological Review, 3* (2, Whole No. 13).

Zhai, S. (2004). Characterizing computer input with Fitts' law parameters--the information and non-information aspects of pointing. *International Journal of Human-Computer Studies, 61*(6), 791-809. doi:10.1016/j.ijhcs.2004.09.006