Fitts' Law as an Explicit Time/Error Trade-Off

Yves Guiard **Halla B. Olafsdottir** **Simon T. Perrault**
Laboratoire de traitement et de communication de l'information, CNRS, Telecom ParisTech
46 rue Barrault, 75013 Paris, France
{yves.guiard; halla; simon.perrault}@telecom-paristech.fr

ABSTRACT
The widely-held view that Fitts' law expresses a speed/accuracy trade-off is presumably correct, but it is vague. We outline a simple resource-allocation theory of Fitts' law in which movement time and error trade for each other. The theory accounts quite accurately for the data of Fitts' (1954) seminal study, as well as some fresh data of our own. In both data sets we found the time/error trade-off to obey a power law. Our data, which we could analyze more thoroughly than Fitts', are consistent with a square-root function with a single adjustable constant. We suggest that the resource-allocation framework should help combine information and energy considerations to allow a more complete account of Fitts' law.

Author Keywords
Fitts' law, resource allocation, speed/accuracy trade-off.

ACM Classification Keywords
H.5.2. User Interfaces: Evaluation/methodology, theory and method.

General Terms
Experimentation, Measurement, Performance, theory.

1. INTRODUCTION

This paper is about Fitts' law, the well-known relation that links the time it takes people to reach a target (e.g., a graphical object) with a pointer and the accuracy of their reaching action. We introduce a new formulation of Fitts' law which specifies one sense in which the law can be said to be a speed/accuracy trade-off, as traditionally assumed in HCI [12,13] as well as psychology [15,18]. A trade-off is a mutual dependency between two utilities that conflict with each other because they both draw on the same limited-resource pool: the better the performance on one front, the worse it is on the other [16,17]. Understanding how this concept applies in the context of Fitts tasks is our goal here.

A Fitts' law equation is an empirical regularity that relates mean movement time μ_T to an index of difficulty *ID* computed as a simple mathematical transform of *D/W*, the ratio of target distance *D* to target width *W*. Here are, among many others [18], four well-known formulations of the law:

$\mu_T = a * \log_2(2D/W) + b$ Fitts (1954) [2] (1)

$\mu_T = a * \log_2(D/W) + b$ Crossman (1956) [1] (2)

$\mu_T = a * \log_2(D/W + 1) + b$ MacKenzie (1992) [12] (3)

$\mu_T = a * (D/W)^b$. Meyer et al. (1990) [15] (4)

μ_T denotes average movement time (technically a mean or possibly a median, at any rate a central-trend statistic), and *a* and *b* stand for adjustable coefficients (a>0). Most popular within HCI is Eq. 3, known as the Shannon version of Fitts' law [11,12]. The starting point of this analysis is that Eqs. 1-4 do *not* describe a speed/accuracy trade-off.

2. THE BASIC MEASURES: TIME AND ERROR

2.1. Time Is Not Speed

First, μ_T, the dependent variable that stands on the left-hand side of Eqs. 1-4, is a time measure. In general μ_T correlates negatively with the average speed of a movement. Nevertheless it is only in casual language that one can tolerate confusion between a time measure, dimensionally [T], and a speed measure, dimensionally [LT^{-1}] [9].

2.2. Accuracy: Neither Information Nor Difficulty

Second, how the quotient of *D*/*W*, which determines the *ID* on the right-hand side of Eqs. 1-4, measures accuracy is unclear. In light of information theory [22], Fitts [2] assumed that the information conveyed by a movement is $\log_2(2D/W)$, a formula which MacKenzie [11,12] corrected to $\log_2(D/W+1)$. The information and the accuracy of movements must be linked somehow, but to our knowledge that link has not been clearly described.

Assuming that the mathematical transforms of *D/W* that feature in Eqs. 1-3 provide estimates of movement *difficulty* rather than movement information does not take us any closer to a measure of movement accuracy. In the Shannonian Fitts-MacKenzie tradition, difficulty is measured in bits and calculated, via Eqs. 1-3, from an objective property of the target layout, namely the ratio of lengths *D* and *W*. But this is just information—for lack of an operational definition of its own, it is hard to see how task difficulty might relate to accuracy.

If one wished to characterize difficulty as *subjective* effort [19], one would have the problem that none of the above *ID*s bear a monotonic relationship with this effort. There is no question that in the upper region of the *ID* spectrum (over 4 bits or so, using the Shannon *ID*), the higher the *ID*, the more difficult the task for participants. But in the lower

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

region of the *ID* spectrum (below 2 bits or so), the *lower* the *ID*, the more difficult the task: no matter their good will, participants will systematically fail to produce large enough spreads of movement endpoints [1,4,12]. This is not too surprising. Since the kinetic-energy cost of movements varies with the square of their velocity, a Fitts task becomes difficult not only with very high but also (in a different sense of the word) with very low *ID*s. The participants' failure to comply with instructions in low-*ID* conditions just reflects their reluctance, or mere inability to produce fast enough movements because of their excessive *energetic* cost. This observation tends to be overlooked in an approach exclusively based on Shannonian information. But from the moment it is recognized that an aimed movement involves not only bits of information, but also joules of energy, it becomes clear that movement difficulty, characterized as subjective effort, can only bear a U-shaped relation with the variable known as the *ID* in Fitts' law research [7]. Information, as captured by any *ID* estimate, cannot be taken as an index of subjective difficulty.

Thus a typical Fitts' law equation expresses a relation, not between movement speed and movement accuracy, but rather between movement time and a certain dimensionless ratio whose relation with both accuracy and difficulty is unclear. We now present some distinctions which we think useful to rephrase Fitts' law as an explicit trade-off.

2.3. Relative Target Distance *D/W* vs. Relative Target Tolerance *W/D*

When Fitts [2] (p. 266) introduced his index of difficulty, he wrote $ID = -\log_2(W/2D)$, rather than $ID = \log_2(2D/W)$. These are two different mathematical writings of the same thing, and so whether the independent variable of Eqs. 1-4 is *D/W* or *W/D* might be judged an idle question. In fact that distinction is quite critical because the quotients of these two divisions designate different measures in the physical world of experimentation. The quotient of *D*/*W* is a measure of *relative target distance* (*RTD*)—i.e., *D* scaled to, or expressed in units of *W*. In contrast, the quotient of *W*/*D* is a measure of *relative target tolerance* (*RTT*)—i.e., target tolerance scaled to, or expressed in units of *D*.[1] Although it has been a tradition to formulate Fitts' law as an equation of the form μ_T=f(*D/W*), there is reason to prefer the inverse writing μ_T=f(*W/D*) [8]. First, there is a *scale of measurement* issue [26]: relative target distance or *D/W* lacks a true zero because the limiting case where *D*=0 and *W*>0, hence *D/W*=0, violates the very definition of a Fitts task—if *D*=0, no movement is required [8]. In contrast, relative target tolerance or *W/D* does enjoy a true zero: the limiting case where *W*=0 and *D*>0, hence *W/D*=0, corresponds to a zero-tolerance aiming task, which makes sense and has indeed been investigated [21]. Thus only *RTT*=*W/D*, and not *RTD*=*D/W*, runs on a *ratio* (equal-interval) scale of measurement [26]. This matters because a higher level of measurement for experimental variables means a more constraining framework for testing theoretical hypotheses [20]. For example, the *y*-intercept of an empirical regression line is interpretable only if the *x* variable has a physically-anchored zero [8].

RTT is also preferable over *RTD* for the statement of Fitts' law because any measure of accuracy, whether absolute or relative, should involve error as a component. It seems much more sensible to ground one's characterization of accuracy on a measure of tolerance (i.e., permitted error) like *RTT* than a measure of distance like *RTD*.

2.4. Task Geometry vs. Movement Performance

This section calls attention to an obvious distinction that has received little attention in the literature. On the one hand experimenters have full control over *D* and *W*, two systematic, deterministic variables that characterize the geometrical layout of targets and that serve to prescribe to participants an average amplitude and a spread of movement endpoints, respectively. On the other hand one needs to characterize the participants' actual performance. Here the elemental measures are movement duration *T* and movement amplitude *A*, from which an endpoint error can be computed as $E=A-D$. Unlike *D* and *W*, variables *T* and *A* (as well as *E*) are random variables, reflecting the natural variability of human performance. We need to distinguish *T*, *A* and *E*, to be measured at the level of individual movements, from central-trend statistics like means μ_T, μ_A, and μ_E, to be calculated over samples of movements.

We deliberately wrote Eqs. 1-4 above as μ_T=f(<u>*D*</u>/*W*) rather than μ_T=f(<u>*A*</u>/*W*) as has been customary since Fitts [2], because the latter notation is somewhat wobbly. If *W* unambiguously designates a property of the target layout (tolerance), it is always unclear whether the conventional symbol *A* stands for *D* (thus referring to the target layout) or μ_A (thus referring to the movement). The writing of Fitts' law becomes particularly ambiguous in this regard when *W* is replaced by *effective* width W_e to denote the tolerance that, in retrospect, would have yielded a pre-specified error percentage, given a certain spread of endpoints. Labeling this variable as "target width" suggests one is talking task geometry while W_e is a random variable of the movement.

In fact the accuracy issue can be approached in Fitts' paradigm from two markedly different, though equally legitimate, angles. In one approach, Fitts' law is all about the dependency of μ_T upon the dimensionless ratio *W/D* (or its inverse *D/W*), as suggested in the formulations we chose for Eqs. 1-4. In this approach μ_T is predicted from the task geometry, and the problem of accuracy must be phrased in terms of *D/W* or *W/D*. In the alternative approach, Fitts' law is all about the mutual dependency of two random variables, movement time and relative variable error. *RVE* seems well represented by σ_A/μ_A, a regular coefficient of variation in which σ_A and μ_A denote the standard deviation and the mean of movement amplitude [8]. Thus Fitts' law can be formulated either as μ_T=f(*W/D*), expressing the causal dependency of a temporal random variable upon a

---

[1] This study requiring a number of non-conventional distinctions and notations, we appended a glossary to the paper (Section 9).

systematically-varied geometrical variable, or alternatively as $\mu_T=f(\sigma_A/\mu_A)$, expressing the mutual dependency of two random variables. These are what we call the *geometrical* vs. the *stochastic* version of Fitts' law.

HCI researchers, who often need to evaluate or predict pointing performance for certain target layouts, naturally adopt the former approach, assuming that movement performance is causally dependent on the target layout. It is the alternative approach, however, that paves the way for a trade-off analysis. If one wants to formulate Fitts' law as a trade-off, one needs to write the law in the form of a *mutual* dependency, with movement time depending on movement error and vice versa—it should not matter whether Fitts' law is written $\mu_T=f(\sigma_A/\mu_A)$ or, reciprocally, $\sigma_A/\mu_A = g(\mu_T)$.

3. A SIMPLE RESOURCE-ALLOCATION THEORY OF FITTS' LAW

Below are listed a set of basic assumptions needed for a resource-allocation theory of Fitts' law. Note that the trade-off under consideration is not between speed and accuracy, but, strictly speaking, between mean or median movement time μ_T and relative variable error *RVE*.

1. *Utility*. Movement time and relative variable error are both *negative* utilities, that is, quantities that must be minimized—the shorter the μ_T, the better the performance; the smaller the *RVE*, the better the performance.

2. *Trade-Off*. The two minimization efforts conflict with each other: the less of one negative utility, the more of the other. This is a trade-off of the min-min category.[2]

3. *Limited Resource Pool*. The trade-off results from the fact that the two concurrent minimization efforts draw from a *common pool of resources*, and this pool is *limited*. This assumption is the counterpart, within the trade-off theoretical approach, of Fitts' [2] limited-capacity channel assumption. We may designate the content of the hypothetical pool, whose nature is unknown, as the *effort*. We just need to assume, using the usual economic analogy, that some generic currency is convertible into speed and/or accuracy and that the available amount of this currency is finite, being a characteristic of every individual placed in a given situation. Devising a method for estimating that amount is our first important challenge here.

4. *Less-than-Total Resource Exploitation*. In Fitts' law experiments participants are instructed to constantly do their best —i.e., to invest 100% of their resources. Human effort, however, is subject to random fluctuations and so the amount of resource actually available to an individual at a given point in time can be less—but never more—than these 100%. The limited resource pool, in other words, must be thought of as an *upper bound*. We believe this realistic assumption, which has escaped researchers' attention until recently [25], is mandatory in any approach to Fitts' law, including the information theoretic approach.

5. *Resource allocation strategy*. Faced by resource scarcity in a Fitts task, participants can deliberately modulate the balance between their concurrent time-minimization and error-minimization efforts. Quantifying that imbalance, estimating its range of variation, and understanding its dependency upon systematically manipulated experimental conditions—different target layouts in Fitts' [2] experiment (Section 4), different verbal instructions in ours (Section 5)—constitute the second challenge of this analysis.

4. FITTS' (1954) TAPPING DATA: EVIDENCE OF A TIME/ERROR TRADE-OFF

This section aims to show that Fitts' data can indeed be reformulated explicitly as a trade-off between two conflicting utilities. Focusing on the min-min trade-off of movement time μ_T and relative variable error $RVE = \sigma_A/\mu_A$, we will introduce a simple geometrical method for characterizing quantitatively the size of the resource pool as well as the strategic imbalance.

At first sight, the suitability of Fitts' experimental protocol for a trade-off analysis of his data might seem questionable. Fitts did not ask his participants to minimize movement time and relative error concurrently—he asked them to minimize a single variable, μ_T, under a number of different constraints of relative tolerance, and so it was a systematic factor that stood for accuracy. One should bear in mind that with such a protocol, still most popular today, error actually remains a negative utility (i.e., the less of it, the better), just like movement time. It is in order to obtain from participants differing levels of *RVE* that experimenters display differing levels of *RTT*. The manipulation of target display as in [2] and instructions as in [4] may be viewed as two alternative methods with the same goal. In the former option participants are to minimize μ_T with a variety of *RVE* constraints, while in the latter they are to jointly minimize μ_T and *RVE* under a variety of speed/accuracy compromises. But both methods boil down to instructions, formulated visually and in words respectively, serving to manipulate the participants' cognitive stance in the face of the fundamental speed/accuracy dilemma.

We will consider the data Fitts [2] obtained in his famous reciprocal tapping experiment, tabulated in his Table 1 (p. 264, *light*-stylus data). The table reports movement times averaged over 16 participants, for each of 16 combination of *D* and *W*. However, Fitts did not actually record the position of movement endpoints, just tabulating percentages of target misses. Capitalizing on Fitts' report (p. 265) that undershoot and overshoot aiming errors were about equally frequent in his light-stylus experiment, we assumed $\mu_A=D$. We inferred endpoint spreads from error rates using the technique described by MacKenzie [10] (Section 2.5). For each combination of *D* and *W* we computed effective width W_e (for a fixed 4% error-rate constraint, under the hypothesis of a Gaussian spread of endpoints) and then calculated $\sigma_A=W_e/4.133$.

---

[2] An example of a *max-max* trade-off is that between speed and accuracy, both positive utilities: the faster and the more accurate the movement, the better the performance.

Note that our analyses below separate the different levels of scale, characterized by D or μ_A, following the recommendation of Guiard [6]. We assume that the form and the size of the target display are specified by $RTT=W/D$ and D, respectively, and that the form and the size of the aimed movements are specified by RVE and μ_A, respectively.

4.1. A Power Relationship Between Movement Time and Relative Variable Error

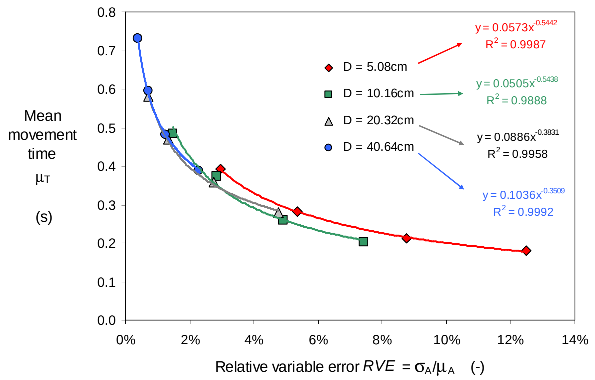

Figure 1. The trade-off of μ_T and RVE in Fitts' tapping data.

As shown in Fig. 1, Fitts' data is closely modeled, for each scale level, as a power function ($.989<r^2<.999$):

$$\mu_T = q * RVE^{\,p} \qquad (5)$$

where p and q represent adjustable coefficients ($p<0$, $q>0$). Note that the logarithmic fit ($r^2=.993$ on average over the four scales) was nearly as good as the power fit ($r^2=.996$).

4.2. Amount of Resources

Eq. 5 may be rewritten as

$$\mu_T * RVE^{-p} = q \qquad (6)$$

or, since we define relative variable error RVE as σ_A/μ_A, as

$$\mu_T * (\sigma_A/\mu_A)^{-p} = q. \qquad (7)$$

Eq. 7 is the statement of a constant product: within each scale condition, the product of μ_T and RVE raised to the power $-p$ was conserved in Fitts' experiment despite his systematic change of the target layout and consequently of μ_T. The conservation of quantity q is illustrated in Fig. 2. For each of the four scale conditions the slope of the regression line is virtually zero—as movement time varied over a range of about 2:1, q remained remarkably stable.

In light of the trade-off theory outlined in Section 3, it is clear that the constant q specifies the average amount of resources that was available to Fitts' participants. Note that the constant q is indicative, not of an amount of resources, but of resource *scarcity*: the smaller the product of the two negative utilities, the better the performance.

The different elevations of the four flat curves of Fig. 2 show that the amount of resources available to Fitts' participants was scale dependent. The constant q reached a minimum in the D=10.16cm condition, presumably a scale optimum for Fitts' particular task conditions.

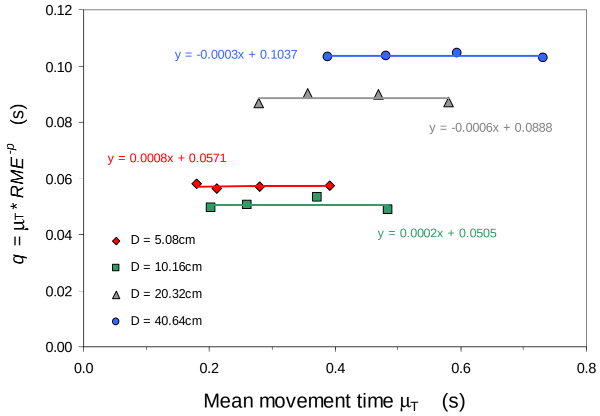

Figure 2. Conservation of the product q of Eq. 7 across the variation of μ_T, for each scale condition.

Fig. 3 plots Eq. 7 for D=40.64cm, whose best fit is $\mu_T=0.1036/RVE^{0.3509}$ ($r^2=.9992$, see Fig. 1). Notice that the rectangle obtained by drawing straight horizontal and vertical lines to the axes from any point of the curve, whether chosen within the actual range of x values like points A, B, C, and D, or extrapolated along the curve like point E, has a constant area (if $y=q/x$, then $xy=q$). This area is no other than the coefficient q of Eq. 7, whose estimate in that particular scale condition is 0.1036s.

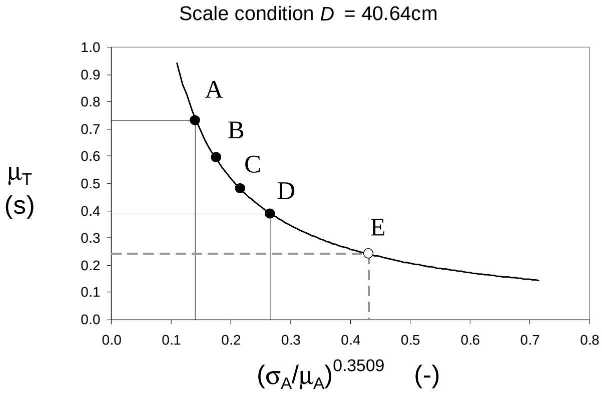

Figure 3. A plot of Eq. 7 for the D=40.64cm scale condition, where exponent p is -0.3509. ABCD are Fitts' actual four data points, E is an arbitrary extrapolation along the curve. All rectangular areas are equal.

4.3. Resource Allocation: Strategic Imbalance

If different points along the curve of Fig. 3 correspond to one and the same amount of resources, they specify different degrees of imbalance between the time- and error-minimization effort. While the product xy (the rectangular surface area in the figure) is conserved all along the curve, reflecting available resources, the ratio y/x (the rectangle's aspect ratio) changes gradually, reflecting different resource-allocation options. For any data point of the curve the actual *strategic imbalance* (SI) of participants can be quantitatively characterized by this aspect ratio, that is,

$$SI = \mu_T \,/\, RVE^{-p}, \qquad (8)$$

or, recalling that relative variable error $RVE = \sigma_A/\mu_A$,

$$SI = \mu_T / (\sigma_A/\mu_A)^{-p}. \qquad (9)$$

With this definition of the aspect ratio, which we chose to compute as y/x, the strategic imbalance *SI* decreases in Fig. 3 from left to right: the less cautious (and the faster) the movement, the lower the index. Thus *SI* correlates positively with—is an index of—the relative strength of the error-minimization component of the participant's effort.

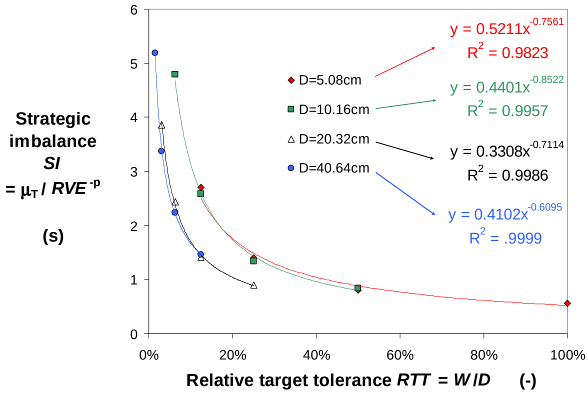

Figure 4. The *SI* as a function of *RTT* in Fitts' data.

Fig. 4 shows the dependency of the actual strategy of Fitts' participants, upon *RTT*, the characteristic of the target layout that Fitts manipulated as an attempt to control his participants' strategy. This dependency is highly non-linear, suggesting that the target-layout manipulation technique that Fitts introduced in his 1954 study actually provided him with mediocre control over the resource allocation strategy of his participants. That mediocrity is apparent in Fig. 5. Although Fitts raised *RTT* up to the point where his two targets touched each other (i.e., $W=D$ hence $W/D=100\%$), *RVE* hardly exceeded 10%.

Back to Fig. 4, notice that Fitts' participants had two different strategic stances in the face of four different movement scales. For $RTT<30\%$, the x ranges covered by the four curves substantially overlap, thus making it possible to compare *SI*s for geometrically-similar target layouts: Fitts' participants had about the same set of strategic imbalances in the two larger-scale conditions $D=20$cm and 40cm, but they apparently had another, more cautious set of strategies in the two smaller-scale conditions $D=5$cm and 10cm. The statistical reliability of this finding cannot be tested for lack of individual data but the pattern seems impressively consistent. One conjecture would be that because high speeds cannot be attained over small amplitudes, a scale reduction might have encouraged participants to adopt a relatively more cautious strategy.

While the pattern is quite conspicuous in the plot of Fig. 4, it is virtually undetectable in the classic plot of μ_T vs. *ID*. Also note that it could not have been deduced from our previous observation that the resource pool was maximal in Fitts' participants for 10cm movements (Fig. 2)—the aspect ratio (the quotient of μ_T/RVE^{-p}) and the surface area (the product $q=\mu_T*RVE^{-p}$) of the rectangles of Fig. 3 are two mathematically independent quantities.

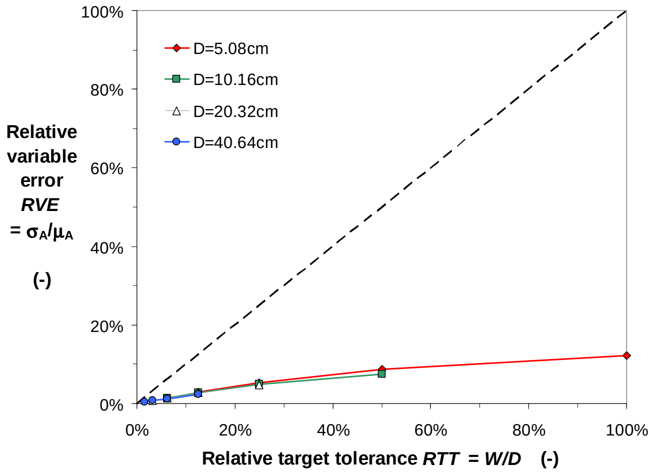

Figure 5. *RVE* as a function of *RTT*. The dashed line represents the theoretical equality *RVE=RTT*.

4.4. Discussion

The foregoing confirms that the classic tapping data of Fitts (1954) can be satisfactorily interpreted as a trade-off between two negative utilities, time and error. Thus, not only can we view Fitts' law as the demonstration that throughput—the inverse of the law's slope, whose dimensions are bits/s—is conserved as the task *ID* is made to vary, we can just as well view the law as evidence that a certain pool of effort resources is conserved in the people across the variation of strategic imbalance. Both the information theoretic approach and the trade-off approach may help us understand Fitts' law.

5. FRESH DATA ON THE TIME/ERROR TRADE-OFF

This section reports a simple experiment which only varied task instructions so as to induce a systematic variation of the participants' strategic imbalance in the face of the concurrent time- and error-minimization efforts. Movement amplitude was invariably a comfortable 150mm.

Among our motivations for running a fresh experiment was the fact that Fitts' [2] individual data are not available. From the standpoint of resource allocation theory, one expects some quantities—notably the coefficient q of Eq. 5—to behave as within-individual constants while at the same time varying from participant to participant. Also of considerable interest is the variability of the strategic imbalance among and within individuals. Whether experimenters manipulate the target layout, as has been customary since Fitts, or speed/accuracy instructions, they face human beings with idiosyncratic strategic styles. No two participants will identically interpret instructions to move, say, as fast as possible. Nor will they show the same degree of flexibility in response to changing instructions.

Our discussion of Fitts' data above did not refer to less-than-total resource exploitation, assumption #4 of our trade-off theory, whose illustration and testing require individual data. Below we will see that this assumption is quite useful to estimate individual trade-offs.

There were two notable differences between our protocol and Fitts'. First, our aiming task was *discrete*, rather than reciprocal, our participants having to return to a fixed home position after each aimed movement. This option makes it possible to clarify the status of our temporal and spatial measures. Whereas in the reciprocal protocol μ_T is the time it takes not only to carry out a movement, but also to evaluate the error inherited from the previous movement and to prepare the next [3], in the discrete protocol μ_T measures the duration of a pure movement-execution process. Furthermore the meaning of the movement's endpoint spread σ_A is interpretable more safely in the discrete case, that variability being generated just by the execution of the movement, whereas in the reciprocal case σ_A must also reflect, to some unknown extent, the variability of the start point [3]. Finally, most pointing actions in real world HCI are discrete.

The other notable difference is that we did not visually specify tolerance *W*. We just specified *D* with two lines indicating the start point and the desired endpoint of the movement, the target being thus displayed as a single line. We manipulated the balance between the two concurrent minimization efforts by means of different sets of instructions that asked the participants to cover their whole spectrum of imbalances, from maximum speed (minimizing μ_T) to maximum accuracy (minimizing *RVE*).

5.1. Method

Participants

Sixteen volunteers participated (all right-handed, median age 27.5years, interquartile range 2.5years, four female).

Speed/Accuracy Instructions

We used five sets of instructions: 1) Max speed, 2) speed emphasis, 3) speed/accuracy balance, 4) accuracy emphasis, and 5) max accuracy. In the max-speed condition the participants were to just minimize movement time, the only requirement regarding accuracy being to refrain from committing a systematic error: no matter the dispersion of movement endpoints, participants were just to manage to terminate their movements at about the target *on average*. At the opposite extreme, the max-accuracy instructions asked participants to try to bring the cursor exactly to the target (zero pixel error), making as many corrective sub-movements and taking as much time as needed—but not more. These two extremes being defined, we simply inserted three intermediate levels of instructions, one unbiased (speed/accuracy balance) and two biased (speed emphasis, accuracy emphasis).

Apparatus and Setup

The experiment involved a 1280x1024-pixel (34.0 x 27.1cm) screen and a Wacom Intuos3 digitizing tablet connected to a PC running Linux Ubuntu. The screen permanently displayed two vertical lines extending from top to bottom, located 150mm apart, which marked the start point (left) and the target (right) of the movement. Both lines were 1-pixel thick and appeared in red color over a white background. Also displayed was a mobile 1-pixel thick cross-hair cursor, black in color, whose motion was controlled by the stylus. The tablet being used in absolute mode with a control-display gain of 1, the hand had to move 150mm from its home position for the crosshair to reach the target line.

The participant was seated at a table supporting the Wacom tablet and the screen, with a viewing distance of about 50 cm. During warm-up trials the participants were allowed to optimize the orientation of the tablet in the horizontal plane to facilitate the execution of the required left-to-right movement, the tablet being often tilted counter-clockwise. On the tablet was secured a horizontal 8-mm thick plastic ruler, along which the stylus tip was to be slid, allowing a strictly one-dimensional hand movement. The ruler offered a mechanical stop at its left end so that the start position of the stylus was standardized to the nearest screen pixel. To help initial positioning, an OK message appeared on the screen when the crosshair coincided exactly with the start line.

We developed our own software, using Lib USB, for tablet-data acquisition, to minimize display latency relative to tablet events and to exploit the full resolution of our input device (5080dpi). The tablet coordinates were translated into pixels using floating values to maximize visual-feedback accuracy. The tablet sampling rate was approximately 100 Hz (in the range 85-125 Hz).

Procedure and Movement Measurement Algorithms

The experiment consisted of 25 blocks of 15 movements, each block being run with a given set of instructions. All five instructions were presented in one order, ascending or descending, the order being reversed from one group of five blocks to the next. The experiment lasted about 40mn per participant, including 10min of warm up.

To begin each trial the participant immobilized the screen crosshair at the start line by positioning the stylus on the tablet at the ruler stop for a few seconds. When ready, the participant moved the stylus to the target position by sliding it against the ruler, finishing up with a dwell, then lifted the stylus and, after a few seconds rest, proceeded to the next trial.

The movement start point corresponded to the place and instant where the crosshair left its home position while exhibiting positive (rightward) acceleration. Determination of the movement *endpoint* in time and space is a more subtle issue and detailed explanations about our offline algorithm were part of the instructions received by the participants. We used two different criteria, depending on the instructions condition. For the max-speed condition, where accuracy was irrelevant, the explanation was that our algorithm would take as the movement endpoint the *first* zeroing-out or zero-crossing of instantaneous velocity, thus ignoring any subsequent episode of velocity, whether deliberate (a corrective sub-movement) or accidental (e.g., a mechanical rebound due to the elasticity of the arm). For the other four conditions the movement endpoint was defined as the beginning of the last dwelling period in the kinematic record, meaning that the algorithm would take into account all corrective sub-movements, if any. The criterion for dwell was the crosshair remaining stationary for at least 100ms at least 50mm away from the home position.

5.2. Results and Discussion

Inspection of the distributions of movement times revealed some skewness, especially in the max-speed condition, hence our recourse below to median movement time (still noted μ_T) as our central-trend statistic, in place of the mean.

Systematic Error, Variable Error

On average over all participants the constant error (μ_A - 150mm) was less than 1 screen pixel in all instructions conditions but one: in the max-speed condition we found a significant 5.5mm overshoot error (t_{15}=4.50, p<.0005). Whatever the reason for this very small (3.7%) if consistent effect, it is quite safe to say that it was the relative *variable* error σ_A/μ_A, rather than the constant error μ_A-D, that was influenced, along with μ_T, by our variation of instructions.

Effect of Instructions on Movement Time and RVE.

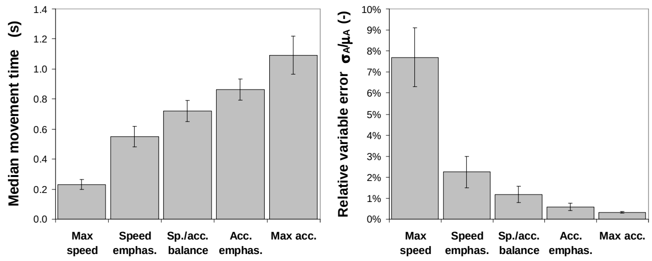

Figure 6. The effect of instructions manipulation on median μ_T and *RVE*, on average over our 16 participants. Error bars are 95% confidence limits based on between-participant *SD*s.

With increasing emphasis on accuracy, μ_T lengthened about linearly and *RVE* decreased non-linearly (Fig. 6). Although the max-speed instructions imposed no constraint whatsoever on the spread of movement endpoints, *RVE* hardly reached 0.08 (or 8%), a finding reminiscent of Fitts' data (see Fig. 5).

Convex Front of Performance

Fig. 7 illustrates the trade-off between μ_T and *RVE* for one representative participant. Panel A plots all the participant's data points, one per trial block. The best fit is a power function, with quite some noise, hence a moderately impressive r^2 of .87. But let us see how assumption #4 of our resource-allocation theory, less-than-total resource exploitation, helped us exploit our data more thoroughly.

The participants were to minimize μ_T and *RVE* in differing proportions. Their data points may be likened with particles attracted to the West and South by two magnetic fields whose relative strengths are modulated by instructions. Viewing the scatter as a mixture of forerunning and dawdling particles, we simply assumed that the forerunners and the dawdlers were the data points under and above the curve. If the resource pool is limited (Assumption #3) then forerunners must have been constrained by a hard wall—the very trade-off curve we are looking for, conceptually the borderline that separates in the graph the region of the *doable* (above the curve) from the region of the *undoable* (below). That hard wall must have prevented forerunners from spreading any further in the South-West direction. But there must be dawdlers (Assumption #4) affected by the hard-wall constraint to an attenuated extent. The data is quite consistent with this view. As shown in Fig. 7B, restricting the fit to the subset of forerunners improved the fit considerably (from .87 to .97 in this individual example).

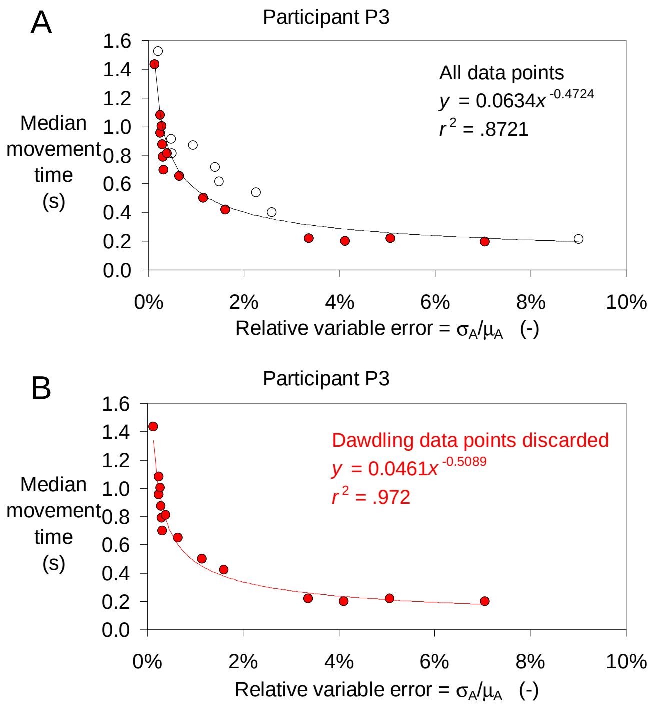

Figure 7. Convex front of performance. Panel A: a power curve is fitted to all the data points delivered by the participant, the data points under the curve (filled discs) are then selected. Panel B: a new power curve is fitted to the selected subset.

An improved fit of the power equation after elimination of dawdlers was observed in all 16 participants (Student t_{15}=8.83, p<.0001), the r^2 increasing on average from .852 to .972. We checked that, in contrast, restricting the fit to the data points resting above the first curve (i.e., selecting the dawdlers instead of the forerunners) did not improve the fit whatsoever, the r^2 changing on average over all participants from .852 to .853 (t_{15}=0.03, p=.97).

The data points that form the South-West quadrant of the scatter plot (shown as filled discs in Fig. 7A) are especially important: they characterize a participant's best performance. These data points must have had more than others their locations affected by resources limitations. In the rest of this report we will focus on this particular subset, which we call the *convex front* of performance, assumed to be the most informative about the location of the trade-off curve we are looking for.

Logarithmic, Exponential, and Power Fits

We tested the simple two-coefficient models that can accommodate the convex-down curvature evident in all our μ_T vs. *RVE* trade-off functions, the logarithmic, the exponential, and the power equation. Fitting the three candidate models to the full individual data sets (i.e., to the 25 pairs of measures originating from the full set of trial blocks) resulted in the power model doing best in 10 cases (mean r^2=.853 over all participants), the exponential model doing best in four cases (mean r^2=.803) and the log model in two (mean r^2=.803). In general, the log and (most blatantly) the exponential equation failed due to insufficient curvature. Fitting the three models again to the convex-front data, the power model turned out to provide the best fit for all but one of our 16 participants, namely Participant #11 (P11), with the r^2 now ranging between .923 and .992 (on average r^2=.972, to be compared with r^2=.937 for the log and r^2=.880 for the exponential model). Therefore we retained the power equation $\mu_T = q*RVE^p$ (i.e., Eq. 5) for modeling the trade-offs.[3]

A Closer Look at Individual Exponents: Evidence for a One-Coefficient SQRT Relation

Thus, a simple power relation turned out to describe quite accurately the time/error trade-off in both Fitts' data and our own, despite our different experimental protocol. Most interestingly, the exponent *p* of the best-fitting power model was similar, considering the two scales conditions of Fitts' study that approximately corresponded to ours. In Fitts' data the exponent was -0.54 for *D*=10cm and -0.38 for *D*=20cm (Fig. 1); in our data, with *D*=15cm, the exponent was -0.47 on average. We further inquired into this issue by asking how the exponent varied with the goodness of fit and with the value taken by *q*, the other adjustable coefficient of Eq. 5 (see Fig. 8).

To reiterate, any individual participant produces a mixture of good and poor performance (assumption #4), but there is an infinity of ways of performing poorly and in principle only the best (convex-front) performance of that participant is informative. It should be realized that as one switches from a within- to a between-individual logic this argument works just the same. Different individuals being unequally able (or willing) to fully concentrate on a repetitive movement task, there is every reason to focus on the data of the best performers in one's quest for a consistent quantitative law. Bearing this in mind, a remarkable suggestion emerges from the data: first, the better the fit of the power model, the closer the exponent to -1/2 (Fig. 8A); second, the smaller the value of the participant's coefficient *q* (i.e., the better the performer, as explained in Section 4.2), the closer the exponent to -1/2 (Fig. 8B). Three individual estimates of the exponent diverged appreciably from -1/2, namely those delivered by participants P6, P8, and P11 (*p*=-0.60, -0.34, and -0.32, respectively), but these happened to be the sample's least credible estimates: P6 and P8 were participants for whom the power fit was distinctively worse than average (see Fig. 8A); as for P11, we can see in Fig. 8B that he ranked 15th/16 for performance, next to P8, who ranked last—a further reason to moderately trust P8's data.

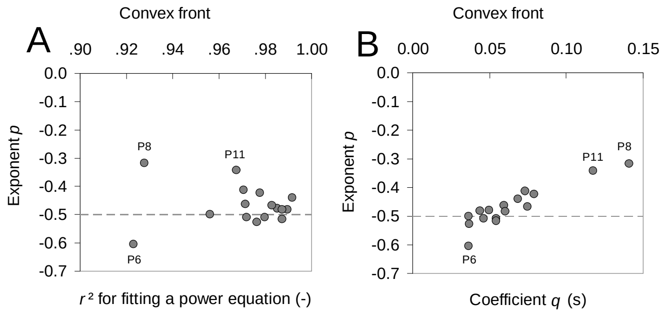

Figure 8. The exponent *p* of Eq. 5 plotted (A) against the r^2 of the power fit and (B) against the value of coefficient *q*. Each data point corresponds to an individual participant.

Thus, focusing on the *best performance* (i.e., the convex front of performance) of our *best performers* (actually 13 of our 16 participants), we found that the trade-off of μ_T and *RVE* can be satisfactorily modeled in our data by a square root equation with a single adjustable constant:

$$\mu_T = q * RVE^{-1/2} \quad \text{or}$$

$$\mu_T = q / \mathrm{SQRT}(\sigma_A/\mu_A) \quad (10)$$

where the multiplicative constant *q* is information about the amount of resources invested by the participant.

Constant Resource Pool and Variable Strategic Imbalance

An immediate implication of Eq. 10 is that the quantity $q = \mu_T*\mathrm{SQRT}(\sigma_A/\mu_A)$ is conserved as the participants modulate their speed/accuracy strategy (Section 4.2). But this strategic imbalance can be quantified quite simply, as the variable ratio $SI = \mu_T/\mathrm{SQRT}(\sigma_A/\mu_A)$ (see Section 4.3).

Fig. 9, which uses data from the same representative participant as Fig. 7, shows examples of the fit obtained with the one-coefficient model of Eq. 10 and of the within-individual conservation of *q* across the variation of μ_T. The orderly pattern of panel A was the rule. On average over our best 13 performers, Eq. 10 yielded r^2=.964 (.889< r < .986). This was significantly less than the r^2=.967 obtained with Eq. 5 (t_{12}=3.61, p=.002, one tailed), but the r^2 reduction (.003) was impressively small, given the sacrifice of one of the two free coefficients of Eq. 5.

[3] The μ_T vs. *RVE* relation involves two random variables neither of which is 'dependent' or 'independent'. In such a case the so-called standard major axis method of curve fitting is known to be preferable over traditional linear regression, which measures errors only along the vertical *y* axis [23]. Here both methods yielded nearly identical estimates (not surprisingly, given the very high correlations found in log-log plot between μ_T and *RVE* [23]), and so we did not depart from the ordinary least-square method.

Panel B displays a scatter plot showing no evidence of any correlation between q and μ_T. Despite a considerable 7-fold variation, μ_T failed to exert any consistent influence on coefficient q. Indeed, considering our 13 reliable data sets, the slope of this relation (-0.0034s on average) did not significantly depart from zero (t_{12}=-1.42, p=.182, two-tailed).

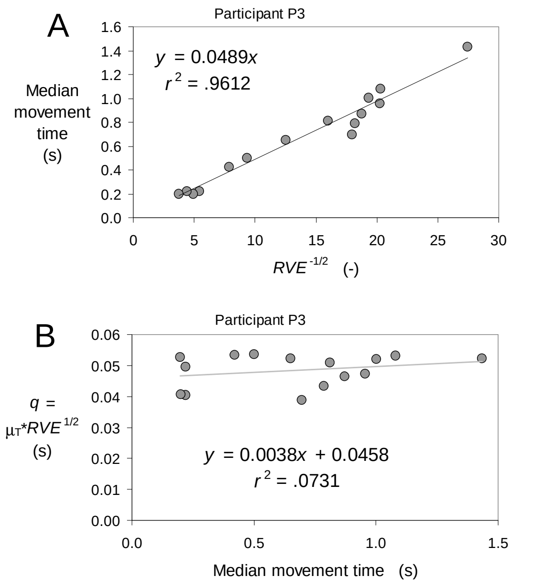

Figure 9. A representative example of (A) the fit of the one-free-coefficient model of Eq. 10 and (B) the approximate conservation of q across a seven-fold variation of μ_T.

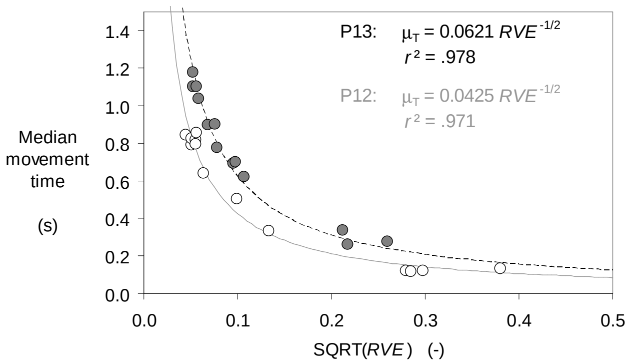

Figure 10. Comparing two individual trade-off curves.

Fig. 10 shows the trade-off curves of two individual participants whose q coefficients took distinctively different values. It is easy to see that P12 invested more resources in the task than did P13 (q=0.0425 and 0.0621, respectively). But the graph, by exhibiting different distributions of data points along their respective curves, also reveals that P12 and P13 had different *strategic preferences* for resource allocation. In response to our max-accuracy instructions P13 climbed his curve higher than P12 did his (SI_{max}=22.5 vs. 18.9), while in response to our max-speed instructions he did not explore his curve as far down as P12 did his (SI_{min}= 1.07 vs. 0.35). Thus the comparison delivers two separate pieces of evidence: beside the fact that P12 invested more resources than P13, he was more speed-biased.

IMPLICATIONS FOR HCI AND BASIC RESEARCH

In this paper we have tried to formulate with some rigor one specific sense in which Fitts' law is a speed/accuracy trade-off. The data show that indeed the law can be recast as a trade-off between two random variables, the time it takes to reach a target and the relative spread of movement endpoints. Note that the trade-off we have described involves time and *in*accuracy, rather than speed and accuracy. Whether Fitts' law can be recast, strictly speaking, in terms of speed and accuracy remains an open question, with some non-trivial difficulties given the multiple possible definitions of each of these two terms.

One potentially important outcome of this research is the suggestion that the mathematical description of Fitts' law can be simplified, without sacrificing much modeling precision, to a square-root equation with a single adjustable constant. For both basic-researchers and practitioners, the fewer the free coefficients of a model, the better. Based on our work currently in progress, our conjecture is that it is only for optimal scales of pointing movement that Fitts' law can be this much simplified.

The resource-allocation approach to Fitts' law helps understand that to obtain a complete understanding of target-acquisition performance we need both an intensive and a qualitative characterization. If the intensive aspect is explicitly addressed by the throughput, it seems that the information-theoretic framework has little to say about the qualitative aspect. The fact that the speed/accuracy balance is variable has been considered a worrisome complication calling for a certain correction—the substitution of effective to nominal width [1,12]—so as to end up with a single synthetic measure of performance. The correction being done, the throughput is quite insensitive to substantial variations of the speed/accuracy imbalance [4,13]. But the fact that the throughput (just like the coefficient q of our trade-off analysis) is conserved through *SI* variations does *not* mean that the latter are unimportant—a conclusion that a rapid reading of [13] might suggest. The cognitive stance controlled by speed/accuracy instructions, which strongly modulates the balance of movement times and endpoint spreads, is indeed an important factor, which does not influence throughput, but does influence another aspect of performance, the *SI*.

Reducing the data of a Fitts' law experiment to just a throughput measure, as recommended in [24], is indeed convenient for a Fitts' law practitioner but there is a cost. Suppose that, comparing two interaction techniques A and B, one finds more throughput with technique A in the presence of more errors. Assuming an appropriate adjustment for errors, the conclusion that A outperforms B may be correct but a full half of the story has been erased by the adjustment procedure. In some research contexts, especially those where safety matters critically, it may be very useful to not ignore that different interface arrangements or interaction techniques may induce different speed/accuracy imbalances in their users.

The resource allocation and the information theoretic approaches are certainly not incompatible with each other, as recognized in practice by Fitts and Radford [4], who (although they did not theorize on strategic imbalance) did manipulate speed/accuracy instructions. The right-hand side of Eq. 10 rewritten as $\log_2(\mu_T)$ less a constant = -½ $\log_2(\sigma_A/\mu_A)$ can be viewed to display bits of information. But the left-hand side of this equation, as already hinted above, is likely to call for a model involving joules of energy. Thus, one intriguing outcome of the present analysis is the suggestion that information theory should certainly participate in, but probably will not suffice to wholly account for Fitts' law.

6. ACKNOWLEGMENTS

This research has been supported by *Institut Telecom*, with a Carnot post-doctoral fellowship to HBO and a *Futur et ruptures* PhD grant to STP, and by Ubimedia, a joint Alcatel and *Institut Telecom* research lab. We warmly thank Eric Lecolinet, Olivier Rioul, Fabrice Rossi, Vincent Hayward, as well as anonymous reviewers for stimulating criticisms and suggestions.

7. REFERENCES

1. Crossman, E.R.F.W. (1956). The measurement of perceptual load. Unpublished PhD thesis, Univ. Birmingham, England.
2. Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psychol., 47, 381-391. Reprint: J. Exp. Psychol.: Gen., 1992, 121, 262-269.
3. Fitts, P.M. & Peterson, J.R. (1964). Information capacity of discrete motor responses. J. Exp. Psychol., 67, 103-112.
4. Fitts, P.M. & Radford, B.K. (1966). Information capacity of discrete motor responses under different cognitive sets. J. Exp. Psychol., 71, 475-482.
5. Guiard, Y. (2001). Disentangling relative from absolute movement amplitude in Fitts' law experiments. CHI'2001. New York: ACM Press, 315-316.
6. Guiard, Y. (2009). The problem of consistency in the design of Fitts' law experiments: Consider either target distance and width or movement form and scale. Proc. CHI'2009. New York: Sheridan Press, 1908−1918.
7. Guiard, Y. & Ferrand, T., 1998. Effets de gamme et optimum de difficulté spatiale dans une tâche de pointage de Fitts. Science et Motricité, 34, 19-25.
8. Guiard, Y. & Olafsdottir, H.B. (2010). What is a zero-difficulty movement? A scale of measurement issue in Fitts' law research. Submitted manuscript. http://hal.archives-ouvertes.fr/.
9. Hoffmann, E.R. (1996). The use of dimensional analysis in movement studies. J. Motor Behav., 28 (2), 113-123.
10. MacKenzie, I.S. (1991). Fitts' law as a performance model in human-computer interaction. PhD thesis available online: http://www.yorku.ca/mack/phd.html.
11. MacKenzie, I. S. (1989). A note on the information-theoretic basis for Fitts' law. J. Motor Behav., 21, 323-330.
12. MacKenzie, I.S. (1992). Fitts' law as a research and design tool in human-computer interaction. Human-Computer Interaction, 7, 91-139.
13. MacKenzie, I.S. & Isokoski, P. (2008). Fitts' throughput and the speed-accuracy tradeoff. Proc. CHI'2008. New York: Sheridan, 1633−1636.
14. Meehl, P.E. (1997). The problem is epistemology, not statistics. In L.L. Harlow et A. (Eds.), What if there were no significance tests? (pp. 393-425). Mahwah, NJ: Erlbaum. http://www.tc.umn.edu/~pemeehl/169ProblemIsEpistemology.pdf.
15. Meyer, D. E., Smith, J. E. K., Kornblum, S., Abrams, R. A., & Wright, C. E. (1990). Speed-accuracy trade-offs in aimed movements: Toward a theory of rapid voluntary action. In M. Jeannerod (Ed.), Attention and performance XIII (pp. 173-226). Hillsdale, NJ: Erlbaum.
16. Norman, D.A., & Bobrow, D.G. (1975). On data-limited and resource-limited processes. Cognit. Psychol., 7, 44-64.
17. Norman, D.A., & Bobrow, D.G. (1976). On the analysis of performance operating characteristics. Psychol. Rev., 83, 508-510.
18. Plamondon, R., & Alimi, A.M. (1997). Speed/accuracy trade-offs in target-directed movements. Behav. Brain Sci., 20, 279-349.
19. Rosenbaum, D. A. & Gregory, R. W. (2002). Development of a method for measuring moving-related effort: Biomechanical considerations and implications for Fitts' Law. Exp. Brain Res., 142, 365-373.
20. Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. Psychol. Rev., 107, 358-367.
21. Schmidt, R.A., Zelaznik, H.N., Hawkins, B., Frank, J.S., & Quinn, J.T. (1979). Motor-output variability: A theory for the accuracy of rapid motor acts. Psychol. Rev., 86, 415-451.
22. Shannon, C.E. (1948). A Mathematical Theory of Communication. Bell System Techn. J., 27, 379-423 and 623-656.
23. Smith, R.J. (2009). Use and misuse of the reduced major axis for line-fitting. Amer. J. Phys. Anthropol., 140, 476-486.
24. Soukoreff, R. W., & MacKenzie, I.S. (2004). Towards a standard for pointing device evaluation: Perspectives on 27 years of Fitts' law research in HCI. IJHCS, 61, 751-789.
25. Soukoreff, R. W., and MacKenzie, I. S. (2009) An informatic rationale for the speed-accuracy tradeoff. Proc. SMC'2009. New York: IEEE, 2890-2896.
26. Stevens, S.S. (1946). On the theory of scales of measurement. Science, 103, No. 2684, 677-680.

9. APPENDIX: GLOSSARY OF BASIC VARIABLES

		Physical Dimension	Unit(*)
Task Geometry			
D	Target distance	[L]	(cm)
W	Target tolerance or width	[L]	(cm)
D/W	Relative target distance (*RTD*)	[-]	
W/D	Relative target tolerance (*RTT*)	[-]	(%)
Elemental Movement Measures			
T	Movement time	[T]	(s)
A	Movement amplitude	[L]	(cm)
$E = A - D$	Endpoint error	[L]	(cm)
Statistical Movement Measures			
μ_T	Mean or median movement time	[T]	(s)
μ_A	Mean or median movement amplitude	[L]	(cm)
$\mu_A - D$	Constant error	[L]	(cm)
σ_A	*SD* of mov. amplitude = variable error	[L]	(cm)
μ_A/σ_A	Relative movement amplitude (*RMA*)	[-]	
σ_A/μ_A	Relative variable error (*RVE*)	[-]	(%)

(*) *Note*. W being much smaller than D, it is convenient to express *RTT*, but not its inverse *RTD*, as a percentage. Likewise, σ_A being much smaller than μ_A, it is convenient to express *RVE*, but not *RMA*, as a percentage [8].