Small Targets: Why are they So Difficult to Acquire?

Olivier Chapuis[1,2] **Pierre Dragicevic**[2]
{chapuis,dragice}@lri.fr

[1]LRI - Univ. Paris-Sud & CNRS
Orsay, France

[2]INRIA
Orsay, France

ABSTRACT
Facilitating the acquisition of small targets is an active area of research. Despite the usefulness of the techniques proposed so far, we show why this topic can benefit from more fundamental work. We investigate three factors that might account for the observed drop in performance when selecting small targets: motor accuracy, visual legibility and quantization. Our findings suggest that visual size is as important as motor scale and that quantization has almost no effect. We also show that – as proposed by Welford in 1969 – motor inaccuracy is well-modeled by subtracting a "tremor" constant from target size. We conclude with implications for design.

ACM Classification Keywords
H. Information Systems H.5 Information Interfaces and Presentation H.5.2 User Interfaces (H.1.2, I.3.6)

Author Keywords
Fitts' law, pointing performance, small targets, quantization.

INTRODUCTION
Common tasks such as window resizing and text selection require pointing at items of only a few pixels. In mobile devices, small items are also common and notoriously difficult to acquire. Such problems have repeatedly been mentioned in the past and a variety of approaches have been proposed to ease the acquisition of small targets. Each of them has advantages and drawbacks and new solutions are still actively being explored.

However, surprisingly little is known about the reasons why small targets are actually difficult to acquire. We argue that answering that question can not only better guide future research and inform future designs, but can also help choosing among existing approaches and refining them.

Since Fitts' law is routinely used to motivate and inform research on pointing facilitation techniques, we first discuss its suitability as a conceptual framework in the context of small targets. We also show why taken together, previous studies on pointing do not provide a clear picture of why small targets are difficult to acquire.

LRI Technical Repport Number 1508
DECEMBER 2008

LABORATOIRE DE RECHERCHE EN INFORMATIQUE
UMR 8623, CNRS Université de Paris Sud
Bât. 490, 91405 ORSAY Cedex France

We then present a study in which three potential sources of problems are investigated: motor accuracy, visual legibility and quantization. Although most "small targets" mentioned in the Human-Computer Interaction literature combine at least the three of these problems, we show how new insights can be gained from a separate assessment of these three factors.

Why Fitts' Law is not Enough
Since Fitts' law has been proved to be an extremely useful theoretical framework, one could question the need for alternative models [27]. In this section we show why Fitts' law is actually not an adequate paradigm for studying small target acquisition techniques, even when used as a first approximation. We then list the benefits one can expect from more systematic investigations into that topic.

Fitts' Law Provides Misleading Arguments
It seems natural to appeal to Fitts' law to argue for a new small target acquisition technique. Fitts' law does predict that, all other things being equal, small targets are harder to acquire than larger ones. But Fitts' law is first of all a law about *scale invariance*: acquisition time depends solely on D/W, the ratio between target distance (D) and target width (W). D is hence as important as W and nothing in Fitts' law justifies a particular focus on target size. Furthermore, nothing justifies a particular focus on *small* target sizes.

Fitts' law additionally predicts that acquisition time increases logarithmically with D/W[1]. This means that reducing target size should not have an extremely strong impact. For example, if $W = 1$ and $D = 64$, the target theoretically needs to be expanded by 8 to yield a movement twice as fast.

Fitts' law is however contradicted by the following observation: performance seems to degrade rapidly when the target size falls below a certain threshold, typically below 4 or 5 pixels [23, 1, 22]. Sometimes, acquisition time and error rate literally explode: in one study [22], users missed a 1-pixel target more than ten times in a row in 25% of the trials.

Such observations suggest a strong scale effect by which small-scale pointing tasks are much more difficult than large-scale ones, D/W being held constant. It is this observed violation of Fitts' law – and not Fitts' law itself – which best justifies small target acquisition techniques.

[1]Here we can neglect the constant often added inside the log term and the non-null intercept often observed in the linear regressions.

Fitts' Law Provides Misleading Design Guidelines

Fitts' law also provides misleading guidelines for the design of small target acquisition techniques. In particular, it exclusively advocates the use of semantic pointing approaches and cannot discriminate between them.

The term *semantic pointing* has been originally used for a technique that lowers mouse gain inside targets [5]. One can however generalize it to any pointing aid that *uses information about targets*. Such techniques typically reduce D/W by manipulating target widths and/or distances in the motor and/or the screen space [2]. Some of them, like sticky targets [9, 7] and area cursors [16, 29], have been proposed as a solution to the small target problem.

Although Fitt's law provides theoretical support for the idea of reducing D/W, it might actually underestimate the improvement obtained on very small targets. Recall that Fitts' law only predicts a weak effect of target expansion on movement time, no matter the original target size. Similarly, Fitts' law does not predict that reducing D might not be as efficient as increasing W, in case W is very small.

Despite their effectiveness, semantic strategies are limited [2]: they essentially consist in redistributing targets to have their D/W ratios better match their probability of acquisition — a special case being removing intervening spaces with zero click probability [13]. Often there is little or no space to remove, one example being the color picker in image authoring applications.

Simple non-semantic approaches nonetheless exist that could facilitate acquisition despite leaving D/W unchanged. They consist in increasing all W's and D's in the same proportions, i.e., a uniform *magnification* solution. Like semantic pointing, magnification can be performed in the screen space, in the motor space, or in both. Examples of screen-space magnification techniques are screen magnifiers [17] and fisheye views [10]. Lowering the control/display (C-D) gain yields a uniform motor-space magnification [6]. Zooming magnifies both screen and motor space [4, 22].

Key Questions

We argue that the following three questions regarding small targets are worth investigating due to their potential implications for design:

1. *Is there any small-scale effect?* If this hypothesis is wrong, it would imply that semantic pointing is the way to go. If it is true, it would imply that a) non-semantic magnification approaches should receive more attention and b) among the possible semantic strategies, expanding W should work best.

2. *What are the causes of the small-scale effect?* Assuming the hypothesis above holds, it is still unclear whether semantic and non-semantic target expansion should be performed in the screen space, in the motor space, or both. In some cases target expansion might not even be enough. Answering these questions requires identifying, isolating and testing potential causes of problems.

3. *What would be a good model of small target acquisition?* Assuming efficient techniques can been identified, it is still unclear how to tune them. For example, will a magnification of 4x be enough? What gain in performance can be expected? Being able to assess movement time according to factors such as D and W will help answering these questions.

Potential Sources of Problems

With respect to our second question, we chose to investigate three potential sources of problem with small targets:

- Small motor scale: Targets that are small in the motor space could be difficult to acquire because they demand too much motor precision.
- Small visual scale: Targets that are small in the screen space are difficult to see and this could make their acquisition more challenging.
- Quantization: Moving 4 pixels to click on a 1-pixel target feels less smooth than moving 400 pixels to click on a 100-pixels target, and this could be a source of problem as well.

The last point is rarely mentioned. We always assume that pointing on a computer screen is a smooth and continuous task, despite the fact that the screen is a discrete space. Although this approximation is correct most of the time, quantization starts to be noticeable when acquiring a target of only a few pixels in size. In fact, problems with high C-D gains have been sometimes attributed to the fact that the mouse cursor jumped by steps of more than one pixel due to its low resolution [15, 6]. This implies that quantization of the visual feedback could result in a drop in performance.

There can be other sources of problems in small-scale target acquisition tasks: physical occlusion [25], parallax [22], and landing/take-off inaccuracies [23] are all common issues in direct-touch and pen-based user interfaces. Even on desktop computers, the mouse cursor can occasionally occlude small targets. In this paper, we however chose to solely focus on the three potential sources of problem enumerated above, which are intrinstic to all small target pointing tasks. We believe that taken together, they form a general and rather complete definition of the concept of "small target".

RELATED WORK

In this section, we briefly recall previous work on small target acquisition techniques. Also relevant to our questions above are previous investigations into scale effects in Fitts' law and some of the proposed alternatives to Fitts' law.

Small Target Acquisition Techniques

There has been considerable interest for pointing facilitation techniques. Studies focusing on small targets have been conducted on different types of hardware.

Desktop computers. A number of studies demonstrated the efficacy of motor-space semantic magnification for small targets [9, 7, 16, 29]. At the same time, it was suggested that in-target feedback in the form of visual or auditory confirmation yield at best marginal improvements [7, 8].

Pen-based devices. Ren et al. studied the use of pen landing and pen take-off for selection and found 5 pixels (1.8 mm) to be a critical target size [23]. Ramos and al. showed that pressure-activated zooming lenses dramatically facilitate pointing for targets less than 4 pixels (1.1 mm) [22].

Touch-screens. By suppressing finger occlusion, the take-off technique was shown to reduce errors but tiny targets were still abnormally long to acquire [25]. Albinsson et al. proposed the use of discrete taps or a levering scheme that magnifies motor space [1]. Both methods were useful for targets of 1 pixel (0.4 mm) but zooming was found to be the fastest of all. The authors still advocated approaches that leave visual content unchanged and argued that for pointing, "the limitation is in control, not visual resolution" [1].

Taken together, these studies provide enough anecdotal evidence for a small-scale effect and suggest ways to overcome it. But besides pointing out a "problematic target size threshold", they have little explanatory power and generalizability. Fitts' law analyses are either absent or deemed inconclusive — with the exception of two studies that showed a good fit but with a single target size condition [9, 7].

Furthermore, it is not clear what problems each of the existing techniques solves, especially on direct pointing devices where occlusion, parallax and landing/take-off imprecision can all affect pointing. Finally, most techniques have several parameters — e.g., magnification factor — whose values seem to be chosen arbitrarily and vary across studies.

Scale Effects in Fitts' Law Studies

Although scale effects have been occasionally studied from the perspective of target distance [19, 11], here we essentially focus on scale effects caused by small target widths. Several early studies on Fitts' law mentioned such small-scale effects (for a review, see [20]), but the direct tap paradigm made it difficult to determine its exact causes.

Some answers can be found in studies on C-D gain, which typically manipulate the scale of pointing tasks in the motor space. Despite conflicting results (for a review, see [6]), some researchers postulated that the drop in performance sometimes observed with high C-D gains was due to quantization [15, 6]. After controlling for quantization, Casiez and al. still observed a slight decrease in performance for high C-D gains that they attributed to a motor accuracy issue [6].

Langolf et al. studied small target acquisition on a stylus-pegboard configuration under microscope magnification. They found that by fixing D, acquisition time followed Fitts' law, suggesting no motor acurracy issue [19]. Similarly, Guiard et al. used a double-scale visual magnification scheme to test higher IDs and found a target size effect for the puck but not for the stylus [12]. He postulated that with a precision grip, "the likely limiting factor for tolerance [W] is vision, not motor control".

Taken together, these studies suggest that motor scale, visual scale and quantization can all contribute to the small-scale effect. However, the relative importance given to these factors varies considerably across studies. Although the use of different muscular groups partly explains that discrepancy [19, 3, 11], we know nothing about the relative influence these three factors can have on a specific hardware configuration, such as a standard desktop computer.

Alternatives to Fitts' Law

A number modifications to Fitts' law have been proposed to improve its fit with observed data (for a review, see [24]). While some of them still express movement time (MT) as a function of D/W, D and $1/W$ must be given asymetric roles to account for small-scale effects. We introduce four such formulae here [2].

A well-known formula is by Welford and assumes different throughputs for the ballistic and the homing phases of movement [28]:

$$MT = a + b.\log_2(D) + c.\log_2(W) \quad (1)$$

Kvålseth also proposed a power model that exhibited a better fit to the original Fitts' data [18]:

$$MT = a.D^b.W^c \quad (2)$$

Note that this is simply a logarithmized version of Equation 1. In order to account for the interactions between W and D, Oel et al. argued for the following refinement [21]:

$$MT = a.D^{b+d.log_2(W)}.W^c \quad (3)$$

Finally, the first part of Welford's paper contains an alternative model, which to the best of our knowledge has never been used in HCI [28]:

$$MT = a + b.\log_2(\frac{D}{W-c} + 1) \quad (4)$$

Where c is an experimentally-determined constant attributed to hand tremor.

EXPERIMENT

We conducted a user study in order to confirm the existence of a small-scale effect and investigate its causes on a standard mouse-screen configuration. We asked subjects to perform 1-D target acquisition tasks and independently manipulated visual scale, motor scale and quantization. We first introduce our experiment design and our use of *scaling methods* as a way to manipulate these three factors. We then discuss our main findings and their implications.

Scaling Methods

For consistency (i.e., the lower value the harder), we will use here the term *continuity* as the inverse of *quantization*. One can think of *visual scale*, *motor scale* and *continuity* as forming a 3-D space that can used to characterize pointing tasks. All three dimensions are *relative*: a visual scale of 2 would magnifying D and W by $2\times$ on the screen according to a nominal pointing task. A continuity of 2 would mean doubling the resolution of the mouse and of the screen.

[2] We have added the missing intercept a to the original Equations 1 and 4, and replaced 0.5 by 1 in Equation 4 to allow easier comparison with Mackenzie's widely-used formulation of Fitts' law [20].

An ideal experiment would evenly sample this 3-D space. However, there are two difficulties. First, all points cannot be measured due to hardware constraints: if a pointing task involves a 1-pixel target, it is not possible to scale it visually by a factor of 0.5 or 1.5, nor is it possible to make it more continuous. Second, evenly paving the space would yield too many conditions for a single experiment. We hence devised an approach based on scaling methods.

A *scaling method* is a specific way of scaling up a pointing task. Suppose the task is to acquire a target of distance and size D_v, W_v on the screen and D_m, W_m in the motor space. Let D_s be the number of discrete steps required for the cursor to progressively reach the middle of the target, and W_s the number of possible cursor locations inside the target. Typically, these are equal to D_v and W_v as expressed in pixels. Finally, let C_v be the cursor's visual size and G the C-D gain, i.e., the ratio between the cursor's and the mouse's travel distances, both expressed in metric coordinates.

Such a target acquisition task can be scaled up by a positive integer S using five possible methods:

- *MotorMag* multiplies D_m and W_m by S and leaves the rest unchanged, except for G which is multiplied by $1/S$. In other terms, this is a motor-space magnification.
- *VisualMag* multiplies D_v, W_v and C_v by S and leaves the rest unchanged, except for G which is multiplied by S. In other terms, this is a visual-space magnification.
- *FullMag* multiplies D_m, W_m, D_v, W_v and C_v by S, and leaves the rest unchanged. This amounts to increase the size of the pixel and of the mouse dot.
- *Zoom* is the same as FullMag except that C_v is left unchanged and W_s and D_s are multiplied by S. This amounts to enlarge the task in both spaces while keeping the original pointing resolution.
- *ZoomBis* is the same as Zoom, but in addition, the cursor's size is multiplied by S.

With respect to the previous 3-D space model, the MotorMag and VisualMag methods cover the *motor scale* and the *visual scale* axes. The FullMag method covers the bisection of these two axes, whereas Zoom and ZoomBis both cover the bisection of the three axes.

We introduced the ZoomBis variant because one consequence of FullMag is to increase the cursor's size, while the cursor remains visually small for Zoom — hence a possibility of confound. In ZoomBis and FullMag, the cursor is the same. However, since subjects still have to bring the whole cursor into the target, W_v must be enlarged by the same amount as C_v in order to preserve W_s and W_m.

Apparatus

We conducted the experiment on a high-end workstation running X Window and Java. The display was a 22" ultra-high-resolution LCD monitor with a native resolution of 3840×2400 and 0.125 mm pixel size (about 200 dpi). The pointing device was a ultra-high-resolution gaming mouse of 83.5 dots per mm (about 2000 dpi). The mouse was teflon-coated and was used on a varnished plywood surface.

We set the display at half its native resolution and programmatically divided by 5 the resolution of the mouse, hence approaching the resolutions and C-D gain of a standard desktop computer while virtually eliminating potential confounds due to display quality and input sensor accuracy. No mouse acceleration was used.

Participants

Twelve unpaid volunteers, ten male and two female, all right-handed, participated in the experiment. Two additional participants failed the vision test (see below) and thus did not proceed with the experiment. Participants were experienced mouse users with ages ranging from 21 to 32 (median 25).

Task and Procedure

A trial was a 1-D discrete target acquisition task decomposed as follows. First, a gray target and two gray markers symbolizing the location of the cursor appeared at the center of the screen (Figure 1a). The participant had been instructed not to move the mouse until the actual (black) cursor appears. After a short random foreperiod, the black cursor was shown (Figure 1b) and the participant had to bring it inside the target (Figure 1c) and press the left mouse button.

To better resemble a typical computer pointing task, a trial had to end successfully even if it included mouse presses outside the target. The next trial then started after the participant had released the mouse button. We recorded movement times from the first mouse movement to the first mouse press, as well as to the first successful mouse press. The participant was allowed to rest every 22 trials.

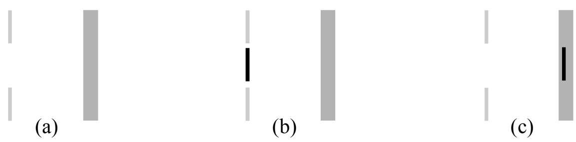

Figure 1. A sample trial: (a) the target and the future location of cursor are shown; (b) the cursor appears after a random foreperiod; (c) the user selects the target.

Before starting the experiment, participants were given written instructions telling them to be as fast and as accurate as possible and to avoid mouse clutching. To allow them to do so, mouse drifting was eliminated by grouping trials by pairs with the same target distance and width, but with opposite movement direction.

Upon reading the instructions, participants were asked to sit in a comfortable position and the screen was moved at a distance of 70 cm to their eyes. A string was then placed in front of their chest to remind them not to lean forward, as mentioned in the written instructions.

Participants were then given a vision test involving 12 labeled representations of a target and a cursor, where the cursor was 1-pixel wide and the targets 1 or 2 pixels wide. The cursor was either inside the target or just next to it. Participants were asked to tell when the cursor was inside the target. The test was considered successful if all the targets were properly identified and there was no false positive.

Design

The experiment was a within-participant design with the following factors:

- five scaling methods METHOD: Zoom, ZoomBis, FullMag, MotorMag and VisualMag;
- four scales SCALE: 1, 4, 16 and 64;
- four nominal widths W: 1, 2, 4 and 8 pixels/dots;
- four nominal distances D: 2, 4, 8 and 16 pixels/dots.

The design was not a full factorial. Instead, we only included all the possible combinations of D $= \{8, 16\}$ and W $= \{1, 2, 4, 8\}$ plus the combinations $(2, 1)$, $(4, 1)$ and $(4, 2)$, for a total of 11 (D,W) couples. (D,W) was fully crossed with METHOD and SCALE, giving a total of 220 conditions. Note that Fitts' IDs remained constant across METHOD and SCALE.

As mentioned before, movement directions were alternated to limit mouse drift. For each METHOD × SCALE condition, participants were presented with 5 blocks of Direction × W × D = 5 × 2 ×22 = 110 trials. The presentation order of W × D was randomized within each block. The first block was for training and the remaining four were recorded.

A pilot study suggested that large variations in C-D gain were the most difficult to accomodate (the nominal gain was 4.15 and ranged from 0.065 to 4.15 for MotorMag and from 4.15 to 266 for VisualMag). We hence blocked by METHOD and sub-blocked by SCALE. For MotorMag and VisualMag, we presented the SCALE conditions monotonically and informed participants of changes in mouse sensitivity through text messages. For the other methods, where the C-D gain is constant, the presentation order of SCALE was randomized.

Additionally, in order to avoid important changes in C-D gain when transitioning between MotorMag and VisualMag, we always presented these two methods at the second and at the fourth position. We computed a Latin square for (Zoom, ZoomBis, FullMag) and crossed it with the two possible orderings of (MotorMag, VisualMag), yielding six different orderings for METHOD, each of which was presented to two participants.

Finally, note that in the case of SCALE 1, all the METHOD conditions are equivalent. We hence decided to present the condition only once and arbitrarily assigned it to the Zoom METHOD. This removed 4 (SCALE) × 5 (Blocks) × 11 (W×D) × 2 (Direction) = 440 redundant trials.

A participant hence performed 5 (METHOD) × 4 (SCALE) × 5 (Blocks) × 11 (W×D) × 2 (Direction) − 440 = 1760 pointing tasks, 1408 of which were recorded. We obtained 96 measures for a full condition and a total of 16896 measures. The experiment lasted approximately 65 minutes.

RESULTS

We first perform a full factorial analysis by removing the data for D=2 and D=4 and by duplicating the data from Zoom METHOD at SCALE=1 to the other methods. The remaining data, e.g., the (D, W) couples (2, 1), (4, 1) and (4, 2), will be dealt with later on.

	All SCALE			SCALE≥ 4		
Factors	DF,Den	F	p	DF,Den	F	p
METHOD	4,44	71.2	< 0.0001	4,44	71.2	< 0.0001
SCALE	3,33	35.7	< 0.0001	2,22	16.9	< 0.0001
W	3,33	166.9	< 0.0001	3,33	560.0	< 0.0001
D	1,11	234.5	< 0.0001	1,11	1441.4	< 0.0001
METHOD×SCALE	12,132	36.3	< 0.0001	8,88	15.9	< 0.0001
METHOD×W	12,132	27.3	< 0.0001	12,132	27.3	< 0.0001
METHOD×D	4,44	23.0	< 0.0001	4,44	23.0	< 0.0001
SCALE×W	9,99	17.7	< 0.0001	6,66	16.0	< 0.0001
SCALE×D	9,33	2.4	0.0829	2,22	11.8	0.0003
W×D	3,33	0.7	0.5509	3,33	2.0	0.1365
METHOD×SCALE×W	36,396	8.0	< 0.0001	24,264	2.4	0.0004
METHOD×SCALE×D	12,132	2.9	0.0013	8,88	0.6	0.7727
METHOD×W×D	12,132	1.6	0.0857	12,132	1.6	0.0857
SCALE×W×D	9,99	1.6	0.1226	6,66	1.9	0.0908
METHOD×SCALE×W×D	36,396	1.3	0.1262	24,264	1.2	0.1951

Table 1. Results of the ANOVA for MT ∼ METHOD × SCALE × W ×D.

Movement Time

In this analysis, we consider the movement time MT to the successful mouse press, as this measure has the advantage to include penalties caused by the errors. We performed a factorial repeated measures analysis of variance for the model MT ∼ METHOD × SCALE × W × D × Random(PARTICIPANT). Table 1 shows the results of this ANOVA for the data, with and without SCALE=1. We first observe a significant effect of METHOD and SCALE on MT and a significant interaction METHOD × SCALE.

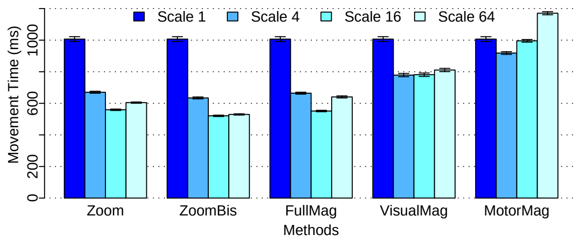

Figure 2. Movement time as a function of SCALE, grouped by METHOD. The five darkest bars represent the same data.

Figure 2 shows the effect of SCALE on each METHOD. We observe that Zoom and ZoomBis follow a similar pattern: MT decreases as scale increases, up to a scale of 16. A post-hoc Tukey test for difference in means confirms these observations. FullMag is similar but we observe a drop in performance for SCALE=64. With MotorMag, we see a small performance improvement from scale 1 to 4 and then a degradation. For VisualMag, we also find an improvement from scale 1 to 4, but higher scales seem to have no effect. All these observations were confirmed by post-hoc tests.

We can also observe in Figure 2 that Zoom, ZoomBis and FullMag seem very close. Indeed, the only difference shown by a post-hoc Tukey test is between ZoomBis and FullMag for SCALE=64.

Now we analyze the factors W and D. Unsurprisingly, we found an effect on MT. Figure 3 illustrates these effects as well as the interactions with the factor METHOD. We can see that the difference between methods decreases as W increases. In particular, if we exclude MotorMag, the difference between VisualMag and other methods is large for W=1

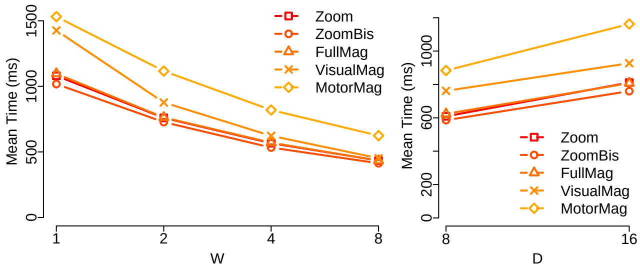

Figure 3. MT **as a function of** W **(left) and** D **(right), for each** Method.

but vanishes for W=8. One can also observe that the impact of D is weaker with VisualMag than with the others methods.

Figure 4 illustrates the interactions between W and Scale, for the Zoom method (left of the figure) and for MotorMag (right). We can see that for Zoom, the difference between Scale=1 and the others scales decreases as W increases. ZoomBis, FullMag and VisualMag exhibit a similar trend, but for VisualMag and Scale$\geq$ 4 the lines are almost confounded (figures not included). However, as can be observed in Figure 4 (right), the situation is completely different with MotorMag.

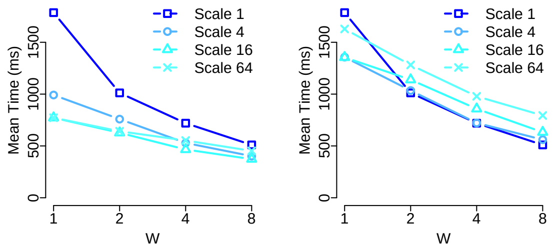

Figure 4. MT **as a function of** W **for Zoom (left) and MotorMag (right), for each** Scale**.**

Error Rate

We found an overall error rate of 7.02% (5.80% for Scale$\geq$ 4). The ANOVA reveals a significant effect of Method and W for both datasets under consideration (all Scale and Scale$\geq$ 4), and a significant effect of Scale for the data with all Scale. No effects were found for D.

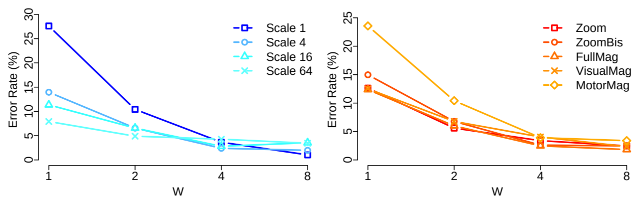

Figure 5. Error rate as a function of W**, for each** Scale **and each** Method**.**

We also observed significant interactions of Method $\times$ W and Scale $\times$ W (all Scale and Scale$\geq$ 4), as shown in Figure 5. For Scale=1 in particular, we have a higher error rate than other scales when W=1 but this difference vanishes as W increases. The same can be said when comparing MotorMag with other methods. For the remaining Method, we only see very small differences in the error rates.

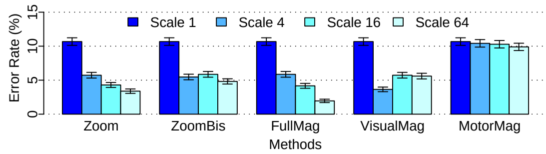

Figure 6. Error rate as a function of Scale **grouped by** Method

We also found a significant interaction of Method $\times$ Scale, but this interaction was no longer significant when restricting the data to Scale$\geq$4. Figure 6 shows the error rate as a function of Scale grouped by Method.

Learning

Regarding learning effects, we performed an analysis of variance for the model MT $\sim$ Block $\times$ Method $\times$ Scale. We found a significant simple effect of Block for the Scale$\geq$4 data ($F_{3,33} = 6.4$, $p = 0.0015$) but not for the full data set. However, we found significant interactions for Block $\times$ Method ($F_{12,132} = 2.5$, $p = 0.0055$) and Block $\times$ Method $\times$ Scale (and no Block $\times$ Scale interaction) for both datasets. In Figure 7 we can observe that the learning effect is essentially caused by MotorMag and that this effect is specially strong for Scale=64. Indeed, post-hoc Tukey tests show a significant difference in means (around 100 ms) between block 5 and blocks 2 and 3 for MotorMag but no difference for the other methods.

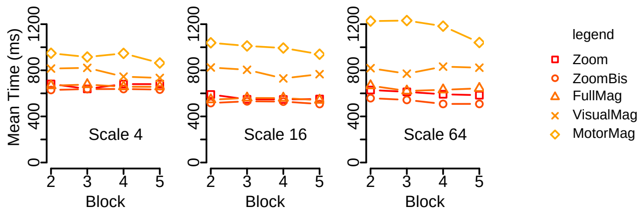

Figure 7. MT **as a function of block by** Method **for each** Scale $>$ **1.**

Small Distances and Small Widths

So far we have excluded the trials where D=2 and D=4. To study these cases, we have performed full factorial analyses of variance for *(i)* the model MT $\sim$ Method $\times$ Scale $\times$ D on the data subset where D$\in \{2,4,8,16\}$ and W=1 and *(ii)* the model MT $\sim$ Method $\times$ Scale $\times$ W $\times$ D on the data subset where D$\in \{4,8,16\}$ and W$\in \{1,2\}$. We obtained results very similar to those discussed previously.

Variants on MT Measures

One may want to stick to a more traditional methodology [20] and either: *(i)* take movement times up to the first click (MT_0) instead of the successful click (MT_1); *(ii)* remove errors ($MT_1 = MT_0$) and outliers (using MT_0). In both cases the analysis yields results that are very similar to those we obtained so far. The results of the ANOVA's are similar, with different p and F values but without change in significance. Of course, conditions with high error rates become faster, but the differences are small and do not affect the results.

Scale Blocks and Duplicate Tasks

In the Zoom condition, some tasks were duplicated across different Scale blocks. In fact, the (D,W) couples $(2,1)$,

$(4,1)$ and $(4,2)$ at SCALE=4 were the exact same tasks as the couples $(8,4)$, $(16,4)$ and $(16,8)$ at SCALE=1. The same is true for scales 16 and 4, and scales 64 and 16. One may expect to find similar movement times and error rates. This was not the case: on average, participants were faster but did more errors in the high-scale block. This can be explained by the fact that duplicated (D,W) couples were the most difficult tasks in the high-scale blocks, but the easiest ones in the low-scale blocks. Hence, the overall difficulty of the block might have influenced the speed-accuracy tradeoff.

MOVEMENT TIME MODELS

As in standard Fitts' law analyses we consider here movement time to the first click [20]. Moreover, we aggregate the data over METHOD, SCALE, W and D, taking the mean over all participants. Target widths and distances denoted by W_m and D_m are expressed in the motor space, in mouse dots ($1\text{dot} = 0.060\text{mm}$). Although the actual unit does not matter in a standard Fitts' law analysis (MT only depends on D/W), it does matter when testing alternatives to Fitts' law.

Visual Magnification

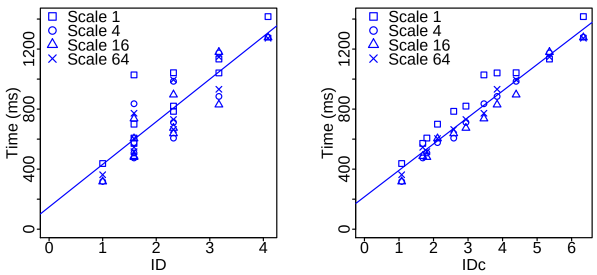

Figure 8. MT as a function of ID and $ID_{c=0.8}$ for VisualMag.

For VisualMag, the left plot in Figure 8 shows movement time as a function of Fitts ID and the corresponding regression line. The fit is not extremely good ($r^2 = 0.767$), with the smallest targets lying way above the regression line. Correcting IDs using Welford's tremor model solves this issue (see Equation 4):

$$ID_c = \log_2\left(\frac{D_m}{W_m - c} + 1\right) \qquad (5)$$

Taking $c = 0.8$ significantly improves the fit ($r^2 = 0.946$), as can be seen in Figure 8. For SCALE$\geq$4, the improvement goes from $r^2 = 0.799$ to $r^2 = 0.982$.

Note that this model does better than the other (more popular) model from Welford (Equation 1) that gives $r^2 = 0.931$ for SCALE≥ 4 and $r^2 = 0.904$ for all SCALE. Moreover, it is as good as Oel *et al*'s model (Equation 3) that gives $r^2 = 0.978$ and $r^2 = 0.942$, despite having one more degree of freedom. This suggests that Welford's tremor model accurately models pointing for targets that are small in the motor space.

Motor Magnification

The left plot in Figure 9 shows the standard Fitts regression with MotorMag for SCALE$\geq$4. Like VisualMag, the fit is

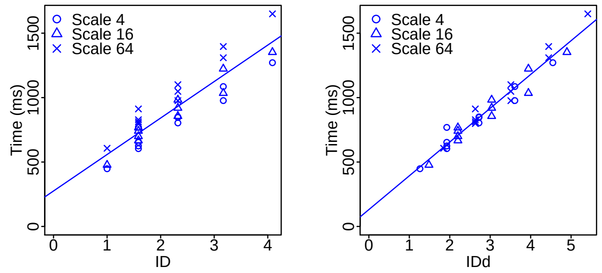

Figure 9. MT as a function of ID and $ID_{d=0.2}$ for MotorMag.

not extremely good ($r^2 = 0.824$), but the alternative models mentioned above do not yield satisfactory results. This suggests a different phenomenon.

For MotorMag, recall we observed an important decrease in performance as scale grows. One may hence postulate problems originating from the stimulus-response (S-R) incompatibility between large-scale mouse movements and a small-scale visual feedback. One way to correct for the ID is to increase target distance as the scale grows and this led us to consider the following model:

$$ID_d = \log_2\left(\frac{D_m + d.\sqrt{\frac{4.15}{CDGain} - 1}\,.D_m}{W_m} + 1\right) \qquad (6)$$

where d is a constant and $4.15/CDGain$ expresses the S-R incompatibility (4.15 is the C-D gain for SCALE=1).

ID_d improves the fit of the MotorMag data for SCALE$\geq$4, as shown in Figure 9 ($r^2 = 0.953$). Again, our results are comparable to Oel et al ($r^2 = 0.931$), despite their extra degree of freedom.

Zooming

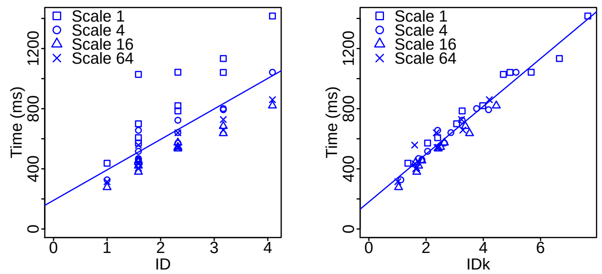

Figure 10. MT as a function of ID and $ID_{k=0.92}$ for Zoom.

The left plot in Figure 10 shows the standard Fitts regression with the Zoom method. The bad fit ($r^2 = 0.534$) is essentially caused by the condition SCALE=1. If we remove this condition we obtain $r^2 = 0.825$. But like VisualMag, adopting Welford's tremor model is enough to improve the overall fit, with $r^2 = 0.881$ and $c = 0.8$.

Interestingly, applying Welford's tremor model after removing SCALE=1 yields a higher tremor constant $c = 2.5$ ($r^2 = 0.933$). This suggests that c might vary with W_m, leading us to consider the following alternative:

$$ID_k = \log_2\left(\frac{D_m}{W_m - k \times \log_2(W_m + 1)} + 1\right) \qquad (7)$$

With $k = 0.92$, we obtain $r^2 = 0.951$ for all scales (see Figure 10) and $r^2 = 0.921$ for SCALE≥4, which is better than ID_c and again similar to Oel et al.

We obtained very similar results for ZoomBis. For FullMag we also obtain relatively good fits, yet not as good. A reason for this is that the ID_c and ID_k models do not account for the degradation we measured for scale 64.

Adjusting for Effective Widths

The idea of effective width, introduced by Crossman in his unpublished doctoral dissertation, is to perform "an adjustment for accuracy": the width of the target is corrected so that (under certain hypotheses) the data yields an uniform error rate of 4%. We refer the reader to [26] for computational details and more motivations for the use of W_e.

One might hope that using effective widths will normalize speed between scale blocks, as we noted important differences in error rates. However, recall that for Zoom for example, faster scale blocks are those with less error rates. In fact, if we compute the motor effective widths, we find they enlarge small motor targets (from 1 to 4), they leave motor targets unchanged from 8 to 64 and shrink motor targets from 128 to 512. That explains why we actually obtained less good fits with ID_e than with ID and suggests that the use of effective widths is not adapted to our study.

However, using the ID_c model together with W_e leads to a slightly better fit than ID_c alone (SCALE≥ 4, Zoom and ZoomBis). This yields higher values of c and this is explained by the fact that W_e measures, in part, motor noise.

DISCUSSION

Here we build upon our findings to address the three questions mentioned in the introduction. We then discuss the limits of our study and suggest possible options for future work.

There is a Small-Scale Effect

Our experiment confirms previous results and intuitions about small-scale effects. This is shown by Zoom and ZoomBis: at high scales, we have normal pointing tasks. At scale 1, however, we have targets that are small in all respects, i.e., typical "small targets" on desktop computers. Task IDs do not change. Still, we observe a clear drop in performance as scale approaches 1. This naturally yields a bad fit with Fitts' law.

The drop in performance is progressive yet very fast: it starts to be observable from scale 16 (targets of 16-256 pixels[3]) to scale 4 (targets of 4-32 pixels). At scale 1 (targets of 1-8 pixels), there is more than twice as much errors, for a speed almost twice as slow. Since we used high-end I/O hardware and controlled for pointer occlusion, these are likely underestimates of what we would obtain on an actual computer.

[3]All figures are given according to our experimental conditions, i.e., 0.06 mm mouse dot size, 0.25 mm pixel size and a distance of 70 cm to the screen.

What Causes the Small-Scale Effect

Our experiment suggests that — at least on desktop computers — the causes are both visual and motor. Quantization does not seem to significantly impact performance.

Visual Causes

For very small targets, the primary causes of the small-scale effect are visual. Data from VisualMag provides strong evidence for this. Visual size has no effect when it is large enough. But we observe an important deterioration from scale 4 (targets of 4-32 pixels) to scale 1 (targets of 1-8 pixels): there are twice as much errors and a reduction in speed of about 25%. Our data actually suggests important differences for targets up to 4 pixels.

The exact origin of these visual problems is unclear, since we actually controlled for visual legibility (we use a sharp monitor and gave a vision test). However, visual legibility is probably not a binary property: participants might have spent more time processing visual information because of its unusually small size. This could have slowed them down.

Motor Causes

We also confirmed the common intuition that motor accuracy is a cause of the small-scale effect. This supported by our data on MotorMag. There is an observable deterioration from scale 4 (targets of 4-32 mouse dots) to scale 1 (1-8 dots). Data suggest problems start to arise at 4 dots. Since we controlled for the accuracy of electronic sensors the problem is likely physiological, although mechanical factors such as mouse/surface friction could also play a role.

The role of motor accuracy is further supported by the data from VisualMag and FullMag: we observed that motor+visual magnification performs slightly better than visual magnification alone, and the improvement goes up to scale 8 while it stops at scale 4 for visual magnification alone.

Quantization

The quantization present in small-scale pointing tasks seems to be a very secondary cause to the small-scale effect. This is suggested by our data on FullMag and Zoom. Recall the first method simply magnifies pixels and mouse dots, whereas the second one improves pointing resolution as scale increases. We found no significant difference between the two methods overall, except for scale 64 (targets from 64 to 512 pixels).

One possible explanation is that although quantization deteriorates the information that is normally available during pointing, it also provides "snapping". With snapping, the cursor is clearly either inside or outside the target. However, it is not clear whether snapping actually helps. In fact, we found dwell time to be higher for FullMag than for Zoom at scale 64. Another possibility is that the human visuo-motor system is able to infer the missing information.

Low C-D gains

When the target is small in only one respect, one extra factor to consider is C-D gain. Recent findings suggest it should not affect performance [6]. Although we found this to be true overall — especially for high C-D ratios (VisualMag) — MotorMag suggests this is not true for very low C-D ratios (despite no mouse clutching). This is supported by the degradation of performance we observed as motor magnification increases (C-D ratios of 0.26 and 0.06).

Participants had trouble performing large and fast movements when the visual feedback provided was very small, as we observed lower peak velocities with MotorMag than with Zoom. However, this is likely a matter of training, as suggested by the learning effect we found on MotorMag.

Is there a Law of Movement for Small Targets?

One important finding from our study is that small target acquisition is complex. It involves different phenomena, and there is probably no simple law for small target acquisition.

Welford's model of human tremor was nevertheless quite successful at modeling motor problems when acquiring small targets. It has a simple interpretation: one can think of an area cursor whose hot spot's location is unknown — to avoid random errors, the user would have to bring the whole cursor inside the target, hence effectively reducing its size.

Note that the tremor constant c is not the same as the motor noise from the traditional impulse variability model, which is linear and scale-independent [27]. $(W - c)$ is also different from the concept of effective width [26]. Effective width is computed from a single target width, whereas c is obtained through model fitting on several target widths.

One difficulty with Welford's model is that c can vary. Indeed, we found c to be higher for Zoom than for VisualMag. We proposed a modified model where c depends on W_m, but this model is essentially computational. It could be that humans can adapt their "constant" motor noise to the overall task demands [27].

Our motor magnification model is also mostly computational, and captures both visual legibility problems and low C-D gain issues. For low C-D gains we observed lower peak velocities, and it hence seemed natural to assume that D_m will penalize MT more than W_m. The good fit we obtained after correcting for D_m suggested this was the case.

Limits of the Study

Distance Range and IDs. We chose to test a wide range of scales rather than large distances. The next step could be to restrict the scale range and explore larger distances in order to further validate and possibly refine our models.

Intermediate Scales. We did a sparse sampling of the scale range. For all scaling methods we observed the largest changes between scales 1 and 4. So one direction for a further experiment could be the study of intermediate scales.

Hybrid Scaling. We only investigated specific values for the visual, motor and quantization factors. As we now know better about visual limits, we can chose a large enough visual scale (4) and then study the effect of motor scale.

Error Rate. Another limit is the differences we obtained in error rates, despite having instructed subjects to be accurate. Pilot studies did suggest that maintaining a 4% error rate would be challenging given the difficulty of some tasks. We hence opted for a solution consisting in having the subject finish the task. This is closer to real computing tasks where failed clicks are frequent, especially with small targets [22].

Generalizability. Out study involves 1-D tasks. With 2-D pointing tasks, some factors — especially quantization — might have behaved differently. Similarly, all our findings concern moused-based desktop computers and the results would have probably been different with, e.g., pen devices or touch-screens.

IMPLICATIONS FOR DESIGN

Our study shows that problems start to arise below a certain target size. This size is at *least 1 mm* on a screen at a distance of *70 cm*, and *at least 0.2 mm* in the motor space (for a computer mouse). These values can be higher depending on the hardware quality.

In these cases, one should consider adding support for *zooming*. Zooming solves both visual and motor issues and preserves C-D ratio. It is especially useful on dense populations of targets, such as in text editors, image processing applications and vector graphics applications. The zoom factor should be chosen so that the smallest targets are enlarged just above the sizes mentioned above (typically 4×).

The traditional solution is to use zoom & pan navigation. For the sole purpose of target selection, however, a better strategy might be *temporary zooming* [22], e.g., a spring-loaded zoom mode. But all known zooming approaches have pros and cons and this topic certainly deserves more research.

Zooming is not necessarily the best strategy in all case scenarios. When targets are isolated and when it is crucial to preserve visual context, *semantic pointing* could be the way to go. Note that semantic pointing works for small targets as well as large targets.

Alternatively, *visual magnification* can also facilitate pointing and is supported natively in some OSs. Variants are magnification lenses and fisheyes, but the motion artefacts they produce can be distracting [14]. *Motor magnification* alone can also be used on low-accuracy input devices, but it is advisable to keep the *C-D gain above 1*.

Another good thing to know is that visual quantization does not harm. For example, one could hide the mouse cursor inside menus to improve text legibility, if roll-over effects are already present. An alternative idea could be to simply make the mouse cursor smaller or less visible.

And finally, Fitts' law does not fit all. When working with small targets, Welford's tremor model (Equation 4) seems more adapted and might better help answering questions such as which scale factor — or pointing technique, or input device — is the best for a specific target layout.

CONCLUSION AND FUTURE WORK

The goal of this paper was to better understand why small targets are so difficult to acquire, a question that has been only partially addressed by previous work. We had subjects perform small target acquisition tasks on a desktop computer, and manipulated the scale of the tasks by various means. Our study confirmed the existence of a "small-scale effect" that violates Fitts' law and whose causes are both visual and motor. One important implication is that more attention should be given to zooming techniques.

We have also found our data to be consistent with Welford's model of human tremor, where IDs can be normalized by removing a constant from target widths. We argued for the adoption of this model instead of Fitts' law when designing for small targets. Although our work is exploratory, we generated testable hypotheses and suggested directions for further investigations. These include testing larger distances and IDs, as well as using a finer sampling of small scales to better understand what happens "near the singularity".

REFERENCES

1. P.-A. Albinsson and S. Zhai. High precision touch screen interaction. In *Proc CHI'03*, 105–112. ACM, 2003.
2. R. Balakrishnan. "Beating" Fitts' law: virtual enhancements for pointing facilitation. *Int. J. Hum.-Comput. Stud.*, 61(6):857–874, 2004.
3. R. Balakrishnan and I. S. MacKenzie. Performance differences in the fingers, wrist, and forearm in computer input control. In *Proc CHI'97*, 303–310. ACM, 1997.
4. B. B. Bederson, J. D. Hollan, K. Perlin, J. Meyer, D. Bacon, and G. W. Furnas. Pad++: a zoomable graphical sketchpad for exploring alternate interface physics. *J. of Visual Languages and Computing*, 7(1):3–32, 1996.
5. R. Blanch, Y. Guiard, and M. Beaudouin-Lafon. Semantic pointing: improving target acquisition with control-display ratio adaptation. In *Proc CHI'04*, 519–526. ACM, 2004.
6. G. Casiez, D. Vogel, R. Balakrishnan, and A. Cockburn. The impact of control-display gain on user performance in pointing tasks. *Hum.-Comput. Interact.*, 23(3):215–250, 2008.
7. A. Cockburn and S. Brewster. Multimodal feedback for the acquisition of small targets. *Ergonomics*, 48(9):1129–1150, 2005.
8. A. Cockburn and P. Brock. Human on-line response to visual and motor target expansion. In *Proc GI'06*, 81–87, 2006.
9. A. Cockburn and A. Firth. Improving the acquisition of small targets. In *Proc. BHCI'03*, 181–196, 2003.
10. G. W. Furnas. Generalized fisheye views. *SIGCHI Bull.*, 17(4):16–23, 1986.
11. Y. Guiard. Disentangling relative from absolute amplitude in Fitts' law experiments. In *Proc. CHI'01*, 315–316. ACM Press, 2001.
12. Y. Guiard, M. Beaudouin-Lafon, and D. Mottet. Navigation as multiscale pointing: extending Fitts' model to very high precision tasks. In *Proc CHI'99*, 450–457. ACM, 1999.
13. Y. Guiard, R. Blanch, and M. Beaudouin-Lafon. Object pointing: a complement to bitmap pointing in GUIs. In *Proc GI'04*, 9–16, 2004.
14. C. Gutwin. Improving focus targeting in interactive fisheye views. In *Proc. CHI'02*, 267–274. ACM, 2002.
15. H. D. Jellinek and S. K. Card. Powermice and user performance. In *Proc CHI'90*, 213–220. ACM, 1990.
16. P. Kabbash and W. A. S. Buxton. The "prince" technique: Fitts' law and selection using area cursors. In *Proc CHI'95*, 273–279. ACM/Addison-Wesley, 1995.
17. R. L. Kline and E. P. Glinert. Improving GUI accessibility for people with low vision. In *Proc CHI'95*, 114–121. ACM Press/Addison-Wesley, 1995.
18. T. O. Kvålseth. An alternative to Fitts' law. *Bull. Psychonomic Soc.*, 16(5):371–373, 1980.
19. G. Langolf, D. Chaffin, and J. Foulke. An investigation of Fitts' law using a wide range of movement amplitudes. *J. Mot. Behav.*, 8:113–128, 1976.
20. I. S. MacKenzie. Fitts' Law as a research and design tool in human-computer interaction. *Hum.-Comput. Interact.*, 7:91–139, 1992.
21. P. Oel, P. Schmidt, and A. Schmitt. Time prediction of mouse-based cursor movements. In *Proc. BHCI-IHM '01*, 37–40, 2001.
22. G. Ramos, A. Cockburn, R. Balakrishnan, and M. Beaudouin-Lafon. Pointing lenses: facilitating stylus input through visual-and motor-space magnification. In *Proc CHI'07*, 757–766. ACM, 2007.
23. X. Ren and S. Moriya. Improving selection performance on pen-based systems: a study of pen-based interaction for selection tasks. *ACM Trans. Comput.-Hum. Interact.*, 7(3):384–416, 2000.
24. M. Schedlbauer. A survey of human cognitive and motor performance models. Technical report, 1997. http://teaching.cs.uml.edu/~heines/techrpts/Papers/TR2007-001.pdf.
25. A. Sears and B. Shneiderman. High precision touchscreens: design strategies and comparisons with a mouse. *Int. J. Man-Mach. Stud.*, 34(4):593–613, 1991.
26. R. W. Soukoreff and I. S. MacKenzie. Towards a standard for pointing device evaluation: Perspectives on 27 years of Fitts' law research in HCI. *Int. J. Hum.-Comput. Stud.*, 61(6):751–789, 2004.
27. G. P. Van Galen and W. P. De Jong. Fitts' law as the outcome of a dynamic noise filtering model of motor control. *Human Movement Science*, 14(4-5):539–572, 1995.
28. A. T. Welford, A. H. Norris, and N. W. Shock. Speed and accuracy of movement and their changes with age. *Acta. Psychol.*, 30:3–15, 1969.
29. A. Worden, N. Walker, K. Bharat, and S. Hudson. Making computers easier for older adults to use: area cursors and sticky icons. In *Proc CHI'97*, 266–271. ACM, 1997.