HAL
open science

On the Limits of the Human Motor Control Precision: the Search for a Device's Human Resolution

François Bérard, Guangyu Wang, Jeremy Cooperstock

To cite this version:

François Bérard, Guangyu Wang, Jeremy Cooperstock. On the Limits of the Human Motor Control Precision: the Search for a Device's Human Resolution. 13th International Conference on Human-Computer Interaction (INTERACT), Sep 2011, Lisbon, Portugal. pp.107-122, 10.1007/978-3-642-23771-3_10 . hal-00759964

HAL Id: hal-00759964

https://hal.science/hal-00759964

Submitted on 3 Dec 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

On the Limits of the Human Motor Control Precision: the Search for a Device's Human Resolution

François Bérard[1], Guangyu Wang[2], and Jeremy R. Cooperstock[2]

[1]University of Grenoble,
LIG, Grenoble-INP
BP 53 - 38041 Grenoble cedex 9, France
[2]McGill University,
Centre for Intelligent Machines,
Montréal, H3A 2A7, Canada
Francois.Berard@imag.fr, {gywang, jer}@cim.mcgill.ca

Abstract. Input devices are often evaluated in terms of their throughput, as measured by Fitts' Law, and by their resolution. However, little effort has been made to understand the limit of resolution that is controllable or "usable" by the human using the device. What is the point of a 5000 dpi computer mouse if the human motor control system is far from being able to achieve this level of precision?
This paper introduces the concept of a Device's Human Resolution (DHR): the smallest target size that users can acquire with an ordinary amount of effort using one particular device. We report on our attempt to find the DHR through a target acquisition experiment involving very small target sizes. Three devices were tested: a gaming mouse (5700 dpi), a PHANTOM (450 dpi), and a free-space device (85 dpi). The results indicate a decrease in target acquisition performance that is not predicted by Fitts' Law when target sizes become smaller than certain levels. In addition, the experiment shows that the actual achievable resolution varies greatly depending on the input device used, hence the need to include the "device" in the definition of DHR.

Keywords: input device, target acquisition, accuracy, device's human resolution, resolution

1 Introduction

Input devices share one fundamental characteristic: their respective limit of resolution (in short, their "resolution"). This is the smallest displacement that they can measure. The resolution is usually expressed as the number of position change reported by the device when displaced by one unit. For example, resolution expressed in "dots per inch", or *dpi*, gives the number of position changes reported for a displacement of one inch.

P. Campos et al. (Eds.): INTERACT 2011, Part II, LNCS 6947, pp. 107–122, 2011.
© IFIP International Federation for Information Processing 2011

Different devices have different resolutions. Computer mice typically have a resolution of 200 dpi. The PHANTOM Omni™ haptic device has a published resolution of above 450 dpi. Resolution of "gaming" computer mice, such as the Logitech G500, can reach as high as 5700 dpi. It is usually assumed that higher resolution is "better" since this provides a higher ratio of information per unit of displacement. For example, for the same displacement of one inch, a 200 dpi mouse can select at most 200 different positions (7.6 bits of information), whereas a 5700 dpi mouse can provide 12.5 bits of information.

It is not clear, however, how such resolution can benefit users if they are not actually able to control the device to select any of the 5700 possible positions within one inch. Given that human motor control capabilities are bounded, the fundamental problem becomes "How small can a target be, so that it remains selectable by a human, using the device, without an inordinate amount of effort?". Our claim is that, at some point, increasing the resolution of input devices is futile because the additional resolution is effectively unusable. We argue, therefore, that a device's *physical* resolution, itself, is fundamentally less relevant a characteristic for HCI than its *human-capable* resolution. Surprisingly, such a human-centered measure of resolution does not appear to have been studied explicitly in the literature.

In this paper, we define a *Device's Human Resolution* (DHR) as the smallest target size that a user can acquire, using the device, with an ordinary amount of effort, that is, matching Fitts' [10] prediction.

Studies have shown that human performance in device control depends on the muscle groups being activated [3,17]. As such, we expect that the achievable precision varies with the choice of device and mode of operation, as these collectively determine which muscle group is used. This motivates the definition of a *Device's* Human Resolution, rather than simply, the more general, "human resolution".

The definition of the DHR, and its measurement, should have significant value when choosing which input device to use for a particular task: for example, 3D positioning may require great accuracy for CAD applications, but less for entertainment applications. The DHR will be paramount in the first case, whereas a high number of degrees of freedom may be more relevant in the second case. Similarly, the DHR can help in the design of novel input devices or interaction techniques by making explicit the achievable resolution of the device *controlled by a human hand*. Touch displays, for example, are obvious candidates for DHR evaluation. The need to interact with an increasing amount of information on progressively smaller displays, e.g., the latest iPod Touch, which packs 960 x 640 pixels into a 3.5 inch diagonal screen, has been the source of a large body of work [2,4,5,14]. In particular, techniques have been designed to improve the precision of selection with fingers [5], for example using two fingers. Evaluation of the DHR of a touch surface would inform us as to whether the sensing device or human precision is the limiting factor. More practically, knowledge of this DHR would indicate the upper limit on useful sensing resolution for touch surfaces. Indeed, the most direct use of the DHR is in directing the allocation of resources where benefits are most likely: no human performance improvement is to be expected by increasing the physical resolution of a device if it already exceeds its DHR. Conversely, an improvement is likely if the device does not yet saturate its DHR.

In the work reported here, we present and justify our definition of a Device's Human Resolution (DHR). We then describe an approach designed to evaluate the DHR. Following this approach, we report on a study of three different devices: a mouse, a stylus attached to an articulated arm, and a free-space device. This study confirms the existence of a DHR, and shows that it varies greatly depending on the form factor of the device and its mode of operation. Moreover, the study provides estimates of the DHR for the three devices. We discuss the implications of this work, in particular to ongoing developments in sensing technologies for input devices.

2 Previous Work

Fitts' Law [10,12] is one of the most widely used tools to study human performance in controlling input devices. The law predicts that the mean time (*MT*) to acquire a target increases linearly with the index of difficulty (*ID*) of the target. In its "Shannon" formulation [12], *MT* is predicted as:

$$MT = a + b \log_2(A/W + 1) \quad \textbf{(1)}$$

where the log part of the equation is the index of difficulty (*ID*), *A* is the initial distance to target (or amplitude), and *W* is the target width. The *a* and *b* parameters are estimated empirically for each device being studied. The Fitts' Law prediction has been shown to be highly accurate, with a correlation coefficient of the linear regression coming very close to 1.0 (MacKenzie [12] offer a review of several such experiments, and describe the computation of an *effective width We* for an improved prediction). One important benefit of a Fitts' Law experiment is that it estimates the rate of information that can be transmitted to the computer with a given device (Zhai [16] discusses the use of $TP = \overline{ID} / \overline{MT}$ or $I_p = 1/b$ to estimate this rate). Fitts' Law experiments, however, are generally run under conditions where the law applies. In our case, we are investigating target acquisition at the limit of human motor control precision, i.e., in the range of target sizes where Fitts' Law may no longer apply.

Fitts' Law allows estimation of the throughput of the tested device. Although throughput is fundamental to performance, this measure alone provides a limited understanding of the suitability of a device to a particular task. To obtain a better understanding of performance differences between devices, MacKenzie et al. [13] introduced seven new measurements, such as the number of target re-entries, meant to analyze the *trajectory* of the pointer during target acquisition. However, these measurements do not directly address the human resolution limits of the device.

The effect on performance of the use of different muscle groups has been observed in several studies [3,17]. For example, for a six degrees of freedom docking task, a device held in the hand and rotated by the fingers is more efficient than the whole hand [17]. However, here again, the performance was only studied in terms of throughput, not in terms of achievable resolution (i.e., the DHR).

One of the very few examples studying the high accuracy attainable by the human hand is the work of Guiard et al. on "multi-scale pointing" [11]. Their experiment showed that users can efficiently acquire targets having very high Fitts' *ID*. The

experiment includes a target of a width of three points at a distance of 14,500 points in motor space units, resulting in a task *ID* greater than 12 bits. They observed that the limiting factor in this case was vision, rather than motor control: in the motor space, the target measures 0.06 mm and the amplitude is 289 mm. Such a small target would require a very high visual acuity if displayed directly. Instead, they introduced a multi-scale pointing technique in which pointer and targets were displayed at two different zoom factors. Participants first focused on the low-zoom display to bring the pointer into the neighborhood of the target, then switched their visual focus to the high-zoom rendering to complete the acquisition. In motor space, however, there was no discontinuity: participants were able to acquire this "high ID" target in a single gesture. This was possible because of the high resolution of the input device, a 1270 lpi[1] Wacom™ tablet, combined with the hand's ability to position the tip of a stylus efficiently in an area as small as 0.06 mm.

When using a stylus on the Wacom tablet, the task followed a Fitts' Law prediction: *IP* remained at approximately 4 bits/s both for an easy target (*ID* = 4 bits, single-scale display) and a difficult one (*ID* = 12 bits, multi-scale display). However, when the same task was performed using a "puck", similar to a computer mouse, *IP* dropped from 5.17 bits/s for the easy target to 3.09 bits/s for the difficult one. One key observation in this experiment is that the stylus and the puck positions were both sensed with the same device (the Wacom tablet). Thus, they shared the same physical sensing resolution and noise in measurement. This rules out the sensing mechanism as the source of performance difference, and points instead to the form factors of the devices themselves. Our interpretation is that the target size of 0.06 mm may have saturated human capacity when pointing with the puck, but not when pointing with the stylus. Thus, the stylus DHR exceeded 423 dpi (equivalent to 0.06 mm), but this is not the case for the puck DHR. This difference is to be expected since, as demonstrated by Balakrishnan et al. [3], the muscle group involved in control of a stylus yields greater efficiency than the muscle group (mostly the wrist) involved in control of the puck.

We note that Guiard et al. were not searching for the limits of human resolution, which could have been done by experimenting with smaller target widths, e.g., of 1 point. However, their results provide strong motivation for the work reported here.

While studying the effect of control/display (CD) gain on pointing performance, Casiez et al [9] acknowledged the existence of "limb precision problems" when using high CD gains, but unrelated to effects of device quantization. They noted "accuracy problems" with 2 mm targets and a CD gain of 12, thus estimating the limb precision at approximately 0.2 mm. However, this accuracy problem was only observed by an increased overshoot: participants were still able to acquire the targets. We are focusing on smaller target sizes that participants are not able to acquire in an ordinary amount of time.

Our work on the limit of human resolution has its origin in a previous study comparing devices for 3D object placement [6]. We found that 3D positioning was more efficient using a regular mouse than directly pointing to the target location in the air with a free-space device. In trying to explain this surprising result, we

[1] Wacom measures their device resolution in lines per inch, which has a slightly different interpretation from dots per inch.

hypothesized that the same task, which was possible using the mouse, became too difficult to perform with the free-space device due to a difference in the hand's precision when moving in the air, as opposed to sliding on a table. We went on to conduct an informal evaluation of the *static stability* of the hand with the free-space device. In comparison, the current work formally acknowledges the existence of a "Device's Human Resolution", provides a definition for it based on the *dynamic stability* of the hand, and attempts to measure it for three devices through an empirical study.

3 Defining the Device's Human Resolution

3.1 Device resolution

According to the The New Oxford American Dictionary, *resolution* is "The smallest interval measurable by a scientific (esp. optical) instrument" [1]. In the case of an input device, its resolution can be defined as the smallest displacement that will trigger a change in its reported position.

For some devices, such as an optomechanical mouse, determining the resolution is easy. The measuring component itself, in this case, a wheel with holes, provides a discretization mechanism that only reports measurements when the device has been moved by an amount of at least its resolution. Many input devices, however, provide a *noisy* raw measurement in which the reported output varies even when the device itself is static. Such devices, for example, optical trackers or accelerometers, are not *statically stable*. In this case, the concept of resolution, as defined above, cannot be applied directly. Instead, the raw measurement is usually filtered until variations in measurement of a static object are not reported.

One way to stabilize noisy device outputs is to estimate the standard deviation of the raw measurements of the position of a static object. Then, a threshold is defined as a multiple of this standard deviation. Variations in measurements are reported only if they exceed the threshold. Assuming Gaussian noise and a threshold of four standard deviations, for example, ensures that no variation is reported for a static object on 99.9937% of the measurements. Since the device only reports an output when it is moved by more than the threshold amount, this value defines the actual device resolution. This is the approach we followed to stabilize the output of our optical tracker for the free-space device.

3.2 Device's Human Resolution

Static stability We initially considered a simple definition for the Device's Human Resolution (DHR), based on the *static stability* of users manipulating the device. This measure can be evaluated in a similar manner to the static stability of the device itself. Users are asked to hold the device as statically as possible for a given duration, and the positions reported by the device are recorded. The standard deviation of the recorded positions is computed and the static resolution of the hand can be defined as

four times the standard deviation. Note that this is only feasible if the recording device resolution exceeds that of the participants. Otherwise, another higher resolution *measuring* device must be used. For example, to determine the human resolution of a low-resolution inertial tracker, the inertial tracker could be augmented with optical markers and tracked by a high-resolution motion capture system.

However, a static stability definition of DHR is not ideal for two reasons. First, it would not be applicable to *self-stabilizing* devices such as a regular mouse, which are perfectly stable when users release their grasp, and hence, provide a null standard deviation in their recorded values. Second, this measure estimates the capacity of users to hold the device stable for a duration, a much less frequent and more demanding task than that of target acquisition. We therefore chose another measure for DHR, inspired by the *dynamic stability* of the hand.

Dynamic stability A dynamically stable system quickly recovers its steady state after being moved away from it. By extension, the dynamic stability of the hand, coupled with a device, can be measured as the time for the hand to bring the device within a specified range of a target position, starting at some initial distance from the target. In this case, the range is equivalent to the target width for a Fitts' acquisition task. Our assumption is that below some specific threshold of target width, we should observe a significant drop in the ability of the user to acquire the target quickly. This threshold is a good candidate for a measure of the DHR.

We thus define the DHR as *the smallest target size that a user can acquire with the device, given an ordinary amount of effort*.

Obviously, ability varies between users, so the DHR should be measured as an average over a population, much in the same way that the throughput of a device is computed.

3.3 On the DHR and the Control/Display gain

It could be argued that human pointing precision can be increased at will, and thus, arbitrarily small targets can be acquired comfortably, simply by lowering the control/display (C/D) gain of the input device. Indeed, the C/D gain is a useful software mechanism that allows the system to adapt the input capabilities of users and devices to the work area. However, this improvement in precision comes at the cost of longer acquisition of distant targets: lowering the gain trades off speed of motion for accuracy. Users have to make larger motions, and *clutch* more often, to reach targets across the work space. As shown by Casiez et al. [9], low C/D gains afford the lowest performance in task acquisition, either using constant gain or pointer acceleration.

Regardless, the DHR is a physical limit that exists only in motor space, not in display space. In particular, this is independent of any technique used to transfer the users' physical motion into movement of the pointer. We interpret the DHR as the quantity of information that the user can transmit *by unit of physical motion* to the system, with a particular input device. Clever transformations may be applied to this information (e.g., pointer acceleration or semantic pointing [7]), but these are irrelevant to the underlying notion of a *physical* human resolution.

4 Measuring the DHR

Assuming a fixed distance to target, an increase of effort is expected when acquiring targets of decreasing sizes. In ordinary circumstances, Fitts' Law tells us that this additional effort should be *proportional* to the variation of task difficulty, as defined by the Fitts' *ID*. Our approach to measuring the DHR is equivalent to finding where this increase is no longer proportional.

In other words, it would be pointless to analyze the difference of mean acquisition times when target size decreases as they are *expected* to increase. The *slope* of the line fitted to this data (the Fitts' regression line) should remain roughly constant, though, and this is why we chose to analyze the variation of this slope, as a function of target size.

In addition, we expect error rate to remain roughly constant before the DHR, but increase thereafter.

The experimental task we use is very similar to a Fitts' task, with the following variations:

- *Inclusion of difficult targets:* Targets are varied from relatively large (easy) to small sizes that are suspected to be below the DHR (difficult).
- *Fixed amplitude:* Rather than vary the distance to target, we chose a reasonable fixed value for this parameter. The target must be sufficiently far away to require a significant displacement of the device, but close enough so that acquisition can be performed in a single gesture. In particular, we ensure that clutching is not required. In our pilot studies, we did not observe any effect of target distance on the DHR, but more formal studies would be required to provide strong recommendations regarding the optimal distance.
- *Requested target width instead of effective target width:* using the effective target width [12] in the Fitts' Law linear regression handles the speed-accuracy tradeoff problem and yields a better fit, and thus a better estimate of the device's throughput. However, the effective width method cannot be applied in our context: it would hide the difficulty for participants to acquire targets whose size is below the DHR, by counting failures as successes on larger targets. Our goal is to witness a drop in performance (either an increase of time or error rate) when participants actually try to acquire targets of the *requested* sizes. Besides, the lack of optimized fit of the regression line is not a problem in our context: in contrast with the throughput, the DHR is not a parameter of the regression line.
- *Validation of acquisition:* Since we are focusing on the capability of the (dominant) human hand to position a device in a small area, we want to avoid any parasitic motion induced by a validation mechanism. We thus allocate the validation action, e.g. clicking, to the participant's other (non-dominant) hand.
- *Parametrization in motor space:* Fitts' experiments have been used to evaluate not only devices, but also interaction techniques, such as semantic pointing [7]. Depending on the study, the two parameters of task difficulty, namely, target size and amplitude, can be defined in motor space, display space, or both. When measuring DHR, these parameters are always defined in motor space.

5 Experimental Verification

We conducted an experiment, following the approach presented above, to verify whether a discontinuity does indeed appear in user performance as target size decreases to very small levels, and if so, to estimate these levels.

5.1 Input devices

Three input devices were tested: a mouse, a free-space device, and a six degrees of freedom (DoF) stylus. All three devices were tested in conjunction with a keyboard, used by the non-dominant hand, for validation of target acquisition.

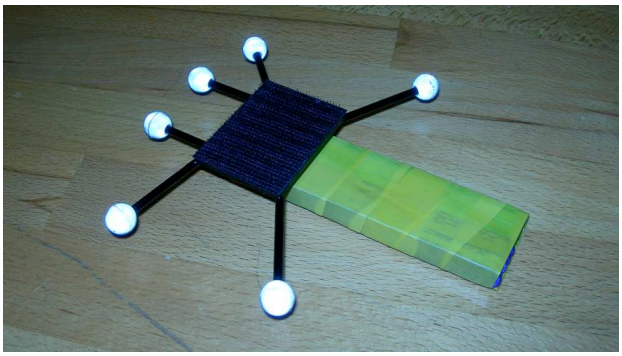


Fig. 1. The tracked free-space device.

- mouse: a Logitech G500™ gaming mouse, with physical controls allowing its resolution to be configured from 200 to 5700 dpi. For our experiments, we used the mouse at its maximum resolution of 5700 dpi, placing it on a typical wooden desk surface as a reasonable compromise between sliding ability and friction. The mouse system driver was bypassed: raw data was used in order to eliminate any mouse acceleration function implemented in the system. We performed a quick verification of the advertised resolution by observing the output of the mouse when translated by a distance of 10 cm, as measured by ruler. We verified that the final output difference corresponded to expectations, and that all in-between values were outputted.
- free-space device: a custom-built, comfortable, graspable rigid object, with six attached reflective markers, as shown in Figure 1. Its position was obtained by a carefully calibrated motion capture system (a six-camera NaturalPoint Optitrack™), placed as close as possible to the workspace in order to maximize device resolution. This resolution was measured with the procedure explained in the above section "Device resolution": static stability was evaluated in different locations of the workspace at a maximal standard deviation of $stddev_{max} = 0.00335mm$. The "tick size" (minimal displacement before motion is reported, i.e., the resolution) was thus set at $4 * stddev_{max} = 0.0134mm$. However, pilot experiments indicated that the human hand stability in free space was

nowhere close to this resolution, making it exceedingly difficult to acquire a target of one tick width. We therefore filtered the motion capture output further to a resolution of 0.3*mm*, equivalent to 85 dpi.

- stylus: the stylus attached to a Sensable PHANTOM Omni™ force feedback device. The PHANTOM was used only for its input capability: the force feedback mechanism was not activated and remained passive. This provided some amount of self-stabilization compared to a free-space device. The device has a resolution of 450 dpi that we checked with a similar procedure as with the mouse.

5.2 The task

A mouse is designed for 2D tasks, whereas the free-space device and stylus are better suited for 3D tasks. Our experimental task, however, was a 1D target acquisition: in measuring human resolution, we were looking for the smallest displacement controlable by hand. This displacement is measured as a 1D distance, even if performed in 2D space (for the mouse) or in 3D space (for the two other devices). As such, we believe that a 1D target acquisition with a constant gain of one is the most fundamental operation of any pointing device, regardless of interaction style. Moreover, this task presents the additional benefit of uniformity across devices.

Although the task was one-dimensional, we wanted to evaluate the DHR of each device in its natural operating space. Thus, we did not physically constrain any of the devices to performing 1D motion as this would interfere with the usual 2D operating space of a mouse or the 3D operating space of the stylus and the free-space device. Moreover, we did not suggest any particular posture be used with all devices: we expected participants to optimize their posture so that they felt most comfortable to achieve these very difficult target acquisitions. Indeed, for the 3D devices, we observed that participants tended to find a way to lock their forearm on the desk so that only the degrees of freedom of the wrist and fingers would be used.

Participants were asked to acquire targets of width 1, 2, 4, 8, 16, 24, and 32 ticks at a distance of 250 ticks, all measured in motor space. For the mouse and stylus, ticks correspond to the smallest displacement measurable by each device, or in other words, its resolution. For the free-space device, ticks were deliberately reduced in resolution, as described above. The corresponding Fitts' *ID* values, using Equation 1, were 3.14, 3.51, 4.06, 5.01, 5.99, 6.98, and 7.97. This range of Fitts' *ID* is typical of many Fitts' studies reported in the literature, hence a very good fit of the regression line to the experimental data would be expected, if not for the unusually small target sizes.

Graphical feedback was displayed on a 1920 x 1080 LCD monitor. As shown in Figure 2, a red pointer, a grey starting area, and a black target were all represented on a white background by vertical rectangles spanning the entire height of the display. The pointer and starting area widths were 1 and 40 ticks, respectively, and the target width was varied between 1 and 32 ticks. To avoid potential issues of visual acuity, we used a tick to pixel factor of 4, i.e., the 1 tick pointer was displayed as 4 pixels wide.

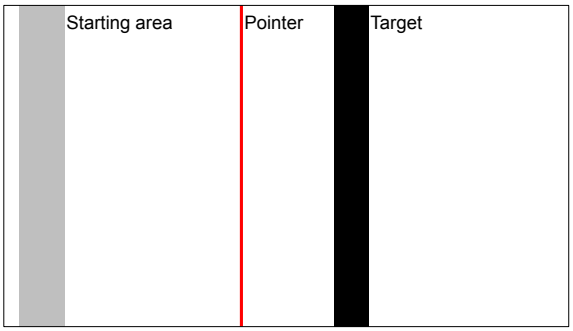
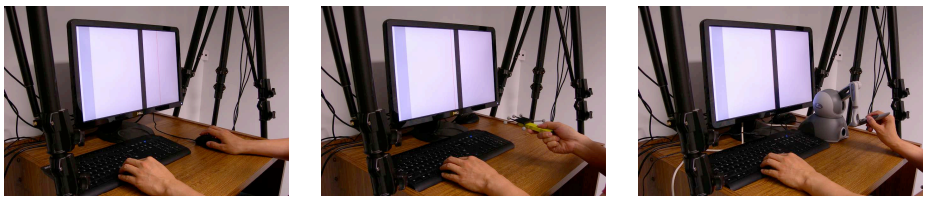


Fig. 2. The user interface of the experiment. Text labels have been added to the figure for clarity.

Participants had to bring the pointer to the starting area and validate by pressing the spacebar key with their non dominant hand. This caused the next target to be displayed, centered exactly 250 ticks from the current pointer position. To terminate a trial successfully, participants then had to move the pointer to the target area and validate again. As visual feedback, the pointer color changed to green when above the target. If the pointer was outside the target upon validation, the system emitted a beep and the user was offered another chance. Upon either successful acquisition or three successive failed attempts[2], the target disappeared. Participants would then begin the next trial by validating within the starting area. The experimental setup is shown in Figure 3.



(a) Mouse (b) Free-space device (c) Stylus

Fig. 3. The experimental setup.

5.3 Experiment design

Eighteen right-handed volunteers, ages ranging from 22 to 38, participated in the study. Each device was tested with six participants, five male and one female each. Participants ran the experiment on only one device in order to limit the effects of

[2] The decision to require three successive attempts in the case of failure was made to discourage participants from ignoring the acquisition of difficult targets. Attempts to rush through this and failing incurred more target selection effort. However, only the first attempts at acquisition were considered successful for the computation of error rate and mean acquisition time.

learning transfer. Each target size was tested for 20 acquisitions. The overall number of trials was thus:

3 devices x
6 participants x
7 target sizes x
20 repetitions
= 2520 target acquisitions

5.4 Results

Results of the experiment are provided in Table 1 and illustrated in Figure 4.

Table 1. Results of the experiment

Device	Target size (ticks)	Fitts'ID	Mean time (s.)	Mean failure rate (%)	Mean slope
Mouse	32	3.14	1.01	1.8	
	24	3.51	1.08	2.5	0.27
	16	4.06	1.24	0.8	0.28
	8	5.01	1.50	3.3	0.43
	4	5.99	2.07	1.7	0.77
	2	6.98	3.06	2.5	1.67
	1	7.97	5.51	2.6	
Free-space	32	3.14	1.03	2.5	
	24	3.51	1.07	3.3	0.26
	16	4.06	1.24	4.2	0.34
	8	5.01	1.70	12.9	0.48
	4	5.99	2.57	20.0	0.69
	2	6.98	3.87	40.8	0.96
	1	7.97	6.89	60.0	
6 DoF Stylus	32	3.14	0.84	0.0	
	24	3.51	0.91	0.0	0.22
	16	4.06	1.05	2.5	0.24
	8	5.01	1.33	4.2	0.33
	4	5.99	1.70	3.3	0.38
	2	6.98	2.32	10.0	0.72
	1	7.97	3.49	11.9	

In order to analyze the deviation of data from Fitts' prediction, we computed Fitts' regression lines on all subsets made of three successive *ID*s for every users (i.e., regressions were computed on sets of 3 x 20 trials minus failures, for each 6 participants x 3 devices). In the above referenced table and figure, slopes are reported for the Fitts' *ID* at the center of the 3- *ID* subset. Hence, there is no slope value for the first and last *ID*. If the mean acquisition time was following Fitts' Law, we would expect the slope to remain roughly constant. In our case, we expect the slope to increase significantly when the target size becomes too small.

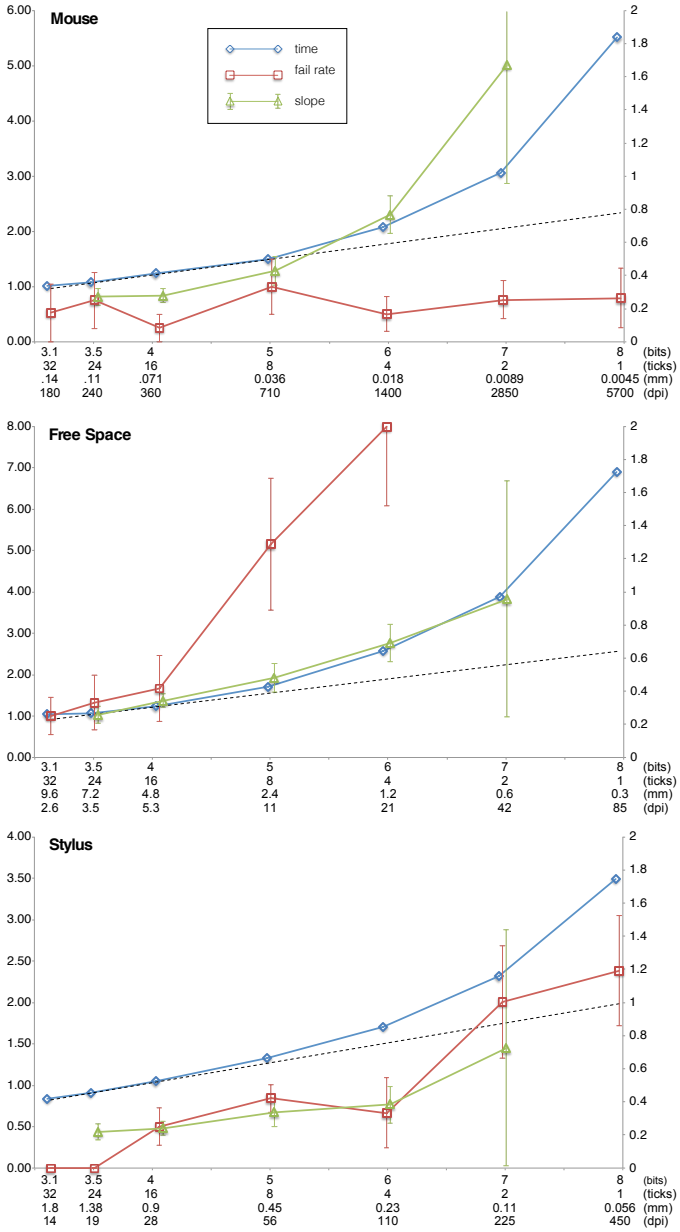


Fig. 4. Results of the experiment: mean time, mean failure rate, and mean slope of the Fitts' regression line (see text for details) as a function of Fitts' ID and target size for the three devices. Mean time (seconds) uses the left scale.
Failure rate (%) and slope (dimensionless) share the right scale, but failure rate was multiplied by 10 for improved legibility. Mean slopes and failure rates are shown with standard error bars. Fitts' regression line for $ID = 4$ is shown as a black dashed line.

Slope Three within-subjects ANOVA were computed, one for each device. Each revealed a significant effect of *ID* on the slope ($F(4,20) > 3.67$, $p < 0.05$).

Pair-wise comparison were computed for all couples of *ID* using one-sided paired t-tests, as we were expecting that slopes at higher *IDs* would be *greater* than the ones at lower *IDs* when passing the DHR. For the mouse, the pair-wise comparison found a significant difference only between the mean slope at $ID = 6$ on one side, and on the other side at $ID = 3.5$, 4.1, 5 ($t(5) < -4.1$, $p < 0.01$). For the free-space device, all difference between mean slopes were significant ($t(5) < -2$, $p < 0.05$), except between $ID = 6$ and $ID = 7$. For the stylus, the only significantly different mean slopes ($t(5) < -2.2$, $p < 0.05$) were at $ID = 3.5$ on one side, and on the other side at $ID = 5, 6, 7$.

Failure rate The within-subjects ANOVA for the mouse and the stylus did not show an effect of *ID* on the failure rate, but the ANOVA for the free-space device did show a strong effect ($F(6,28)=13.61$, $p < 0.001$).

Par-wise comparison were computed for the free-space device only. No significant increase was detected on all the couples made from the first three *IDs* ($ID = 3.1$, 3.5, 4), nor from ($ID = 5$, $ID = 6$). All other difference were significant: $t(5) < -2.1$, $p < 0.05$ for ($ID = 3.5$, $ID = 5$) and ($ID = 4$, $ID = 5$); $t(5) < -3.4$, $p < 0.01$ far all remaining couple of *ID*.

5.5 Analysis

As expected, target acquisition performance collapses as target size become smaller. This can be observed either on mean acquisition time, mean failure rate, or both.

Mouse In the case of the mouse, the slope remains roughly stable on the first three data points, but then it increases significantly at $ID = 6$. Even though the curve shows a very steep slope for the last data point ($ID = 7$), the lack of significant difference is due to the high variability of the data at this point. This variability can be explained by considering that participants were asked to perform target acquisitions that were far beyond their motor control capabilities.

The unexpected result was the mean failure rate which remained at less than 4% across all targets, with no significant effect of *ID*. This came as a surprise as we anticipated an increased error rate past the DHR. Participants managed to acquire the single-tick, 0.0045 mm target. The explanation lies in the self-stabilizing capability of the mouse: after several trials and errors, once the pointer was shown on target, participants can release the pressure on the mouse and validate with their non dominant hand. However, successful acquisition through this trial and error approach comes at the expense of a large increase in acquisition time, and higher variability. Clearly, at this size level, target acquisition no more appears as an automatized motor control.

If based on the last point that does not significantly diverge from Fitts' Law, the DHR of this mouse can be estimated at above target size 0.036 mm and below 0.018, or between 700 and 1400 dpi.

Free-space device The data for the free-space device is quite different. For this, the mean time exhibits a curved shape that does not clearly reveal a discontinuity. This translate to a steadily increasing slope, with all difference but the last being significant. Here again, the last estimation of the slope ($ID = 7$) shows a much larger variability than the previous ones.

Here, the failure rate is far more informative of performance sensitivity to target size. Failure rate remain at a usable level below 5% with no significant difference with the first three target size. At target size 2.4 mm, however, failure rate jumps to 12% (a significant increase from the three previous failure rates), and eventually, as high as 60% for the smallest target size.

Clearly, the participants' strategy was different from that for the mouse. The lack of self-stabilization made it very frustrating to try to hold the pointer inside the small targets for validation. When target size became too small, participants gave up and opted to fail rather than fight for success. This demonstrates the complementarity of error rate and acquisition time in determining the DHR. In the case of the free-space device, the bottleneck is error rate. The DHR appears to be bracketed in the range of 5-10 dpi.

Stylus The stylus exhibits the same trend as the mouse: a very good fit of the mean time to prediction on the first three points, and no significant difference detected on the slope for these points. At $ID = 5$, the slope is significantly higher than the one at $ID = 3.5$. However, at $ID = 5$ and $ID = 6$, the mean time is less than 13% above the prediction from previous points, and the error rate remains low below 5%. The next smaller target size is more problematic with a 10% error rate. We thus estimate the stylus DHR in the range of 100-200 dpi.

6 Discussion

One of the unexpected discoveries of our experiment is the benefit of self-stabilizing devices for very accurate target acquisition. The property of such devices that they come to perfect rest when users release control allows for a trial and error approach. Users will make an attempt, release their grasp, and check the outcome of the attempt, repeating as necessary until success. At every attempt, users get a chance to rest. This is in contrast with non-self-stabilizing devices, where users have to maintain firm control of the device to ensure that it remains on target at the moment of validation. Even though we did not try to control this parameter, we realized a posteriori that our three devices cover a spectrum from completely self-stabilizing (the mouse) to not self-stabilizing at all (the free-space device), and the stylus providing a middle-ground by way of its mechanical arm, which provides some amount of self-stabilization. This could provide an explanation to the different trends observed in the data: failure rate remaining low for the mouse, but shooting up for the free-space device, and a middle-ground observed with the stylus.

While we were looking for a clear threshold in performance, what we observed was rather a gradual degradation. This is not surprising considering the fact than we are estimating a *human* parameter, and that the estimation comes from a population of

variable individuals. Similarly to the processors' cycle times and the short-term memory capacity in the Model Human Processor [8], a range is a better depiction than a single value for these parameters. Still, we were able to witness, in the case of the mouse and stylus, an initial alignment with the mean time prediction followed by a significant departure from the prediction. In the case of the free-space device, it is the error rate that revealed the collapse in performance.

The most direct application of the DHR is a better allocation of the efforts to improve physical resolution of devices. In the case of the mouse, we found the DHR in the range of 700-1400 dpi. Even using the higher boundary, this raises the question of whether, for non-experimental purposes, there is any benefit beyond marketing for the 5700 dpi mouse we used. The results of our experiments suggest that further efforts to increase the dpi of the mouse would be wasteful.

There are, however, new types of devices for which the DHR can be used to define how much improvement in resolution is required for optimal tracking. Mobile devices, for example, are typically equipped with accelerometers, allowing them to be used in new forms of gestural interaction. The peep-hole display [15], for example, is a clever technique to allow efficient navigation in a large virtual document with a small screen device. This interaction technique requires sensing of the device's position in space, using its translation to navigate the document. This is very similar to the free-space interaction we evaluated, for which we determined that the DHR was in the range of 5-10 dpi. Thus, the accelerometers or other sensors used to track the device position for the peep-hole display would offer no improvement in control with a sensing resolution notably larger than 10 dpi.

Our work also confirms what was hinted at by Guiard et al. [11] in their work on multi-scale pointing: there is room for interaction techniques that exploit the accuracy of the hand when used in conjunction with appropriate devices, e.g., the high resolution mouse used in our study or the stylus used in Guiard et al.'s study. Their experiment included a 12.2 bit target acquisition task with a target size of 0.06 mm. Our experiment shows that target size can be reduced as far down as 0.036 mm (a 13 bit target using the same 289 mm amplitude) while maintaining good performance, i.e., in line with the Fitts' Law prediction.

We believe that the DHR provides an important measure of the information capacity of devices. It was shown that our stylus can be used efficiently to select targets of size down to 0.5 mm, whereas for a free-space device, the limit was 5 mm. Thus, for an equivalent operating range, the stylus can be used to select from ten times as many potential targets (in one dimension). However, the operating range of the stylus attached to the PHANTOM™ device is much smaller than the operating range of a free-space device (i.e. the space comfortably reachable around the user). Hence, the operating range of a device should also be taken into account when discussing the information capacity of the device. DHR and operating range define some form of *static* ability of a device: how many different targets can be fit inside the device's operating range. This should be studied in relation with the *dynamic* ability of the device: how efficiently can these targets be selected, i.e., the Fitts' throughput.

7 Conclusion

While it could be anticipated that target acquisition performance collapses when reaching a human limit of accuracy, we are not aware of any work specifically studying this phenomenon and addressing the evaluation of this limit. Our experiment highlights the fact that there is a significant drop in human performance when the target size decreases below a measurable threshold. We call this threshold the Device's Human Resolution (DHR). The DHR is highly dependent on the form factor of the device and its mode of operation: our experiment revealed DHRs spanning three orders of magnitude. The DHR of the mouse was in the order of thousands of dpi, the 3 DoF stylus in the hundreds, and the free-space DHR in the tens of dpi.

This work was a first attempt to demonstrate the existence of a Device's Human Resolution and at evaluating it for some specific devices. The DHR values that we estimated, however, were only one-dimensional. Further studies will be required to evaluate the variation of the DHR in higher dimensions, and evaluate the DHR for other devices. We hope that our initial effort in this regard inspires ongoing research by others to further explore these issues.

References

1. F. R. Abate and E. Jewell, editors. The new Oxford American dictionary, second edition. Oxford University Press, New York, USA, 2005.
2. P.-A. Albinsson and S. Zhai. High precision touch screen interaction. In ACM Conference on Human Factors in Computing Systems (CHI), pages 105–112. ACM, 2003.
3. R. Balakrishnan and I. S. MacKenzie. Performance differences in the fingers, wrist, and forearm in computer input control. In ACM Conference on Human Factors in Computing Systems (CHI), pages 303–310. ACM, 1997.
4. P.Baudisch and G.Chu. Back-of-device interaction allows creating very small touch devices. In ACM Conference on Human Factors in Computing Systems (CHI), pages 1923–1932. ACM, 2009.
5. H. Benko, A. D. Wilson, and P. Baudisch. Precise selection techniques for multi-touch screens. In ACM Conference on Human Factors in Computing Systems (CHI), pages 1263–1272. ACM, 2006.
6. F. Bérard, J. Ip, M. Benovoy, D. El-Shimy, J.R. Blum, and J.R. Cooperstock. Did "minority report" get it wrong? superiority of the mouse over 3d input devices in a 3d placement task. In IFIP Conference on Human-Computer Interaction, INTERACT, volume 5727 of Lecture Notes in Computer Science, chapter 45, pages 400–414. Springer, 2009.
7. R. Blanch, Y. Guiard, and M. Beaudouin-Lafon. Semantic pointing: improving target acquisition with control-display ratio adaptation. In ACM Conference on Human Factors in Computing Systems (CHI), pages 519–526. ACM, 2004.
8. S. Card, T. P. Moran, and A. Newell. The Psychology of Human-Computer Interaction. Lawrence Erlbaum Associates, 1983.
9. G. Casiez, D. Vogel, R. Balakrishnan, and A. Cockburn. The impact of control-display gain on user performance in pointing tasks. Human-Computer Interaction, 23(3):215–250, 2008.
10. P.M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47(6):381–391, 1954.

11. Y. Guiard, M. Beaudouin-Lafon, and D. Mottet. Navigation as multiscale pointing: extending fitts' model to very high precision tasks. In ACM Conference on Human Factors in Computing Systems (CHI), pages 450–457. ACM, 1999.
12. I. S. MacKenzie. Fitts' law as a research and design tool in human-computer interaction. Human-Computer Interaction, 7:91–139, 1992.
13. I. S. MacKenzie, T. Kauppinen, and M. Silfverberg. Accuracy measures for evaluating computer pointing devices. In ACM Conference on Human Factors in Computing Systems (CHI), pages 9–16. ACM, 2001.
14. D. Vogel and P. Baudisch. Shift: a technique for operating pen-based interfaces using touch. In ACM Conference on Human Factors in Computing Systems (CHI), pages 657–666. ACM, 2007.
15. K.-P. Yee. Peephole displays: pen interaction on spatially aware handheld computers. In ACM Conference on Human Factors in Computing Systems (CHI), pages 1–8. ACM, 2003.
16. S. Zhai. Characterizing computer input with fitts' law parameters–the information and non-information aspects of pointing. International Journal of Human-Computer Studies, 61(6):791 – 809, 2004. Fitts' law 50 years later: applications and contributions from human-computer interaction.
17. S. Zhai, P. Milgram, and W. Buxton. The influence of muscle groups on performance of multiple degree-of-freedom input. In ACM Conference on Human Factors in Computing Systems (CHI), pages 308–315. ACM, 1996.