

Positional Variance Profiles (PVPs): A New Take on the Speed-Accuracy Tradeoff

ANONYMOUS AUTHOR(S)

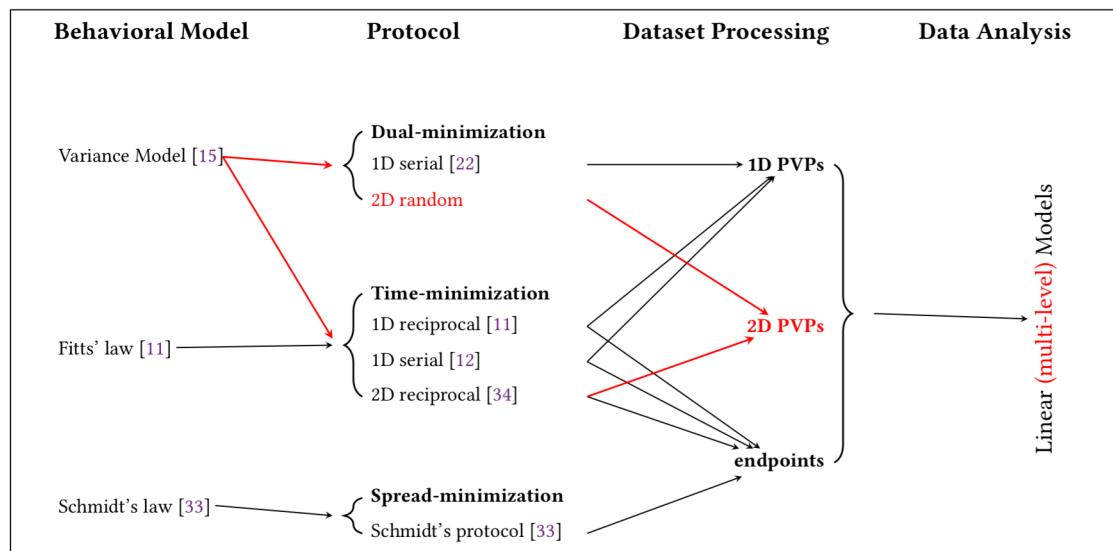


Fig. 1. Design space for experiments studying the speed-accuracy tradeoff: A behavior model guides the protocol of an experiment to evaluate input performance, as well as the ensuing data analysis. In this work (red), we transpose a behavior model of computational neuroscience [15] called the Variance Model, to propose a novel 2D random *dual-minimization* protocol, and the 2D PVP method which extracts features from a set of whole trajectories. This provides an alternative to classical Fitts' law experiments, which use *time-minimization* protocols, and endpoint information only.

Fitts' law is a behavioral model, used to design protocols and analyze data from pointing experiments. These are usually conducted in HCI to evaluate input performance. Recently, an alternative method to characterize input performance, called the method of PVPs, was proposed in one-dimension, based on 1) a dual-minimization protocol, and 2) an analysis of the variability of entire trajectories. After presenting the method of PVPs, we extend it to two-dimensions (2D), and show how it explains Fitts' law and how it predicts its parameters. We also present the results of a controlled experiment where 2D-PVPs are fitted to empirical data from two protocols, the classical protocol and the dual-minimization protocol, to validate the method. We contrast the comparison of three input devices (mouse, touchpad, controller) via a Fitts' law experiment and the method of PVPs, showing that current protocols can be simplified while providing more informative output.

CCS Concepts: • Human-centered computing → HCI theory, concepts and models.

Additional Key Words and Phrases: Fitts' law, PVP, pointing, evaluation

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Association for Computing Machinery.

Manuscript submitted to ACM

53 **ACM Reference Format:**

54 Anonymous Author(s). 2018. Positional Variance Profiles (PVPs): A New Take on the Speed-Accuracy Tradeoff. In . ACM, New York,
 55 NY, USA, 24 pages. <https://doi.org/XXXXXXX.XXXXXXX>

56 **1 INTRODUCTION**

57 In human motor control, an increase in speed is invariably made at the cost of a decrease in accuracy – a speed-accuracy
 58 tradeoff that has been captured by several behavioral models. One of the most popular in HCI is Fitts' law; it predicts
 59 the time MT it takes a participant to reach a target of size W, located at a distance D as fast as possible.

$$60 \quad 61 \quad 62 \quad 63 \quad 64 \quad 65 \quad 66 \quad 67 \quad 68 \quad 69 \quad 70 \quad 71 \quad 72 \quad 73 \quad 74 \quad 75 \quad 76 \quad 77 \quad 78 \quad 79 \quad 80 \quad 81 \quad 82 \quad 83 \quad 84 \quad 85 \quad 86 \quad 87 \quad 88 \quad 89 \quad 90 \quad 91 \quad 92 \quad 93 \quad 94 \quad 95 \quad 96 \quad 97 \quad 98 \quad 99 \quad 100 \quad 101 \quad 102 \quad 103 \quad 104 \quad MT = a + b \log_2 \left(1 + \frac{D}{W} \right) = a + b ID. \quad (1)$$

70 Ever since the seminal study by Card *et al.* [4] which used Fitts' law as a basis to evaluate input performance, it has
 71 become a staple of HCI research: in so-called Fitts' law experiments, it serves as a reference behavioral model, and is
 72 used to design experimental protocols and fit data when researchers evaluate input performance [17, 34] (see Fig. 1)

73 However:

- 74 (1) There is little knowledge on why Fitts' law is a good description of human behavior. The theoretical basis of the
 75 law is not well understood: the experimental protocol and data analysis are based on Fitts' loose analogy between
 76 an aiming task and a noisy channel [11, 34]. That analogy however has been critiqued multiple times [7, 17, 28],
 77 which makes some Fitts' law practices questionable.
- 78 (2) Fitts' law experiments use time-minimization protocols, which poorly control the accuracy of movements [22, 23]:
 79 movements may end up outside the target –usually for small targets– or be clustered within a small part of the
 80 target –usually for large targets [11, 34].¹ In practice, the independent variable W is often corrected *post-hoc*
 81 to reflect the actual spread of endpoints via the so-called effective version of Fitts' law [34], a practice that is
 82 contrary to the purpose of a controlled experiment –the independent variable should not be adjusted after the
 83 experiment– and has further been shown to be arbitrary [17].

84 As a result, many problems around Fitts' law experiments remain unresolved: Several versions of ID in Eq. (1)
 85 exist [17, 31, 34]. Which one should we use and why? What is the interpretation of the intercept *a* in Eq. (1)? Should it
 86 be equal to 0 [21, 34]? Interpreting the intercept is important because it conditions whether an assessment of input
 87 performance should consider the intercept or not, and whether a non-zero value is indicative of a poor experiment
 88 or not as suggested by Soukoreff and Mackenzie [34]. Can throughput, a one-dimensional score that estimates the
 89 “performance” of a device or technique, be unequivocally defined, and what then is its interpretation [17, 34, 38]? Should
 90 the experimenter encourage the participants to aim for a 4% miss rate as often encouraged [17, 34], or should the
 91 experimenter strive for a null miss rate?

92 In this work, we build on an information-theoretic transmission model with feedback from the computational
 93 neurosciences, which we call the *Variance Model* [15], and the corresponding data analysis method called the *method of*
 94 *PVPs* to provide an alternative to Fitts' law experiments for input performance evaluations. The method of PVPs is a
 95 data aggregation method, where a time series of the variance of a *set* of trajectories is computed to obtain a variance
 96 profile. PVPs display two phases, of which the second is described by the *theoretical* variance model, and the first was
 97 *empirically* characterized [15].

100 ¹Fitts himself was aware of this. His 1954 paper [11] features 3 experiments: The first one is the one who was popularized, but the latter two featured
 101 variations where the width was properly controlled, by means of pegs and discs.

The method of PVPs can be applied to several protocols, including ones where there is no pre-specified width and is also more expressive than Fitts' law, because it characterizes whole trajectories, and not just endpoints. A further advantage of PVP is that Fitts' law can essentially be derived from them. Hence, by transposing the method of PVPs in HCI, we also alleviate some drawbacks associated with Fitts' law:

- Fitts' law is expressed as the interaction between the two phases of the PVP. As a result, Fitts' law parameters can be predicted and can be interpreted within the information-transmission scheme. This provides new arguments to the aforementioned unresolved problems with Fitts' law experiments.
- We propose a new protocol, amenable to the method of PVPs, which does not try to enforce endpoint width constraints to generate data, thereby removing the need for a *post-hoc* correction. The protocol is also simpler to conduct, since it does not require crossing D and W.

Currently, however, the method of PVPs can not be applied off-the-shelf to most HCI evaluation opportunities: the method was described for 1D data only, whereas most situations in HCI feature two-dimensional pointing. A similar problem affected Fitts' law experiments, which was also initially described in 1D only, and which attracted significant effort [1, 19, 26, 29, 36]. Thus, in this work we extend 2D-PVPs and in doing so make the connection to existing literature on 2D Fitts' law. Further, the effect of D on PVP features was weakly assessed (only 2 levels) in the original work on the method of PVPs [15] —their focus being on the second phase— but this effect is crucial towards predicting Fitts' law. Therefore, we conducted an empirical study of the effect of distance on PVP features. We also compare a Fitts' law experiment featuring three devices (mouse, touchpad, controller) with a PVP-based experiment featuring the new protocol. In summary, our contributions are:

- transposing a theory and method from the neurosciences in HCI that describes the evolution of a set of pointing trajectories;
- proposing a new experimental protocol for evaluating input performance, that is simpler than the one used in Fitts' law experiments;
- extending the method of PVPs to the 2D case, proposing 3 different profiles;
- estimating the effect of D on PVP features, validating the 2D-PVPs;
- comparing a Fitts' law experiment with our new experiment and method for the comparison of three devices (mouse, touchpad, controller);
- providing new arguments to open problems surrounding Fitts' law experiments.

2 BACKGROUND

2.1 Fitts' law and Fitts' law experiments

The original formulation of Fitts' law, called its nominal version [39] is given Eq. (1). Fitts' law in its *effective* or *corrected* version [5, 28] reads

$$MT = a + b \log_2 \left(1 + \frac{D}{4.133\sigma} \right) = a + b \text{ID}_e, \quad (2)$$

where σ is the standard deviation of endpoints measured for each condition. This is a popular amendment to Fitts' original formula, and is meant to account for divergent participant speed-accuracy strategies [5, 17, 34], as some participants emphasize speed but miss the target more often in the process. As explained in the introduction and by Guiard [22, 23], such a post-hoc adjustment is not in line with controlled experiments, where it is assumed that the independent variable is controlled by the experimenter, and not up to the participant or the environment.

When Fitts' law is used to assess the performance of a participant or a device –*i.e.*, a Fitts' law experiment is performed— one usually runs a controlled experiment where the factors D and W are fully crossed. The law's parameters, a and b in Eq. (1), are then empirically estimated from endpoint information (time and position), see Fig. 1. Overall, lower values of a (intercept) and b (slope) indicate a better efficiency since they lead to shorter movement times.

A one dimensional score known as the *ISO-throughput* [34]

$$TP = \frac{1}{N} \sum_{n=1}^N \frac{ID_{e,n}}{MT_n}, \quad (3)$$

may be used to condense all measured input performance into a single score, where N is the number of different experimental conditions, and where the data is aggregated by experimental condition. That metric however has been criticized on practical and theoretical grounds [17, 38].

2.2 Evaluating Input Performance: Protocols

Fitts' law experiments are part of a wider family of controlled studies that are used to investigate the speed-accuracy tradeoff. These experiments may use different protocols to produce data, of which Guiard [23] distinguishes three classes (see Fig. 1):

- time-minimization protocols, where participants are instructed to hit targets as fast as possible. Most protocols used in Fitts' law experiments, such as Fitts' reciprocal paradigm [11], Fitts' serial paradigm [12], and ISO's multi-directional tapping task [34] are time-minimization protocols.
- spread-minimization protocols, where participants are instructed to hit targets under a movement time constraint while minimizing the spread of endpoints. One well known spread-minimization protocol is due to Schmidt [33], and is associated with an empirical law named as Schmidt's law, but it has not been very popular in HCI.
- dual-minimization protocols, where participants should minimize both pointing time and spread of endpoints. For these protocols, the experimenter may direct the participant strategy (*e.g.*, emphasize speed over accuracy). Guiard and colleagues [22, 23] introduced a 1D serial dual-minimization protocol.

Contrary to Fitts' law, which requires a target width to be predetermined, and thus a time-minimization protocol to be evaluated, the method of PVPs can be applied irrespective of the protocol, since it is based only on computing a time series of the variance of a set of trajectories. In the original work on PVPs [15] the method of PVPs was tested on a time-minimization paradigm, and a dual minimization paradigm, with similar outcomes (see PROPERTIES in the next section).

3 METHOD OF PVPs

In this section, we review how Positional Variance Profiles (PVPs) are constructed, how they are fitted to extract key features, and their connection to an information-transmission scheme [15] that explains human behavior as the result of optimal transmission over a noisy Gaussian channel with feedback information.

3.1 Constructing PVPs

A trajectory is a time series of coordinates *e.g.*, mouse cursor or finger-tip positions, for a single movement. PVPs are computed for all trajectories of a given block by:

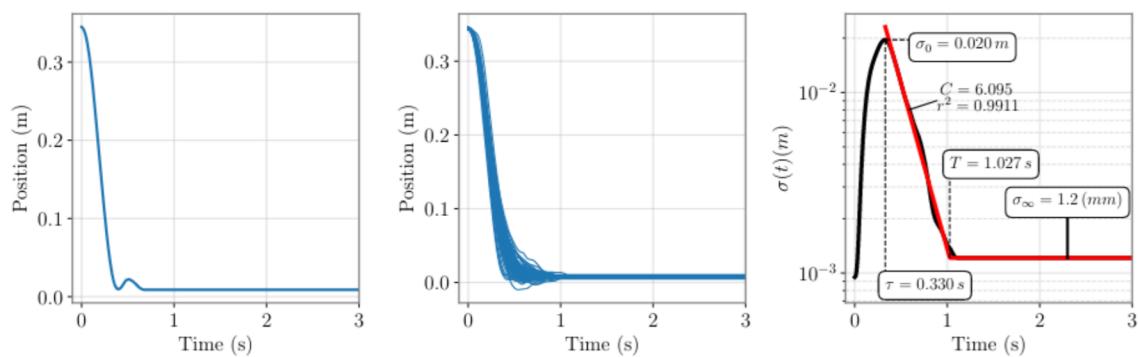


Fig. 2. PVP construction, illustrated on trajectories from a block of the Pointing Dynamics Dataset [30] ($P_1, D = 0.353, ID = 6$). Left panel: An individual trajectory, with position plotted against time. Middle panel: A set of pooled, synchronized trajectories, with position plotted against time. Right panel: PVP of the set of movements used in the middle panel, where standard deviation of the position plotted against time is plotted in black in a lin-log scale. The spline fit is displayed in red, as well as meaningful features of that fit: τ the duration of the first phase, σ_0 the initial variance at the start of the second phase, σ_∞ the end variance at the end of the second phase, C the estimated rate of the second phase, and T , the effective movement time.

- (1) Extending each trajectory so that all movements have the same duration, by *padding* the final value of position (to, say, 3 s), and resampling them to have equally spaced samples (Left panel in Fig. 2). By extending, we don't mean to warp the trajectory as in e.g., Dynamic Time Warping [32]; instead we do as if the recording lasted a little longer while movements had already ended.
- (2) Synchronizing all trajectories, using their starting time as the new time origin (Middle panel in Fig. 2).
- (3) Computing the standard deviation² $\sigma(t)$ of the position signal over time (Right panel, black curve in Fig. 2). The profile is best represented in a lin-log scale (see Property 6 below).

3.2 Properties of PVPs

PROPERTY 1. *PVPs can be decomposed into three phases: A first phase where positional variance increases rapidly, a second phase where it decreases steadily, and a third phase where positional variance is constant.*

Note that while [15] discuss a two-phase model, PVPs actually have three phases: this is because the third phase is due to how trajectories are extended. It does not have any meaning beyond the fact it is needed to fit the PVPs correctly.

The features of the first phase are (see Right panel of Fig. 2)

- τ , the duration of the first phase;
- D_τ , the distance covered during the first phase (not shown);
- σ_0 , the standard deviation of position at the end of the first phase (i.e., start of the second phase).

The first phase of the PVPs has not been described by a theoretical model, but was instead empirically characterized [15]

²While the name PVP suggests we deal with variance, the PVP actually is a time series of standard deviations. The difference is slight: in log-scale the two are related by a factor 2, that factor canceling out conveniently when using standard deviation in the theoretical model.

261 PROPERTY 2. *The duration of the first phase, τ , is linear in D :*

$$262 \quad 263 \quad 264 \quad 265 \quad 266 \quad 267 \quad 268 \quad 269 \quad 270 \quad 271 \quad 272 \quad 273 \quad 274 \quad 275 \quad 276 \quad 277 \quad 278 \quad 279 \quad 280 \quad 281 \quad 282 \quad 283 \quad 284 \quad 285 \quad 286 \quad 287 \quad 288 \quad 289 \quad 290 \quad 291 \quad 292 \quad 293 \quad 294 \quad 295 \quad 296 \quad 297 \quad 298 \quad 299 \quad 300 \quad 301 \quad 302 \quad 303 \quad 304 \quad 305 \quad 306 \quad 307 \quad 308 \quad 309 \quad 310 \quad 311 \quad 312$$

$$\tau = \tau_0 + \alpha D + \varepsilon \quad (4)$$

PROPERTY 3. *The distance covered during the first phase, D_τ , is proportional to D*

$$D_\tau = \beta D + \varepsilon' \quad (5)$$

PROPERTY 4. *The standard deviation at the end of the first phase (maximum standard deviation), σ_0 , is proportional to D*

$$\sigma_0 = \gamma D + \varepsilon'' \quad (6)$$

Here, ε , ε' and ε'' represent Gaussian deviations (error models). Deviations from the average trajectory in the second phase were also characterized:

PROPERTY 5. *Deviations from the average trajectory are Gaussian distributed.*

In [15] the link between the second phase and a biologically plausible information-transmission model is made. It is predicted that for a Gaussian input, a feedforward channel with Gaussian noise, and a feedback information channel with delay, the standard deviation of the output can decrease exponentially over time *at best*, at a rate C called the *capacity*, which is a participant-dependent constant, independent of task parameters such as D and W .

PROPERTY 6. *During the second phase, the PVP can decrease at best at an exponential rate C that is independent of the task properties:*

$$\log \sigma(t + \Delta t) = \log \sigma(t) - C\Delta t \quad (7)$$

This property is the direct result of an optimization procedure that is agnostic of the task—if the controller (in this case, the human) behaves optimally, has feedback information of the location of the endpoint they are controlling and has their actions subjected to Gaussian noise, then this is the best they could achieve. Users could perform sub-optimally for many reasons e.g., poor feedback information, but could only do better if assumptions the model makes are violated e.g., if they had access to another source of information or if they were not subjected to Gaussian noise.

3.3 Fitting PVPs

PVPs are fitted based on PROPERTIES 1 and 6. The standard deviation profile is logarithmically transformed to simplify the estimation of C ; a linear spline is then fitted to the second and third phases. A least-squares method determines the knot of the spline (T in the right panel of Fig 2), as well as the slope of the first part of the spine (C in the right panel of Fig. 2); the slope of the third phase is assumed null, movements having terminated.

3.4 Link between PVPs and Fitts' law parameters

The empirical model of the first phase interacts with the variance decreasing model of the second phase, through σ_0 : the first phase will lead to more or less variability, that needs to be reduced during the second phase. The task's parameters influence the two phases: a larger D will lead to higher speeds early in the movement, and thus larger σ_0 [33]. A smaller W will lead to a smaller endpoint variance. Both cases will tend to increase the duration of the second phase, because variance is reduced at a constant rate.

313 *Effective Fitts' law.* Putting what we just described (see Eq. 27 in [15] for more details) we get:

$$314 \quad 315 \quad T = \tau + c(\gamma) + 1/C \log\left(\frac{D}{4.133\sigma_\infty}\right), \text{ or, in shorter form,} \quad (8)$$

$$316 \quad 317 \quad 318 \quad T = \tau' + 1/C \log\left(\frac{D}{4.133\sigma_\infty}\right), \quad (9)$$

319 where $c(\gamma)$ is a linear function of γ in Eq. (6).

320
321 *Nominal Fitts' law.* The nominal Fitts' law is recovered by computing the miss rate ε of a Gaussian distribution of
322 endpoints with standard deviation σ_∞ centered on a target of size W (see e.g. [17] for a detailed computation). This
323 naturally makes the inverse Gaussian error function $\text{erf}^{-1}(1 - \varepsilon)$ appear in the formula (see Eq. 29 in [15] for more
324 details):

$$325 \quad 326 \quad 327 \quad T = \tau + d(\gamma, \text{erf}^{-1}(1 - \varepsilon)) + 1/C \log\left(\frac{D}{W}\right), \quad (10)$$

$$328 \quad 329 \quad 330 \quad T = \tau'' + 1/C \log\left(\frac{D}{W}\right), \quad (11)$$

331 where d is a linear function of γ in Eq. (6) and the inverse Gaussian error function.

332
333 *Identification.* Comparing Eqs. (9) and (11) with Eqs. (2) and (1), it is straightforward to identify:

- 334 (1) T with MT, where T acts as a proxy for MT [15, Fig. 7],
- 335 (2) C with $1/b$, the inverse of the slope in Fitts' law,
- 336 (3) σ_∞ with the endpoint standard deviation σ ,
- 337 (4) τ' and τ'' with a , the intercept in Fitts' law.

338 Thus, Fitts' law parameters can be predicted by PVP features and interpreted within the information-transmission
339 model.

340 4 EXTENDING THE METHOD OF PVPS IN 2D

341 In most HCI applications, movements are performed in two dimensions e.g., moving a mouse cursor on a computer
342 screen, which means we ought to extend the method of PVP in 2D. Instead of proposing a multidimensional information-
343 transmission model, we look for a 1D proxy of the 2D spread of data. Note that 2D versions of Fitts' law were faced
344 with a similar problem of matching 2D data to a 1D model, but with endpoints only; we therefore discuss parallels
345 between our work and existent work in Fitts' law literature when we found them.³

346 4.1 Parametrization of movement

347 We introduce the following quantities to parametrize any trajectory (see Fig. 3):

- 348 • The start location $O = (x_0, y_0)$, defined by the cursor location at time $t = 0$;
- 349 • The target location $T = (x_T, y_T)$, which remains fixed throughout the movement;

350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
³A naive solution would consist in using the remaining euclidean distance to the target as proxy signal upon which to compute PVPs. However, this would necessarily produce an underestimate of the spread of data, since undershoots and overshoots would in that case not be distinguished, as recognized previously by Wobbrock *et al.* [36] in the case of endpoints. A simple workaround to differentiate undershoots from overshoots is to consider the distance from the starting point rather than the distance to the target. Although this solves the previous issue when the movement is always perfectly aligned with the start-target axis (like in the 1D case), it does not in all other cases.

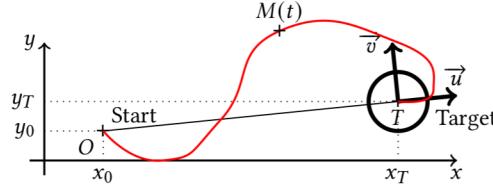


Fig. 3. Origin of movement (0), target (T), and current position ($M(t)$) of the cursor. The orthogonal basis used to define α and β is also given.

- The primary direction of movement defined as \overrightarrow{OT} is normalized to $\vec{u} = \overrightarrow{OT}/\|\overrightarrow{OT}\|_2$. This vector is dependent on the initial conditions only.
- The direction orthogonal to movement is defined such that $\{\vec{u}, \vec{v}\}$ form the standard orthonormal basis in the plane.

We define \overrightarrow{MT} as the vector relating the current position $M(t)$ to the (fixed) target position T at each time, and compute its projections $\alpha(t)$ and $\beta(t)$ on the basis $\{\vec{u}, \vec{v}\}$

$$\alpha(t) = \overrightarrow{MT} \cdot \vec{u}, \quad (12)$$

$$\beta(t) = \overrightarrow{MT} \cdot \vec{v}. \quad (13)$$

One recovers the 1D case with $\beta = 0$ and $x(t) = \alpha(t)$.

A set of trajectories from pilot data in the $\{\vec{u}, \vec{v}\}$ plane is illustrated Fig. 8 (Appendix). It shows that the $\beta(t)$ component of movement is roughly symmetric. This suggests that the experimental protocol used to acquire this pointing data (presented in the next section) successfully balanced the initial directions of movement.

4.2 Accounting for all directions of movement

The covariance matrix of a 2D vector captures the spread of that vector in all directions, it is the extension of variance in the multidimensional case. The vector $\overrightarrow{MT}(t)$ has covariance matrix $\Gamma(t)$

$$\Gamma(t) = \begin{bmatrix} \Gamma_{\alpha\alpha}(t) & \Gamma_{\alpha\beta}(t) \\ \Gamma_{\alpha\beta}(t) & \Gamma_{\beta\beta}(t) \end{bmatrix}. \quad (14)$$

Each term of that covariance matrix can be estimated by the sample covariance estimator:

$$\Gamma_{\alpha\beta} = \frac{1}{N-1} \sum_{i=1}^N (\alpha - \mu_\alpha)(\beta - \mu_\beta), \quad (15)$$

where μ_α (resp. μ_β) is the *sample* average for α (resp. β) and where i indexes trajectories. We have also dropped the time variable (t) for lighter notations here, and so will we for all similar quantities in the remainder of the paper.

4.3 Candidate 2D PVPs

To get a 1D metric, one still has to compress the covariance matrix to a single dimension term that effectively conveys the spatial spread of the data. We propose the three following PVPs, where $\tilde{\sigma}$ denotes the time series of the profile:

417 *PVP_α: Projection in the direction of movement.* Here, we simply take

$$418 \quad 419 \quad \tilde{\sigma}(t) = \sqrt{\Gamma_{\alpha\alpha}} \quad (16)$$

420 For Fitts' law experiments —who deal with endpoints only— the ISO standard computes σ by considering only the
 421 deviation in the direction of movement [34]. In our terms this amounts to considering only $\Gamma_{\alpha\alpha}$, with the square
 422 root being there to have a quantity with the same physical dimension as a standard deviation, hence our first PVP
 423 proposition.
 424

425 Note that the ISO standard provides no justification for this method, and even though it is known that errors
 426 orthogonal to the principal direction of motion are usually of smaller magnitude than in the direction of motion [24], it
 427 is unclear why deviations in the orthogonal direction should be ignored altogether [36]. One possible explanation could
 428 be the *minimum intervention principle* [35], which states that deviations that do not interfere with the task performance
 429 are not corrected. The next PVPs do account for deviations in all directions.
 430

432 *PVP_{Tr}: distance from centroïd.* Here, we take

$$434 \quad 435 \quad \tilde{\sigma}(t) = \sqrt{\text{Tr}(\Gamma)} = \sqrt{\Gamma_{\alpha\alpha} + \Gamma_{\beta\beta}}, \quad (17)$$

436 where we directly sum up the variances due to movement in the primary direction of movement and in the orthogonal
 437 direction in light of PVP_α's criticism. By plugging in the sample estimator Eq. 15, we recognize that
 438

$$439 \quad 440 \quad \tilde{\sigma}(t) = \sqrt{\frac{1}{N-1} \sum_{i=1}^N [(\alpha - \mu_\alpha)^2 + (\beta - \mu_\beta)^2]}, \quad (18)$$

442 which is what Wobbrock *et al.* [36] coined the 2D distance-from-centroïd formula in the case of endpoint deviation. The
 443 PVPs computed with this method are called PVP_{Tr}, the trace Tr of a matrix being the sum of its diagonal terms.
 444

445 *PVP_{det}: Determinant of the covariance matrix.* Here, we take:

$$447 \quad 448 \quad \tilde{\sigma}(t) = (\det(\Gamma))^{1/4} = \left(\Gamma_{\alpha\alpha} \Gamma_{\beta\beta} - \Gamma_{\alpha\beta}^2 \right)^{1/4}, \quad (19)$$

449 where $\det(\Gamma)$ is the determinant of the matrix Γ . Informally, the determinant of a matrix measures the area spanned by
 450 its basis vectors. Hence, the square root of the determinant is the characteristic length of a measure of the area spanned
 451 by its elements. Since the terms in the covariance matrix are homogeneous to variances, an extra square root is required
 452 to have a characteristic length homogeneous to a standard deviation, which explains the quartic root. This formulation
 453 is the only one of the three to account for cross-correlation terms ($\Gamma_{\alpha\beta}$).
 454

455 *Other PVPs.* We have selected and rationalized three PVP computations. There are opportunities for others, say,
 456 $\sqrt{\Gamma_{\alpha\beta}}$ or $\sqrt{\Gamma_{\alpha\alpha} + \Gamma_{\alpha\beta} + \Gamma_{\beta\beta}}$, but at this point, we are not interested in systematically investigating the “best” 2D-PVP.⁴
 457 However, one can easily use our definitions to introduce PVPs in any dimension (for example 3D): in N-dimensions, we
 458 obtain an $N \times N$ matrix, of which we select either the first term, the trace, or the determinant to obtain respectively
 459 PVP_α, PVP_{Tr} and PVP_{det}. Also note that because the determinant and trace of a matrix are rotation-invariant and
 460 because any shift in the origin of the basis is canceled out in the covariance matrix, one can consider any orthonormal
 461 basis in the plane on which to project \overrightarrow{MT} with the same results for PVP_{Tr} and PVP_{det} (but not PVP_α).
 462

463
 464
 465
 466 ⁴And as the results will show, there may be practically little interest in selecting either one of them in a “typical” situation. Differentiating between the
 467 profiles would likely be interesting in situations where variance is large and/or unusually high in the orthogonal direction.

469 4.4 Theoretical Comparison of 2D PVPs

470 Before an empirical study, we can compare PVPs on formal grounds. First, it is obvious from Eqs. (16) and (17) that
 471 PVP_{Tr} will always be larger than PVP _{α} .

472 Then, we found by observing the data presented in the next section, that the cross-covariance term $\Gamma_{\alpha\beta}$ is usually
 473 quite small compared to the variance terms. Making the approximation that $\Gamma_{\alpha\beta} \sim 0$, the determinant of the covariance
 474 matrix thus reduces to $\Gamma_{\alpha\alpha}\Gamma_{\beta\beta}$. When taking the logarithm of the quartic root and using the properties of the logarithm,
 475 one gets:

$$476 \quad 477 \quad \log(\Gamma_{\alpha\alpha}\Gamma_{\beta\beta})^{1/4} = \frac{1}{2} \left[\log \sqrt{\Gamma_{\alpha\alpha}} + \log \sqrt{\Gamma_{\beta\beta}} \right]. \quad (20)$$

478 Hence, PVP_{det} is obtained by averaging PVP _{α} and PVP _{β} in the log-lin plane. The well-known inequality of arithmetic
 479 and geometric means also implies that PVP_{det} is always smaller than PVP_{Tr}.⁵

480 If we further assume that $\Gamma_{\beta\beta} \leq \Gamma_{\alpha\alpha}$, which is reasonable since variability in the direction orthogonal to movement
 481 is usually lower than in the direction of movement, as explained previously, we have

$$482 \quad 483 \quad \text{PVP}_{\text{det}} \leq \text{PVP}_{\alpha} \leq \text{PVP}_{\text{Tr}}. \quad (21)$$

484 The actual magnitude of these differences is entirely due to the difference in magnitude between $\Gamma_{\alpha\alpha}$ and $\Gamma_{\beta\beta}$. We
 485 hypothesize that the difference in Γ is in part related to the choice of experimental protocol, especially whether the
 486 direction of approaches are balanced or not. The inequality is illustrated in Fig. 9 (Appendix).

487 5 EMPIRICAL STUDY: METHOD

488 We report on a controlled experiment to validate 2D-PVPs, where we investigate whether:

- 489 (1) Results from the 2D-PVP match those of the 1D-PVP in Gori and Rioul's work [15]. In particular, we hypothesize
 490 that
- 491 • The 2D-PVP profiles display three phases: A first phase where the PVP increases rapidly, followed by a
 492 second phase where the PVP decreases at an exponential rate, and a third phase where the PVP is constant.
 - 493 • Distance significantly affects τ , D_τ , and σ_0 , but not C , and with considerable effect sizes.
- 494 (2) Comparisons performed with a traditional Fitts' law experiment are equivalent to comparisons performed with
 495 a new protocol analyzed with 2D PVPs. We namely investigate whether, in a classical protocol, linear regression
 496 parameters can be predicted with features of the 2D-PVPs.

500 We performed a single controlled experiment to acquire data for this investigation, as described below.

501 5.1 New Protocol

502 We devised a new dual-minimization protocol, where the participant aims towards a very small dot which appears
 503 at a given distance and a pseudo-random angle. We selected the smallest dot possible that was still attainable with
 504 reasonable effort (less than 1.33 mm wide). Participants were instructed to select the target by emphasizing accuracy,
 505 but taking no more time than needed. We used a pseudo-random method to place targets, to remove prediction⁶ and
 506 learning effects. The protocol we propose is essentially a balanced, reciprocal, 2D version of Guiard's dual-minimization
 507 protocol [22, 23].

508 509 ⁵The inequality reads $\sqrt{xy} \leq \frac{x+y}{2}$, which gives $\frac{1}{2} \log \sqrt{\Gamma_{\alpha\alpha}\Gamma_{\beta\beta}} \leq \frac{1}{2} \log \frac{\Gamma_{\alpha\alpha}+\Gamma_{\beta\beta}}{2}$. The left hand side is PVP_{det}, while the right hand side is always
 510 smaller than PVP_{Tr}.

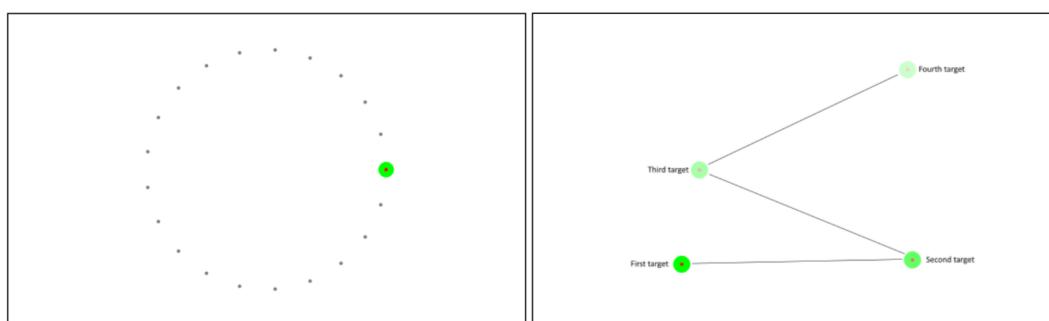
511 512 ⁶where the participant already knows where the next target is going to appear

521 **5.2 Procedure**

522 We conducted a two factor full factorial design, with factors PROTOCOL \times DEVICE, where PROTOCOL is either ISO or
 523 DualMin and DEVICE is one of {Mouse, Touchpad, Controller}. In the ISO condition, participants were instructed to
 524 perform the ISO ring of rings task as quickly as possible while aiming for an error rate of about 4%, as prescribed by the
 525 ISO standard [34]. In the DualMin condition, participants were instructed to perform the new protocol.
 526

527 Each participant performed 6 blocks, where each block is a combination of PROTOCOL \times DEVICE. To counterbalance
 528 order and carry-over effects for our 12 participants, we used two (different) 6-by-6 latin squares. To keep the experiment
 529 length less than 30 minutes per participant, we did not use a full factorial design for D and W in the ISO condition.
 530 Instead, we determined a set of 6 (D,W) pairs that were randomly shuffled for each new block: {(200, 124), (440, 66),
 531 (1160, 90), (920, 38), (680, 18), (1400, 14)}, resulting in ID's {1.4, 2.9, 3.8, 4.7, 5.3, 6.7}. These pairs were determined to
 532 satisfy a multi-objective optimization problem with following aims: a) low correlation between D and W, to estimate
 533 effects and associated standard errors of D and W properly [14] b1) equally spaced ID, and b2) maximum range of ID
 534 spanned, to have a good range for the covariates c) reasonable values for D and W (no risk of hitting the edge of the
 535 screen, target attainable). More details can be found in the Supplementary material. The 6 D conditions were similarly
 536 shuffled in the DualMin condition.
 537

538 The experiment starts with a training session that ends when the participant signals that they are accustomed to the
 539 setup. For each condition, and for each (D,W) pair, participants were asked to perform 21 trials. The first trial was
 540 systematically dropped. Optional 1-minute breaks were offered to participants between each condition (each condition
 541 consisting of 126 movements—which take between 1 and 2 second per movement for most participants—this amounts
 542 to a potential 1-minute break every 2 to 5 minutes).
 543



559 Fig. 4. Left: Task in the ISO condition. All targets are visible. We have represented the condition with smallest W. Right: Task in the
 560 DualMin condition. Only one target is visible at a time during the experiment, but 4 images have been stacked here. For both panels,
 561 the user should aim towards the red target; the green disk is only here to facilitate finding small targets which may be hard in the
 562 DualMin condition.
 563

564 **5.3 Apparatus**

565 Participants were seated on a chair during the experiment, which they could adjust as needed. The experiment
 566 was performed on a Windows Microsoft Surface Pro 4 (2736x1824), running PopOS with the Linux Surface Kernel.
 567 Participants were asked to use a standard commercial external mouse, the Surface's native touchpad or an external
 568 Xbox controller in a random order. In the case of the mouse and touchpad, the native cursor acceleration was enabled,
 569

573 and sensitivity was set to PopOS's default setting. For the controller, we used a custom transfer function, as we judged
 574 the default on PopOS to be poor in a pilot study.

575 In the *ISO* condition, participants were shown a gray arrangement of circles, one of which was red to indicate the
 576 target to select. In the *DualMin* condition, participants were shown only a red target. In both cases, a larger green circle
 577 surrounded the target, so that participants wouldn't spend time searching for the target, see Fig. 4.
 578

579 5.4 Data Collection

580 Our software captured cursor position at 100Hz, and registered full trajectory information (position and time), target
 581 information, and bookkeeping information. In total, we performed $12 \times 6 \times 6 = 432$ blocks, and $432 * 21 = 9072$
 582 movements. This dataset is available online at (repository will be added in final version).
 583

584 5.5 Data Analysis

585 5.5.1 *2D-PVP profiles*. We computed PVP profiles for each block using a custom library (library repository will be
 586 added in final version). First, we discarded the first trajectory of each block. To synchronize the trajectories, we started
 587 by re-interpolating the time series at 100 Hz to account for unequal sampling times. We then applied a low-pass filter
 588 on the time series (kaiser window) with cut-off frequency at 10 Hz. Trajectories were then pooled (per condition), and
 589 extended. Finally, we computed PVP_{det} , PVP_α and PVP_{Tr} as described in Section 4.
 590

591 5.5.2 *Effect of task parameters on 2D-PVP features*. To investigate the level of association between PVP features and
 592 task parameters, we ran linear multilevel regressions [14] on τ , D_τ , σ_0 and C using *lme4* [2]. These models allow
 593 fitting an intercept and slope per device while controlling for the variance induced by multiple participants and the
 594 dependency of data in repeated measurements. We defined D, W and the interaction $D \times W$ as fixed effect. We also
 595 controlled for dependence of data by adding a random effect PARTICIPANT and controlled the variation due to DEVICE
 596 when needed. The various statistical models we fitted are described using the syntax of R formulae. An example formula
 597 is $\text{tau} \sim D*W + \text{Device} + \text{Device}:D + (1|\text{Participant})$. Here, Device is a categorical factor with n categories, and
 598 will add $(n - 1)$ constants in the model (the baseline device being integrated in the model's intercept), Device:D is
 599 an interaction term which will add $(n - 1)$ device-dependent slopes (the baseline device being in the model's slope),
 600 $(1|\text{Participant})$ adds a random intercept per participant, and $D*W$ is equivalent to $D + W + D:W$.
 601

602 5.5.3 *Comparison between Fitts' law experiment and 2D-PVP*. We estimated nominal and effective version of Fitts'
 603 law parameters for each device separately. We then estimated the intercept and slopes for each device based on our
 604 predictive formulae Eqs. (9) and (11), and we compared these values against each other.
 605

606 5.6 Participants

607 We recruited 12 participants, based on a power analysis, where we used data from the Pointing Dynamics dataset [30]
 608 as pilot data. We extracted PVPs and corresponding parameters with our library, on which we fit a general linear
 609 multilevel model as described above. For our target effect size, we used a conservative estimate, namely the lower bound
 610 of the 95% confidence interval of the effect size of the coefficient associated with the D covariate for regressions on τ ,
 611 D_τ , σ_0 . That means that even if D's effect turns out to be 2 standard deviations smaller than what was estimated from
 612 the pilot data, we would still have the appropriate statistical power to detect it. To computer power curves, we used
 613 *simr* [18]. Taking the least favorable of the three regressions, with a target power $\beta = 0.8$, yields $N = 9$ participants as
 614

625 our target sample size. We ended up recruiting 12 participants to make balancing the experiment design, which has
 626 $2 \times 3 = 6$ levels, easier. The full analysis can be found in the supplementary material.
 627

628 6 EMPIRICAL STUDY: RESULTS

630 In total, we computed 1292 profiles across 432 blocks (4 missing due to a bug in the software).
 631

632 6.1 Data presentation

633 PVP $_{\alpha}$ for the 6 DEVICE \times PROTOCOL conditions for a single participant are displayed Fig. 5. The mouse produces the
 634 cleanest profiles, with PVPs quite packed, the controller produces very long and gentle but noisy slopes in the second
 635 phase, while the touchpad has profiles that are affected by D much more than the others, as visible by how PVPs are
 636 separated. We also see the difference in target sizes in the ISO condition, where we see the PVPs level off at different
 637 values of standard deviation. Overall, we see that the profiles are more noisy than what was presented in [15]; this is
 638 likely due to the fact that we only have 20 movements per block, whereas they used about 40 to 50 movements per
 639 block. Also, the controller and touchpad produce more variable trajectories than the mouse; this is especially true for
 640 the controller.
 641

642 6.2 Three phases

643 To identify the three phases of the profiles, we fitted each profile as described in Subsection 3.3. We determined that 8
 644 out of 1092 profiles were not a good match with the three phase description (either $r^2 < 0.8$, or those who produced a
 645 positive slope ($C < 0$)). The bad profiles were all in the ISO protocol, mostly for the smallest ID (5/8 for ID = 1.4).
 646

647 6.3 Comparison between PVPs

648 A visual comparison of the three PVPs is shown Fig. 9 (Appendix), which illustrates Eq. (21): PVP $_{\text{Tr}}$ is always above
 649 the others, while PVP $_{\alpha}$ is usually above PVP $_{\text{det}}$. Overall, PVP $_{\alpha}$ follows PVP $_{\text{Tr}}$ during the early parts of the profiles
 650 and then crosses over to PVP $_{\text{det}}$. In the remainder of the work, we present results only for PVP $_{\alpha}$: effect sizes may be
 651 marginally different between PVPs but no significant difference emerged between the profiles in our analyses.
 652

653 In the remainder, we thus focus on PVP $_{\alpha}$. After visualizing the data, we fit a linear multilevel model to each of τ ,
 654 D_{τ} , σ_0 and C, with the goal of estimating the effect sizes. We won't be focusing on statistical significance: given the
 655 large sample size (N=430) many of the effects we estimate will be significant.
 656

657 6.4 Effects of D on τ , D_{τ} , σ_0 and C

658 We first visualize the effects of D on τ , D_{τ} , σ_0 and C via boxplots grouped by device in Fig. 6, since we hypothesize
 659 that different devices may lead to different phase features. Boxplots for the effects of W and ID are shown in the
 660 Supplementary material.
 661

662 Results show that

- 663 • τ is about constant for the mouse and touchpad, with the touchpad leading to somewhat larger τ . The effect of
 664 D on τ is however marked for the controller, for which τ is also on average much larger.
- 665 • D_{τ} shows a strong increase with D. We see that the increase levels off for the controller, which suggests the
 666 slope for the controller will be different from the other devices.

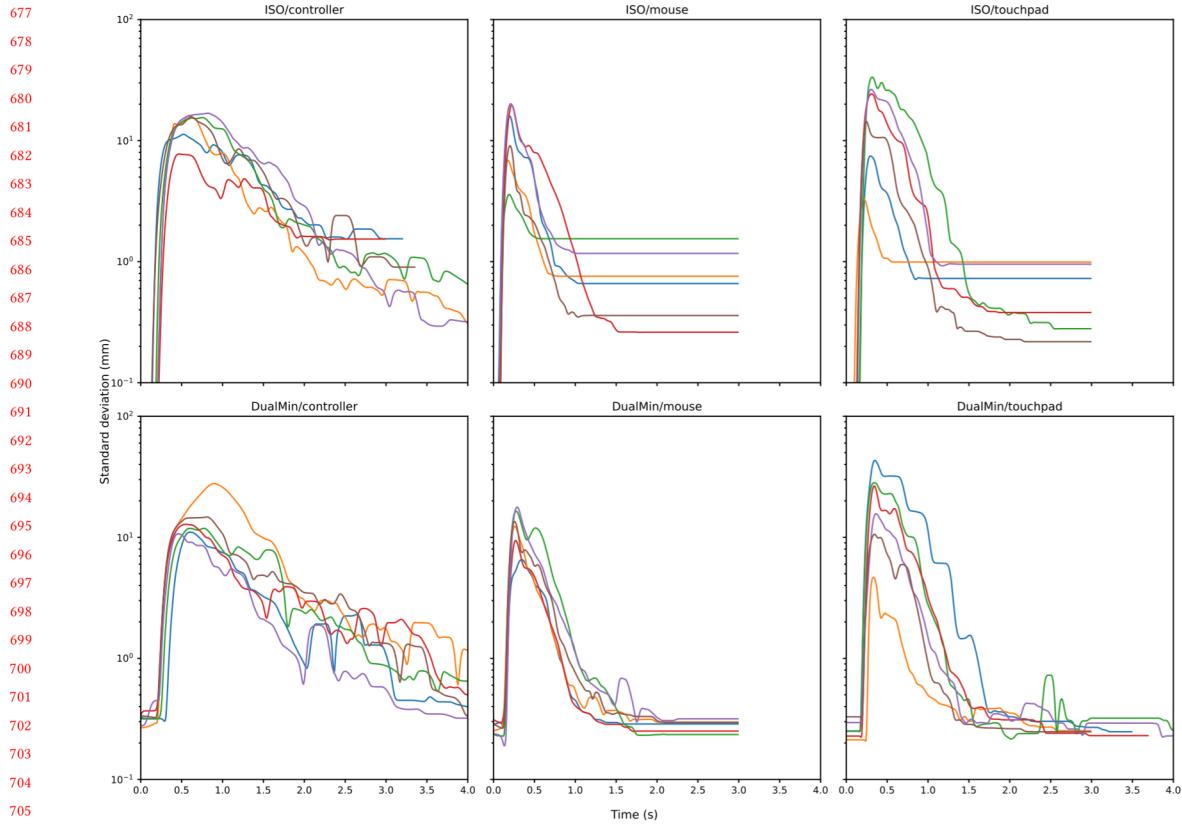


Fig. 5. PVP_α for Participant 136418 for the 6 PROTOCOL \times DEVICE conditions. Each plot displays the PVP (standard deviation in mm) against time (s) with identical scales.

- For σ_0 , the case is similar to D_τ , although this time slope for touchpad is quite different from the two other devices.
- C seems constant across values of D. There is a notable difference between the controller, which has values for C two to three times lower than the mouse and touchpad.

We then used multilevel models to estimate the effect sizes of D, because they allow us to control for variations induced by the other factors by including random intercepts and slopes, based on the boxplots. The estimated effect sizes are given Tab. 1, and the full fit results are delayed to Tab. 3 in the Appendix.

6.4.1 Duration of the first phase: τ .

We fitted the following statistical model

$\text{tau} \sim D * W + (1 | \text{Participant}) + \text{device} + \text{device:D}$.

The intercept contributes by far the most to the value of τ , ($\tau_{\text{controller}} = 0.58$ s, $\tau_{\text{mouse}} = 0.36$ s, $\tau_{\text{touchpad}} = 0.40$ s), with other effects being small, except for the slope in the controller⁷.

⁷D \times mouse represents the difference in D slope between the baseline group (controller) and the mouse group. Summing the two slopes gives a null slope, showing that D has virtually no effect on τ in the mouse group. The same is true for the touchpad group

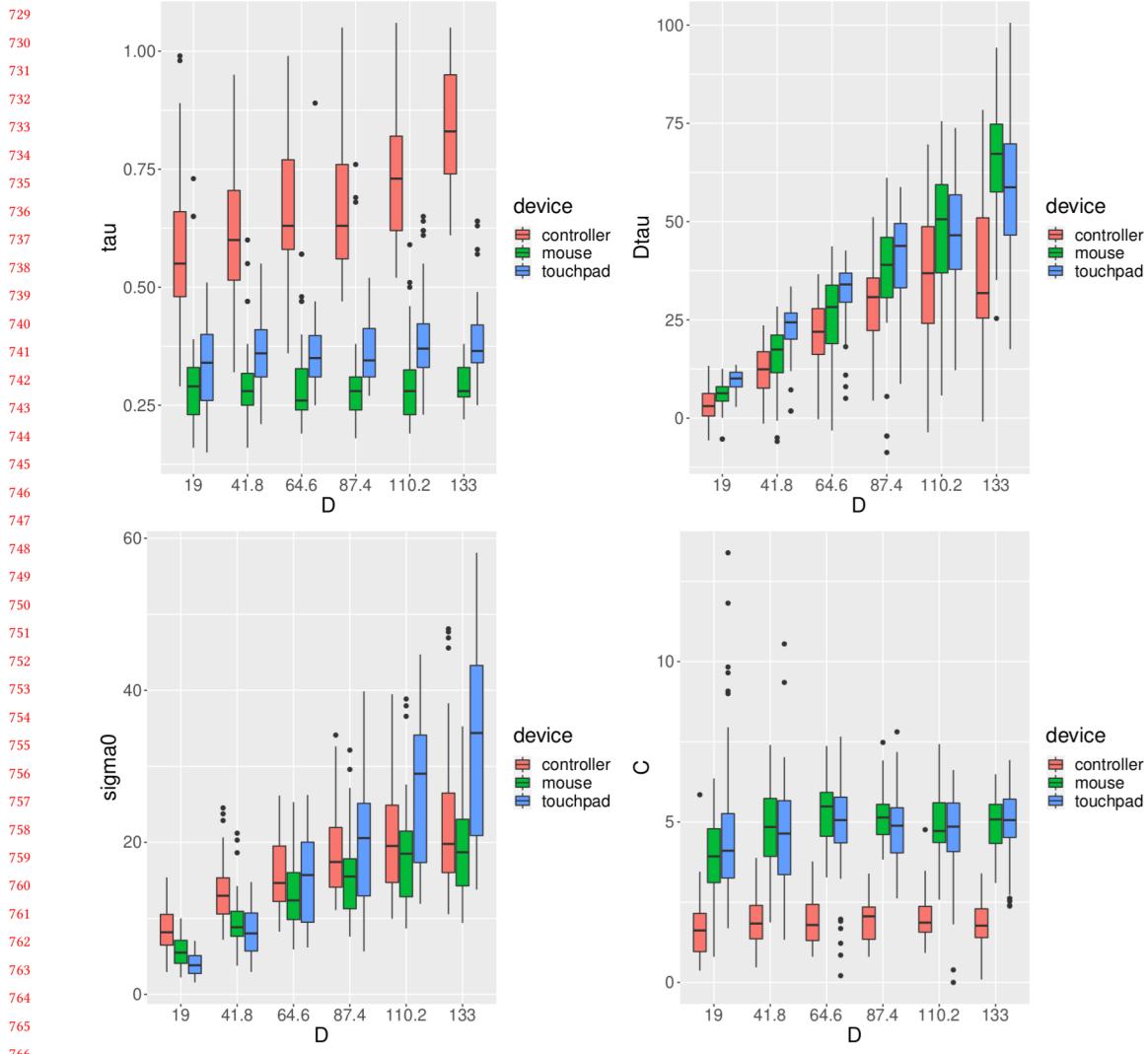


Fig. 6. Boxplots grouped by device, showing the effect of D on: top left – τ , top right – D_τ , bottom left – σ_0 , bottom right – C .

Table 1. Estimated effect sizes for PVP features. For each covariate, we provide the estimated effect size, as well as average contribution to τ in percentage.

	Mean	Intercept	D	W	mouse	touchpad	$D \times W$	$D \times \text{mouse}$	$D \times \text{touchpad}$
τ (s)	0.470	0.579 (123%)	0.003 (49%)	-0.013 (-9%)	-0.222 (-47%)	-0.177 (-38%)	$3.5e^{-5}$ (2%)	-0.003 (-49%)	-0.003 (-49%)
D_τ (mm)	32	1.99 (6%)	0.248 (60%)	-0.326 (-4%)			0.003 (2%)	0.232 (56%)	0.216 (52%)
σ_0 (mm)	17.7	4.28 (29%)	0.16 (69%)	-0.013 (-3%)			$2.7e^{-4}$ (0%)	-0.023 (-10%)	0.089 (38%)
C (bit/s)	4.11	1.66 (40%)	0.004 (7%)	0.059 (1%)	3.06 (75%)	3.21 (79%)	-0.001 (-6%)		

781 Table 2. Estimated parameters for nominal and effective Fitts' law using the classical method, compared with \hat{a} 's and \hat{b} estimated
 782 from Eqs. (9) and (11).

	ε	Nominal b	Effective b	\hat{b}	a Effective	\hat{a} Effective	a Nominal	\hat{a} Nominal
785 <i>mouse</i>	0.8%	0.15	0.19	0.21	0.25	0.21	0.37	0.27
786 <i>controller</i>	2%	0.40	0.35	0.54	1.10	0.60	0.71	0.58
788 <i>touchpad</i>	0.8%	0.24	0.21	0.21	0.51	0.41	0.39	0.51

790
 791 6.4.2 *Distance covered during the first phase: D_τ .* We fitted the following statistical model

$$792 \text{Dtau} \sim D * W + (1 | \text{Participant}) + \text{device}:D.$$

793
 794 There is a strong effect of D on D_τ , which varies per device: the effect of D is about twice as strong for *mouse* and
 795 *touchpad* compared with *controller*, while other effects contribute marginally ($D_{\tau,\text{mouse}} = 0.48D$, $D_{\tau,\text{touchpad}} = 0.47D$,
 796 $D_{\tau,\text{controller}} = 0.25D$).

797
 798 6.4.3 *Standard deviation at the end of the first phase: σ_0 .* We fitted the following statistical model

$$800 \text{Sigma0} \sim D * W + (1 | \text{Participant}) + \text{device}:D.$$

801
 802 There is a strong effect of D on σ_0 , with D having a stronger effect for *touchpad* than the other two devices
 803 ($\sigma_{0,\text{controller}} = 4.3 + 0.16D$, $\sigma_{0,\text{mouse}} = 4.3 + 0.14D$, $\sigma_{0,\text{touchpad}} = 4.3 + 0.25D$). The intercept contributes significantly,
 804 accounting for about 30% of σ_0 's average value.

805
 806 6.4.4 *Participant performance: C.* We fitted the following statistical model

$$807 \text{C} \sim D * W + (1 | \text{Participant}) + \text{device}.$$

808
 809 The only strong effects are the intercept's, where C for the controller is much lower than the other two devices
 810 ($C_{\text{mouse}} = 4.72 \text{ bit/s}$, $C_{\text{touchpad}} = 4.87 \text{ bit/s}$, $C_{\text{controller}} = 1.66 \text{ bit/s}$).

811 6.5 Comparison with Fitts' law

812 To compare Fitts' law evaluations with our new evaluation, we first fitted Fitts' model, with separate intercept and
 813 slopes for each device and for both the nominal and the effective version:

$$814 \text{MT} \sim \text{IDe} + \text{device} + \text{device}: \text{IDe}$$

$$815 \text{MT} \sim \text{ID} + \text{device} + \text{device}: \text{ID}$$

816 The parameters estimated via the linear model are shown Fig. 7 and Tab. 2 (also see Tab. 4 in the Appendix).

817
 818 We then estimated the intercept and slope \hat{a} and \hat{b} for the nominal and effective laws based on Eqs. (9) and (11), and
 819 effect sizes from Tab. 1. The predictions are accurate for the mouse and touchpad both in the effective and nominal
 820 versions of the law for the intercepts (with a maximum error of about 120 ms for the touchpad in the nominal version).
 821 For the slope estimates, the estimates are particularly precise for the effective version of the law (error of 0.02 s/bit for
 822 the mouse, below 0.005 s/bit for the touchpad), but a little less precise for the nominal version (0.04 s/bit error for the
 823 mouse, 0.03 s/bit for the touchpad). The predictions are less accurate for the controller (about 0.2 s/bit for the slope, and
 824 half a second of error for the intercept in the effective version).

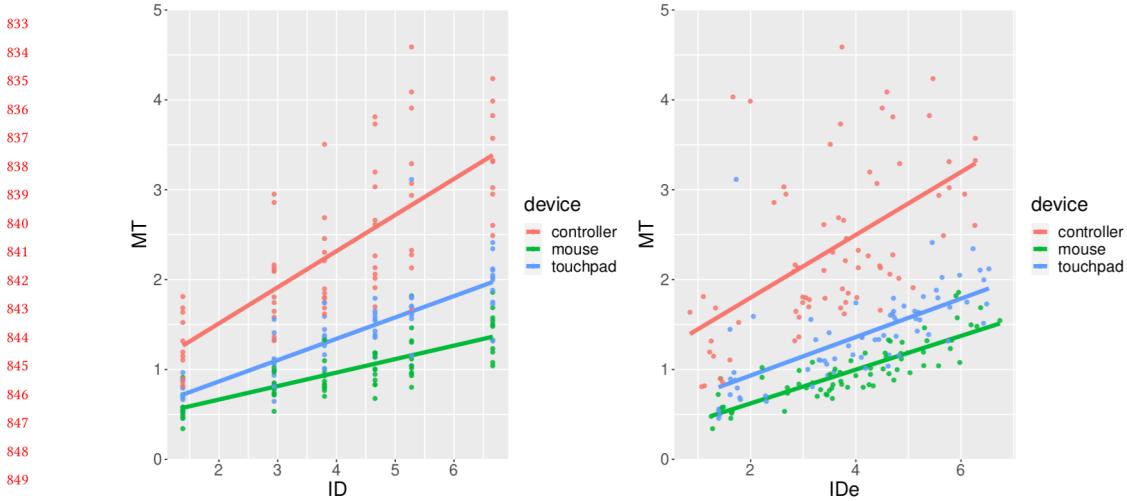


Fig. 7. Fitts' law in nominal (left) and effective (right) versions fitted on data from the *ISO* condition.

7 DISCUSSION

7.1 Two(three) phase model

The controlled experiment supported that the two⁸ phase model was applicable in almost all blocks. We included two popular devices (mouse and touchpad), as well as a less popular one, with a custom transfer function, to evaluate this property broadly. The fact that even with the controller we clearly see the two phases supports the idea that this property is due to a low-level human process, and suggests it will be applicable in many HCI situations.

Two-phase models are actually widespread in the study of human movement, starting with Woodworth [37], who hypothesized that a ballistic phase primarily used to cover distance was followed by a second homing-in phase. This two-phase description was further developed, particularly by Elliott and colleagues [8–10] who suggested that the two phases were a feedforward control phase followed by an online (feedback) control phase.

PVPs are aligned with that description: the first phase of the PVP is a fast distance covering mechanism in which variance increases – properties usually associated with feedforward control [3, 13] – while the second phase of the PVP is a longer one based on a variance reduction model with feedback information.

7.2 Which PVP version should we use?

Our study showed little difference between the three PVP versions that we proposed. Most of the variance is along the $\Gamma_{\alpha\alpha}$ component at the start of the movement, it gradually becoming closer to the $\Gamma_{\beta\beta}$ component, as shown by the fact PVP_α follows PVP_{Tr} at the start of the movement, but then ends up following PVP_{det} towards the end of the movement. While we could have selected a criterion to compute a score and select the “best one”, we feel there is little sense in doing that: we believe that the user of PVPs should reflect which variance component they want to account for ($\Gamma_{\alpha\alpha}, \Gamma_{\beta\beta}, \Gamma_{\alpha\beta}$) in order to determine which one is the most appropriate. We suppose that difference in PVP versions will be meaningful only for data with high variance, in particular when in the direction orthogonal to movement.

⁸We intentionally play on the fact that while the model is a two-phase one, the method introduces a third phase for the purpose of estimation.

885 7.3 Effect of task parameters on PVP features

886 Overall, we replicated the results from the 1D study [15]. All three devices shared the same predominant effects of
 887 D on PVP features. The only discrepancy that we found with [15] is an intercept on σ_0 . These features inform us on
 888 the device in more detail than what can be achieved with only the intercept a and slope b from a traditional Fitts' law
 889 experiment. For example, in comparison with the mouse and the touchpad, for the controller, τ was much higher, while
 890 D_τ was lower: the first phase, while being longer, covered less distance. This suggests that the transfer function of the
 891 controller could be better tuned (we believe that the maximum speed of the cursor was too low, and hence a higher
 892 gain in the high speed condition would be beneficial).
 893

894 7.4 Link with Fitts' law experiments and parameters

895 At a high level, our results paint the same picture between Fitts' law experiments and PVPs: performance between the
 896 touchpad and the mouse are similar, while the touchpad scores comparatively lower. Interestingly though, the protocol
 897 that we used is much simpler: instead of crossing D and W as in Fitts' law experiments, we only need a condition with
 898 a very small target and large distance for the *DualMin* protocol. In this work, we used 6 different conditions in the
 899 *DualMin* protocol, but this was only to provide enough values of D to validate the properties of Section 3 for 2D PVPs:
 900 a single condition would have been enough. We were also able to predict with a good accuracy the parameters of Fitts'
 901 law parameters from the PVP features, despite the predictive formulae incorporating highly non-linear terms (e.g.,
 902 $\log_2(\text{erf}^{-1}(1 - x))$). To our knowledge, this is the first time a model is able to predict the values, and not just the shape
 903 associated with Fitts' law.
 904

905 8 IMPLICATIONS FOR FITTS' LAW AND FITTS' LAW EXPERIMENTS

906 In the introduction, we listed some open problems and discussions regarding Fitts' law experiments. Below, we revisit
 907 that list under the light of our results.
 908

909 8.1 Which version of ID should we use, and why?

910 Eqs. (9) and (11) suggest the nominal index $ID = \log_2 \frac{D}{W}$ and the effective index $ID_e = \log_2 \frac{D}{\sigma_\infty}$ where σ_∞ is equivalent
 911 to σ in Eq. (2). Thus, the Variance Model [15] is consistent with the received expression for ID except for the "+1" term
 912 inside the logarithm. How much difference does that "+1" cause? As previously discussed [15, 25, 38], that term is of
 913 little interest, because it induces changes only for very low values of the ratio D/W where Fitts' law is known to be a
 914 poor model [6]. In fact, for low D/W, the proper approach would be to better model the first phase, and build a model
 915 based on that. As a result, we suggest using $ID = \log_2 \frac{D}{W}$ or $ID_e = \log_2 \frac{D}{W_e}$, while maintaining $ID > 2$ or 3.
 916

917 The *Variance Model*, via Property 6 also suggests a *local* index of difficulty id, valid between any two instants t_1 and
 918 t_2 in the variance decreasing phase
 919

$$920 \quad \text{id} = \log_2 \frac{\sigma(t_1)}{\sigma(t_2)}, \quad (22)$$

921 According to Guiard [21], ID runs on a non-ratio scale of measurement. In contrast, the local id is a well-defined ratio
 922 between two measurable standard deviations.
 923

937 8.2 What is the interpretation of ID?

938 ID or id are, contrary to the received interpretation [34], not expressed as channel capacities; instead they measure the
 939 information that needs to be transmitted to *reliably* perform the aiming task, independently of the distance covering
 940 mechanism. So, rather than being “rates”, ID and id are “loads”, which actually matches the notion of difficulty inherent
 941 to the name “index of difficulty”: a more difficult task requires more information to be transmitted. This interpretation
 942 shows that MacKenzie’s argument for adding “+1” to the logarithm “so that $\log_2(1 + \frac{D}{W})$ better matches the formula
 943 for the Shannon capacity” [27] is thus flawed, since ID is not a capacity in the first place.
 944

945 8.3 How should the intercept be interpreted?

946 The intercept of Fitts’ law can be expressed from the PVP features, see Eqs. (4), (6), (8) and (10), and [15, Eqs. (27), (29)]
 947

$$948 \quad a = \tau_0 + \alpha D + \frac{1}{C} \log_2(2\sqrt{2}\gamma \times \text{erf}^{-1}(1 - \varepsilon)) \quad (\text{nominal}), \quad (23)$$

$$949 \quad a = \tau_0 + \alpha D + \frac{1}{C} \log_2(4.133\gamma) \quad (\text{effective}), \quad (24)$$

950 where c and d are some functions. Thus, τ is the contribution of
 951

- 952 (1) An incompressible positive time τ_0 , which differs from one participant to another (see Eq. (4)).
- 953 (2) A scaling effect due to distance, via the term α in Eq. (12). This scaling effect accounts for about XXX
- 954 (3) A second scaling effect, due to γ , that relates the standard deviation at the end of the first phase to D . This effect
 955 also includes:
 - 956 • the arbitrary 4.133 constant in case of the effective law;
 - 957 • the miss rate via $\text{erf}^{-1}(1 - \varepsilon)$ in case of the nominal law.

958 Judging from the many contributions to the intercept, it becomes clear why interpreting the intercept is so hard [38]⁹
 959 However, contrary to what is typically argued [34, 38], there seems to be no reason to expect a positive or null intercept:
 960 in fact negative intercept may be predicted, for example if the effect of γ is very small.

961 This expression for intercept is also interesting, in that it expresses what Guiard [20] calls the *weak version* of Fitts’
 962 law, the strong version being the one presented here. In the weak version, an effect of scale shows up, and MT has
 963 a term linear in D . For some experiments, such as Fitts’ disc-transfer experiment [11], that effect is noticeable. The
 964 expression for the intercept shows that both the weak and strong versions of Fitts’ law can be predicted (depending on
 965 the value of α).
 966

967 8.4 Can throughput be unequivocally defined?

968 The throughput is an effective rate of information transmission; given a high effort by the participant such as that
 969 produced in a controlled experiment, it is hypothesized that throughput can reach the capacity (*i.e.*, the highest
 970 theoretically attainable rate) [17]. In the literature, two throughput propositions dominate, namely $1/b$ and the ISO-
 971 throughput (Eq. (3)), with no real consensus as to which one is preferable [17].
 972

973 The capacity in *Variance Model* matches the $1/b$ throughput according to Eq. (9), exactly as advocated by Zhai [38].
 974 Hence, throughput, in the sense of information transmission, is defined as $1/b$. However, such a throughput is likely
 975 not what the practitioner is looking for: it characterizes only the transmission part via the second phase of movement,
 976 without accounting for the distance covering part that would include the first phase. To better understand this, one can
 977

978 ⁹Even excluding things like reaction time or extra cognitive processes that may occur.
 979

989 think of a device, where C is relatively high *i.e.*, variance reduction is very effective, but which creates a lot of variance
 990 at the end of the first phase. The result would be a device with high throughput, but an overall poor performance, as
 991 measured by MT (and/or T). Unfortunately, there is little that information theory, and thus the *Variance Model*, can
 992 offer to provide an all encompassing measure of performance that includes both phases. Hence, we believe that both C
 993 and ISO-throughput Eq. 3 are useful metrics. However, we think that ISO-throughput should be renamed to make it
 994 clear it is not a throughput related to information transmission — for example the *ISO performance measure*.
 995

996 8.5 Should participants aim for a 4% Miss Rate?

997 It has long been argued, based on an information-theoretic formula, that a 4% participant miss rate was something the
 998 experimenter should strive for in a Fitts' law experiment [17, 34]. This argument is based on the entropy of a Gaussian
 999 distribution, which introduces the 4.133 factor in effective versions of Fitts' law 2. However, as shown by Eq. (24) τ ,
 1000 that term is arbitrary and actually carries over to the intercept; one could introduce any other value, with the only
 1001 effect of changing the intercept. In the nominal version of the law, Eq. (23), it even disappears. Hence, there can be
 1002 no information-theoretic reason to instruct a participant to aim for a 4% miss rate.¹⁰ The dual-minimization protocol
 1003 solves the issue by considering errors rather than misses (for the difference between misses and errors, refer to [16]).
 1004

1005 9 CONCLUSION

1006 In this work, we proposed an alternative to the evaluation of input performance via Fitts' law experiments, by building
 1007 on a theoretical model from the computational neurosciences. We successfully extended the associated method —called
 1008 the method of PVPs— to deal with 2D data. The method is associated with a simpler protocol, yet leads to more
 1009 expressive results, while remaining compatible and even predicting Fitts' law parameters. Because it is associated
 1010 with a theoretical model, it becomes easier to interpret and guide choices made during the design and the analysis
 1011 of controlled experiments. Future work will consist in further exploring the method of PVPs. Interesting questions
 1012 would be to determine the minimum number of samples needed to estimate PVP features reliably, to further reduce the
 1013 duration of experiments destined to measure input performance. It would also be interesting to study more extreme
 1014 cases of input performance, such as very low accuracy techniques (*e.g.*, eye/head tracking), or for participants with
 1015 physical disabilities. On the whole, we believe the method to be a promising alternative to Fitts' law.
 1016

1017 REFERENCES

- 1018 [1] Johnny Accot and Shumin Zhai. 2003. Refining Fitts' law models for bivariate pointing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 193–200.
- 1019 [2] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1 (2015), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- 1020 [3] Bastien Berret, Adrien Conessa, Nicolas Schweighofer, and Etienne Burdet. 2021. Stochastic optimal feedforward-feedback control determines timing and variability of arm movements with or without vision. *PLOS Computational Biology* 17, 6 (2021), e1009047.
- 1021 [4] S. K. Card, W. K. English, and B. J. Burr. 1978. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics* 21, 8 (1978), 601–613. <https://doi.org/10.1080/00140137808931762>
- 1022 [5] E. R. F. W. Crossman. 1957. *The speed and accuracy of simple hand movements*. Technical Report .
- 1023 [6] E. R. F. W. Crossman and P. J. Goodeve. 1983. Feedback control of hand-movement and Fitts' law. *The Quarterly Journal of Experimental Psychology* 35, 2 (1983), 251–278.
- 1024 [7] H. Drewes. 2010. Only One Fitts' Law Formula Please!. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) (CHI EA '10). ACM, New York, NY, USA, 2813–2822. <https://doi.org/10.1145/1753846.1753867>
- 1025 [8] Digby Elliott, Steve Hansen, Lawrence EM Grierson, James Lyons, Simon J Bennett, and Spencer J Hayes. 2010. Goal-directed aiming: two components but multiple processes. *Psychological bulletin* 136, 6 (2010), 1023.

1026 ¹⁰It also happens to be very hard to enforce in practice.

- 1041 [9] Digby Elliott, Werner F Helsen, and Romeo Chua. 2001. A century later: Woodworth's (1899) two-component model of goal-directed aiming. *Psychological bulletin* 127, 3 (2001), 342. <https://doi.org/10.1037/0033-2909.127.3.342>
- 1042 [10] Digby Elliott, James Lyons, Spencer J Hayes, James J Burkitt, James W Roberts, Lawrence EM Grierson, Steve Hansen, and Simon J Bennett. 2017. The multiple process model of goal-directed reaching revisited. *Neuroscience & Biobehavioral Reviews* 72 (2017), 95–110. <https://doi.org/10.1016/j.neubiorev.2016.11.016>
- 1043 [11] P. M. Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381. <https://doi.org/10.1037/h0045689>
- 1044 [12] Paul M Fitts and James R Peterson. 1964. Information capacity of discrete motor responses. *Journal of experimental psychology* 67, 2 (1964), 103. <https://doi.org/10.1037/h0045689>
- 1045 [13] K. Gan and E. R. Hoffmann. 1988. Geometrical conditions for ballistic and visually controlled movements. *Ergonomics* 31, 5 (1988), 829–839.
- 1046 [14] Andrew Gelman and Jennifer Hill. 2006. *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.
- 1047 [15] Julien Gori and Olivier Rioul. 2020. A feedback information-theoretic transmission scheme (FITTS) for modeling trajectory variability in aimed movements. *Biological Cybernetics* 114, 6 (2020), 621–641.
- 1048 [16] J. Gori, O. Rioul, and Y. Guiard. 2017. To Miss is Human: Information-Theoretic Rationale for Target Misses in Fitts' Law. In *CHI'17: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. ACM, Denver, USA, 5.
- 1049 [17] J. Gori, O. Rioul, and Y. Guiard. 2018. Speed-Accuracy Tradeoff: A Formal Information-Theoretic Transmission Scheme (FITTS). *ACM Trans. Comput.-Hum. Interact.* 25, 5, Article 27 (Sept. 2018), 33 pages. <https://doi.org/10.1145/3231595>
- 1050 [18] Peter Green and Catriona J. MacLeod. 2016. simr: an R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution* 7, 4 (2016), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- 1051 [19] Tovi Grossman and Ravin Balakrishnan. 2005. A probabilistic approach to modeling two-dimensional pointing. *ACM Transactions on Computer-Human Interaction (TOCHI)* 12, 3 (2005), 435–459.
- 1052 [20] Yves Guiard. 2019. Polar and Cartesian Structure in the Data of Fitts's (1954) Classic Experiments-With a Criterion for Distinguishing a Strong and a Weak Version of Fitts' Law. *Journal of Motor Behavior*, (2019), 1–23.
- 1053 [21] Y. Guiard and H. B. Olafsdottir. 2011. On the measurement of movement difficulty in the standard approach to Fitts' law. *PloS one* 6, 10 (2011), e24389. <https://doi.org/10.1371/journal.pone.0024389>
- 1054 [22] Yves Guiard, Halla B Olafsdottir, and Simon T Perrault. 2011. Fitt's law as an explicit time/error trade-off. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, ACM, New York, NY, USA, 1619–1628. <https://doi.org/10.1145/1978942.1979179>
- 1055 [23] Y. Guiard and O. Rioul. 2015. A mathematical description of the speed/accuracy trade-off of aimed movement. In *Proceedings of the 2015 British HCI Conference*. ACM, New York, 91–100.
- 1056 [24] Peter A Hancock and Karl M Newell. 1985. The movement speed-accuracy relationship in space-time. In *Motor Behavior*. Springer, , 153–188.
- 1057 [25] E. R. Hoffmann. 2013. Which version/variation of Fitts' law? A critique of information-theory models. *Journal of motor behavior* 45, 3 (2013), 205–215. <https://doi.org/10.1080/00222895.2013.778815>
- 1058 [26] Richard J Jagacinski and Donald L Monk. 1985. Fitts' Law in Two dimensions with hand and head movements movements. *Journal of motor behavior* 17, 1 (1985), 77–95.
- 1059 [27] I. S. MacKenzie. 1989. A note on the information-theoretic basis for Fitts' law. *Journal of motor behavior* 21, 3 (1989), 323–330. <https://doi.org/10.1080/00222895.1989.10735486>
- 1060 [28] I. S. Mackenzie. 1992. *Fitts' law as a performance model in human-computer interaction*. Ph.D. Dissertation. University of Toronto.
- 1061 [29] I. S. MacKenzie and W. Buxton. 1992. Extending Fitts' law to two-dimensional tasks. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, New York, 219–226.
- 1062 [30] Jörg Müller, Antti Oulasvirta, and Roderick Murray-Smith. 2017. Control theoretic models of pointing. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 4 (2017), 27. <https://doi.org/10.1145/3121431>
- 1063 [31] R. Plamondon and A. M. Alimi. 1997. Speed/accuracy trade-offs in target-directed movements. *Behavioral and Brain Sciences* 20, 02 (1997), 279–303. <https://doi.org/10.1017/S0140525X97001441>
- 1064 [32] Hiroaki Sakoe and Seibi Chiba. 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing* 26, 1 (1978), 43–49.
- 1065 [33] R. A. Schmidt, H. Zelaznik, B. Hawkins, James S. F., and J. T. Quinn Jr. 1979. Motor-output variability: a theory for the accuracy of rapid motor acts. *Psychological review* 86, 5 (1979), 415.
- 1066 [34] R William Soukoreff and I Scott MacKenzie. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International journal of human-computer studies* 61, 6 (2004), 751–789. <https://doi.org/10.1016/j.ijhcs.2004.09.001>
- 1067 [35] Emanuel Todorov. 2004. Optimality principles in sensorimotor control. *Nature neuroscience* 7, 9 (2004), 907–915.
- 1068 [36] J. O. Wobbrock, K. Shinohara, and A. Jansen. 2011. The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, ACM, , 1639–1648.
- 1069 [37] R. S. Woodworth. 1899. Accuracy of voluntary movement. *The Psychological Review: Monograph Supplements* 3, 3 (1899), i.
- 1070 [38] S. Zhai. 2004. Characterizing computer input with Fitts' law parameters—the information and non-information aspects of pointing. *International Journal of Human-Computer Studies* 61, 6 (2004), 791–809.

- 1093 [39] S. Zhai, J. Kong, and X. Ren. 2004. Speed–accuracy tradeoff in Fitts’ law tasks—on the equivalency of actual and nominal pointing precision.
 1094 *International journal of human-computer studies* 61, 6 (2004), 823–856.

1095

A EXTRA FIGURES

1097

1098

1099

1100

1101

1102

1103

1104

1105

1106

1107

1108

1109

1110

1111

1112

1113

1114

1115

1116

1117

1118

1119

1120

1121

1122

1123

1124

1125

1126

1127

1128

1129

1130

1131

1132

1133

1134

1135

1136

1137

1138

1139

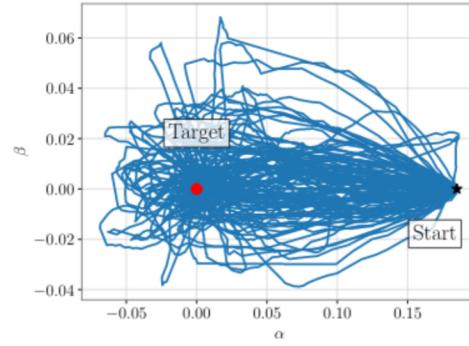
1140

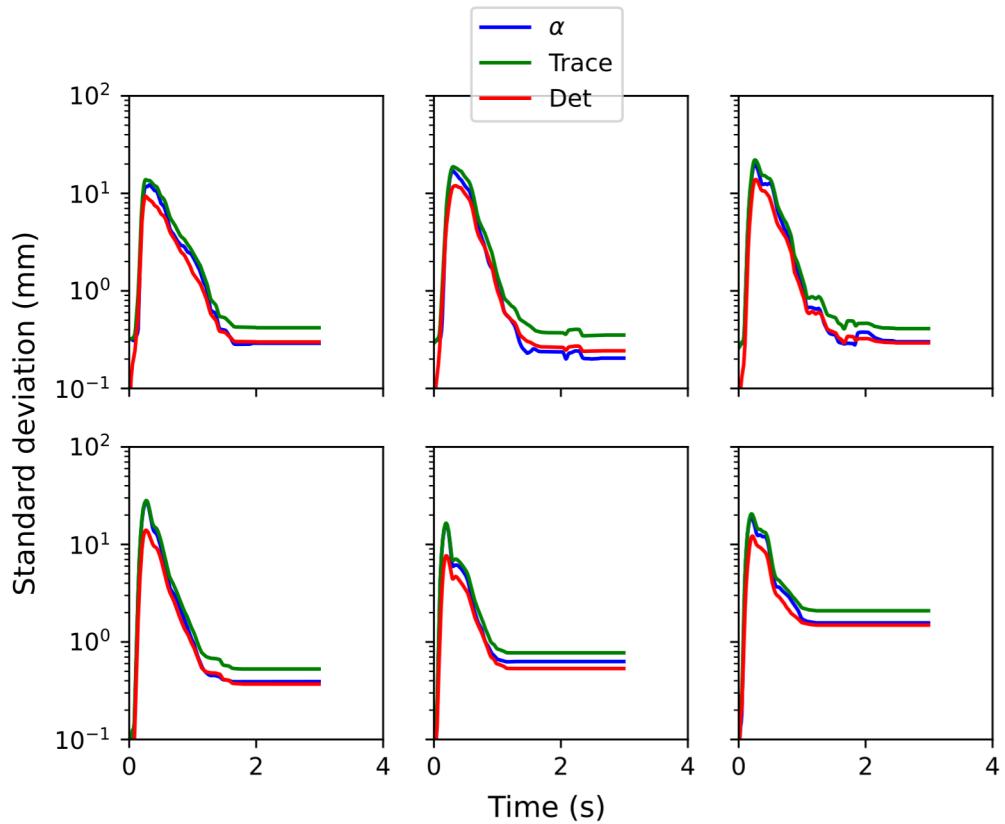
1141

Fig. 8. Left: A set of trajectories from pilot data, decomposed by direction of movement (α , x-axis) and direction orthogonal to movement (β , y-axis). Units for both axes are m .

1143

1144





B EXTRA TABLES

1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196

Table 3. Results of the multilevel fits for PVP features.

	τ	D_τ	σ_0	C
(Intercept)	0.579*** (0.039)	1.987 (1.867)	4.288*** (0.904)	1.658*** (0.257)
D	0.003*** (0.000)	0.248*** (0.021)	0.160*** (0.010)	0.004 (0.002)
W	-0.013*** (0.003)	-0.326 (0.264)	-0.137 (0.126)	0.059* (0.027)
devicemouse	-0.222*** (0.037)			3.064*** (0.136)
devicetouchpad	-0.177*** (0.037)			3.213*** (0.136)
$D \times W$	0.000 (0.000)	0.003 (0.004)	0.000 (0.002)	-0.001 (0.000)
$D \times devicemouse$	-0.003*** (0.000)	0.232*** (0.015)	-0.023** (0.007)	
$D \times devicetouchpad$	-0.003*** (0.000)	0.216*** (0.015)	0.089*** (0.007)	

Significance: *** $\equiv p < 0.001$; ** $\equiv p < 0.01$; * $\equiv p < 0.05$

N = 430, Participant = 12

Table 4. Results of the linear model fits for the Fitts' law model on data from the ISO condition.

	Effective	Nominal
(Intercept)	1.100*** (0.161)	0.708*** (0.131)
IDe/ID	0.350*** (0.042)	0.402*** (0.030)
device: mouse/controller	-0.852*** (0.236)	-0.343 (0.184)
device: touchpad/controller	-0.595* (0.234)	-0.316 (0.184)
IDe/ID \times device: mouse/controller	-0.162** (0.059)	-0.252*** (0.042)
IDe/ID \times device: touchpad/controller	-0.136* (0.058)	-0.165*** (0.042)
R-squared	0.653	0.765
N	210	210

Significance: *** $\equiv p < 0.001$; ** $\equiv p < 0.01$; * $\equiv p < 0.05$