

# Pointing Models for Users Operating Under Different Speed Accuracy Strategies

ANONYMOUS AUTHOR(S)

This work proposes and compares models for pointing tasks, focusing on users operating under different speed-accuracy strategies. Traditional models like Fitts' law capture the relationship between movement time and target difficulty but often overlook strategic aspects and only describe average performance. We address this by 1) capturing the statistical dependence between speed and accuracy 2) developing parametric models that describe the entire distribution of movement times for different user strategies. We namely utilize copulas, a novel approach in HCI, alongside other more established techniques, to model pointing data comprehensively. We use existing datasets to measure dependencies and compare our models. This work advances the knowledge of pointing models and provides the required resources to implement the suggested models.

## ACM Reference Format:

Anonymous Author(s). 2024. Pointing Models for Users Operating Under Different Speed Accuracy Strategies. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 11 pages. <https://doi.org/XXXXXXXXXXXX>

## 1 INTRODUCTION

Pointing towards a target *e.g.*, selecting an icon with a desktop mouse, is one of the fundamental interactions in HCI [2, 32]. Pointing *models* are regularly used to evaluate input techniques [32] and to design interfaces [4]. Classically, HCI researchers model pointing time with Fitts' law [8, 16, 32], and refer to it as *movement time* MT

$$MT = a + b \log_2(1 + D/W), \quad (1)$$

$$\equiv a + b ID. \quad (2)$$

Here,  $D$  and  $W$  are geometric parameters (size and distance of the target),  $ID = \log_2(1 + D/W)$  is the index of difficulty, and  $a$  and  $b$  are parameters estimated from empirical data. Fitts' law captures the speed-accuracy tradeoff of human movement: decreasing  $W$  will lead to larger MT *i.e.*, a higher accuracy can be obtained only at the cost of longer pointing times.

Since Fitts' seminal 1954 work [8], and its introduction by Card and colleagues in HCI in 1978 [2], movement modeling has greatly advanced. While Fitts' law only describes *endpoints*, motor control models now propose description of entire *trajectories* produced during pointing tasks [7, 12, 14, 31, 33]. While Fitts' law only describes *means* of movement times, several models now describe *distribution* of movement times, usually using asymmetric distributions [3, 13, 23, 27, 30, 40].

---

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](mailto:permissions@acm.org).

© 2024 Association for Computing Machinery.

Manuscript submitted to ACM

One aspect that has been less studied in models of pointing is its strategic aspect. Most people will regularly over or under utilize the target width  $W$  [32], creating, according to Zhai and colleagues, *an additional subjective layer of speed-accuracy tradeoff* [39], which the current formulations of Fitts' law do not capture *a priori*. But models that describe how users point under different speed-accuracy strategies are essential when addressing users with varying incentives. For example, people will favor accuracy in contexts where safety matters critically [20].

One notable exception is the Weighted Homographic (WWho) model of Guiard and Rioul [21], which quantitatively relates movement time, movement precision, strategy, and effort (a detailed explanation of the model is provided Appendix C). The WWho model was recently used as a low level control module in hierarchical models of user behavior [22, 25], where a high level controller learns to be more cautious in contexts where the cost to recover from errors is higher. However, we believe the WWho model is not an appropriate tool to describe user performance, since it was built as a model of *ultimate* performance. It does not describe the entire distribution of data, but instead, it describes the limit speed and accuracy performances one can obtain (the *front of performance*, in Guiard and Rioul's words, also see subsection 2.5). In Guiard and Rioul's work, the WWho model is fitted only on the best 10 measures, out of 400! Hence, works that use the WWho model use an overly optimistic model of user behavior.

The goal of this work is to propose and compare candidate models for pointing for users operating under different speed-accuracy strategies. Our constraints are as follows:

- The model should describe the entire distribution of movement times, and not just the means (as in Fitts' law) or extrema (as in WWho, also see [13, 17, 27]). This ensures the model will be widely applicable.
- The model should be parametric, so that it can be easily interpreted, re-used and adapted by others *e.g.*, as a low-level control module as in [22, 25] without too much effort.
- One of the parameters of the model should map directly to the concept of strategy *i.e.*, a bias towards speed or accuracy when producing movements, filling the gap identified beforehand.

In this work, we place the emphasis on characterizing the relationship between speed and accuracy, with particular attention to their (statistical) dependence. When participants prioritize accuracy, there is on average an increase in pointing time, which leads to a wider pointing time distribution [13, 27, 40]<sup>1</sup>. Understanding precisely how this spread increases is central to our analysis. Therefore, we make use of copulas, a statistical tool well-suited for capturing the dependence between variables which to our knowledge has never been used in HCI before, to model the pointing data, among other better-known techniques.

*Notations.*  $\mathcal{N}(\mu, \Sigma)$  refers to the (multivariate) Gaussian distribution with mean  $\mu$  and covariance matrix  $\Sigma$ ,  $\mathcal{E}(\lambda)$  refers to the exponential distribution with mean  $\lambda$ ,  $\mathbb{E}[X]$  refers to the mathematical expectation of  $X$ ,  $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$  refers to the (population) variance of  $X$ .

<sup>1</sup>As hypothesized by [10, p. 132] in analogy to reaction time distributions [34].

## 2 BACKGROUND

### 2.1 Participant strategies and the effective index of difficulty

Fitts' law experiments are usually conducted to investigate/evaluate the speed accuracy tradeoff. To do so, the experimenter manipulates accuracy by modifying  $W$  and observes speed of the response by measuring MT in Equation 1. However, it is well known that this manipulation is imperfect: users underutilize *easy* (low ID) targets and over-utilize *hard* (high ID) targets [19, 39]. A solution, due to Crossman and popularized in HCI by Mackenzie [16] is to compute an effective index of difficulty  $ID_e$  as a *post-hoc* adjustment of difficulty as a function of the effectively measured standard deviation of endpoints  $\sigma$

$$MT = a + b ID_e \quad (3)$$

$$= a + b \log_2 \left( 1 + \frac{D}{4.133\sigma} \right). \quad (4)$$

Several researchers, beginning with Mackenzie [28] have investigated how people perform Fitts' task when instructed to emphasize either speed or accuracy. As expected, participants are able to modify their strategies, as reflected by changes in  $ID_e$  for a given nominal ID level. For example in Mackenzie's experiment, on average participants erred up to 20% in the speed condition, but erred close to 0% in the accuracy condition. Guiard [19] later performed a similar experiment, but without explicit width (participants had to aim towards a line) with 5 instructions (Ultra Fast, Fast, Balanced, Accurate, Ultra Accurate). Other similar experiments include [1, 37]. But Fitts' law, even with  $ID_e$ , does not link strategy,  $ID_e$  and MT *a priori*:  $ID_e$  can only be defined *after* the experiment has been conducted.

### 2.2 Measures of dependence between two random variables

**2.2.1 Measures of association.** Measures of association [29] quantify the strength and direction of the relationship between two variables; they are (in most cases) weaker than measures of dependence in the sense that independence implies an absence of association, but the converse is not true.

**Pearson's  $r$ .** Pearson's  $r$  is a well-known measure of association that captures the *linear relationship* between two random variables. A value of 1 (resp.  $-1$ ) means that the data from the two variables are perfectly aligned with a positive (resp. negative) slope. One well-known weakness of Pearson correlation is that a null value only indicates the absence of a linear association, even though non-linear associations are common.

**Spearman's  $\rho$  and Kendall's  $\tau$ .** Spearman's  $\rho$  between two variables equals Pearson's  $r$  between the *ranks* of the variables; as a result it allows capturing monotonic (increasing or decreasing, which includes linear but many more) relationships between variables. It will have a value of 1 (resp.  $-1$ ) when the order of the sorted values of one variable is the same (resp. opposite) as the other variable. Kendall's  $\tau$  is a different rank correlation metric, which, for our purpose, has the same properties as  $\rho$ .<sup>2</sup>

<sup>2</sup>... but in general does not have the same value [29, Chapter 5]. It is more sensitive to smaller rank inversions or ties. Both  $\rho$  and  $\tau$  can be derived from the same general correlation coefficient [24].

**2.2.2 Bivariate copulas** . Bivariate copulas are a statistical tool that describe the dependence between two random variables [29]. They “couple” [29, Chapter 1] two marginals into a single joint distribution, “parametrically”, where the parameter of the copula controls the strength of the association between the variables. Different types of copula introduce different types of dependence; for example a Gumbel copula introduces so-called upper tail dependence [29], where if one variable takes on a high value, the other is also likely to take on a high value. Identifying the best fitting copula fully describes the dependence structure between two variables, which include linear/non linear and tail dependences. Copulas often encapsulate rank correlations such as  $\tau$  and  $\rho$  *e.g.*, a Gumbel copula of parameter  $\theta$  has  $\tau = \frac{\theta-1}{\theta}$ . To fully characterize data, copulas have to be used in conjunction with the marginals of each variable, which specify their full distribution.

### 2.3 Fitts’ law and regression

The expression for Fitts’ law Equation 4 masks an underlying statistical model that assumes additive, centered noise<sup>3</sup>:

$$MT = a + bID_e + \text{noise (additive)}, \quad \mathbb{E}[\text{noise}] = 0 \text{ (centered)}. \quad (5)$$

In such linear models, Pearson’s  $r$  is commonly used as a goodness-of-fit measure, particularly because in this case,  $r^2 = R^2$ , reflecting the proportion of variance in the dependent variable that is explained by the model and confirming the model’s linearity [26]. However, as shown by Drewes [6] and Gori *et al.* [18], the Pearson correlation between MT and  $ID_e$ ,  $r(MT, ID_e)$ , often yields low values in practice, with some cases showing correlations close to zero [16].

To mitigate variability in MT and increase the goodness-of-fit measure, researchers typically compute the sample average of MT, denoted in the remainder of the work as  $\overline{MT}$ <sup>4</sup>, before calculating  $r$ . Consequently, the model actually considered by researchers is:

$$\overline{MT} = a + bID_e + \text{noise (additive)}, \quad \mathbb{E}[\text{noise}] = 0 \text{ (centered)}, \quad (6)$$

with the goodness-of-fit measure being  $r(\overline{MT}, ID_e)$ .

### 2.4 Asymmetric MT distributions and the Exponentially Modified Gaussian noise model

Rather than averaging out the noise as in Equation 6, recent works model the noise distribution instead, and show that the conditional distribution of MTs for a given ID or  $ID_e$  level is distributed according to an asymmetric, positive (right) skewed [3, 10, 13, 23, 27, 30, 40] distribution. While several distributions may be possible, including gamma, Gumbel and lognormal, we follow Gori and colleagues [10, 13, 27], and Zhao *et al.* [40] who considered an Exponentially Modified Gaussian (EMG) noise model as MT distribution: it satisfies all the mentioned properties, has parameters that are easy to interpret, has shown good performance (including when modeling steering data [35]), and there exist tools to aid in parameter estimation [11, 27].

<sup>3</sup>Because most researchers applying Fitts’ law rely on ordinary least squares (OLS), they also implicitly assume the noise is uncorrelated and homoscedastic, as OLS is only optimal under these conditions.

<sup>4</sup>See [17] for further discussion on considering the sample average.

Gori [10] further showed that a model where the standard deviation of the MT distribution increases linearly with its mean provided a good fit to empirical data:

$$MT = G + E, \quad (7)$$

$$G \sim \mathcal{N}(\beta \mathbf{x}, s) = a + b ID_e + \mathcal{N}(0, s), \quad (8)$$

$$E \sim \mathcal{E}(\lambda \mathbf{x}), \quad (9)$$

$$\mathbf{x} = [1, ID_e]^t, \quad \beta = [a, b], \quad \lambda = [\lambda_0, \lambda_1]. \quad (10)$$

Although the EMG model aforementioned has a quite complex density, it has very simple conditional mean

$$\mathbb{E}[MT|ID_e] = \beta \mathbf{x} + \lambda \mathbf{x}, \quad (11)$$

and conditional variance

$$\text{Var}(MT|ID_e) = s^2 + (\lambda \mathbf{x})^2. \quad (12)$$

Zhao *et al.* [40] also confirmed the quality of the EMG fit with linear conditional mean using a similar model.

## 2.5 Two Fitts' laws

Gori and colleagues [15, 16] have argued that there exist two version of Fitts' law: an average Fitts' law—which we call the “mean” law, computed via linear regression on  $\overline{MT}$  as in Equation 6, and a minimum Fitts' law, which we call the “min” law. Fitts' min law represents the best performance possible, whereas Fitts' mean law represents a more central tendency. The EMG model presented just before can be used to identify the min law [13] via the  $\beta$  parameters.

## 3 DEPENDENCE BETWEEN $ID_e$ AND MT WHEN ACCURACY IS MANIPULATED BY W

In this section, we used an existing dataset by Jude *et al.* [23] (JGP). The JGP dataset was generated using the well-known 2D multi-directional tapping task [32] via the FittsStudy software [36], with the experiment replicated six times over three days, twice a day, for 15 participants. We noticed a bug in the original dataset, where the 2 last replications are identical to the two first, so we discarded the 2 last replications from the dataset, leaving us with 4 replications. In total, we thus have  $60 = 15 \times 4$  individual datasets. We start with a theoretical analysis.

### 3.1 Theoretical analysis of $r(\overline{MT}, ID_e)$ correlation

It has long been established that  $r(\overline{MT}, ID_e)$  is very close to 1 in experiments when accuracy is manipulated by W [32], including the in the JGP dataset [23]. We first investigate the type of dependence between  $ID_e$  and MT that is required to have data that is compatible with Fitts' law *i.e.*,  $r^2(ID_e, \overline{MT})$  close to 1. The dependence between  $ID_e$  and MT is fully specified by the conditional distribution of MT given  $ID_e$   $p(MT|ID_e)$ .

We now give a sufficient condition such for *any* conditional MT distribution to reach  $r^2(ID_e, \overline{MT}) = 1$ .

**PROPOSITION 3.1.** *Any conditional distribution  $p(MT|ID_e)$  will reach  $r^2(ID_e, \overline{MT}) = 1$  if it has a conditional expectation  $\mathbb{E}[MT|ID_e]$  that is linear in  $ID_e$ .*

The proofs is delayed to Appendix B. Below, we give some examples that fulfill this condition

- The EMG distribution [10, 13, 27, 40] Equation 7 has mean value  $\beta\mathbf{x} + \lambda\mathbf{x} = \beta_0 + \lambda_0 + (\beta_1 + \lambda_1)\text{ID}_e$ .
- A gamma [23] distribution with shape  $k$  and scale  $\theta$  has mean value  $k\theta$ , so having  $k$  or  $\theta$  increase linearly with  $\text{ID}_e$  will satisfy the condition.
- A generalized extreme value [3] distribution  $\text{GEV}(\mu, \sigma, s)$  has a mean value that increases linearly with  $\text{ID}_e$  if  $\mu = a + b\text{ID}_e$ .

Hence, a high  $r^2(\text{ID}_e, \overline{\text{MT}})$  is easily reachable by imposing a linear conditional mean, and in practice does not characterize the dependency between MT and  $\text{ID}_e$  well enough, because it only considers the first moment of the conditional distribution<sup>5</sup>. We thus turn to the other classical measures of dependence introduced subsection 2.2.

### 3.2 Measures of association

For each participant and each iteration, we computed Pearson's  $r$ , Spearman's  $\rho$  and Kendall's  $\tau$ . Histograms of the three measures are provided Figure 1, where we notice values typically associated with moderately/strong (rank) correlation. We also computed intraclass correlation coefficients (ICC), which all turned out significant ( $\alpha = 0.011$ ) (results available in the supplementary materials), indicating individual differences in association. Hence, the JGP dataset pictures moderately strong correlations between MT and  $\text{ID}_e$  for all participants, and differences between participants.

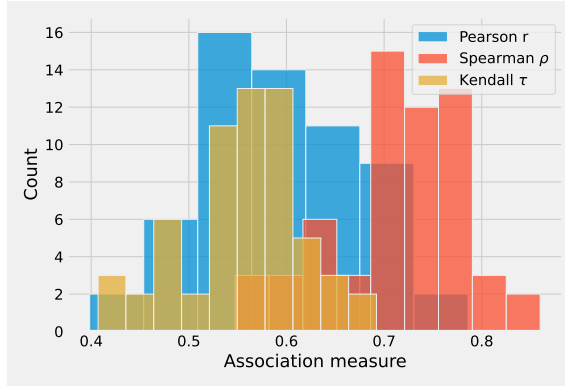


Fig. 1. Histogram of the measures of association (Pearson's  $r$ , Spearman's  $\rho$  and Kendall's  $\tau$ ).

### 3.3 Fitting copulas

The previous measures of association are helpful as a diagnostic, but they are difficult to exploit for creating a generative distribution for movement times. One possibility is to play with variance models for the conditional distribution of MT knowing  $\text{ID}_e$ , and determine the best model based on maximum likelihood estimation or bayesian estimation [40]. Here, we take a more principled approach, and try to match the structure of the dependence between  $\text{ID}_e$  and MT with known copulas.

The methods used to fit copulas for all 60 individual datasets of the JGP dataset is described Appendix A. The log-likelihoods, averaged over the 4 replications, and normalized by the maximum average log-likelihood

<sup>5</sup>One sees from Appendix B that  $r(\text{MT}, \text{ID}_e)$  additionally considers the second moment, but still not the entire distribution.

for each participant are displayed Figure 2. It shows the rotated Gumbel copula to systematically outperform other copulas. The rotated Galambos copula follows very closely; this is not surprising considering Gumbel and Galambos copula are very similar [9]. The t-copula and rotated Hüsler-Reiss (HR) copulas also score decently close to the rotated Gumbel. The Gumbel and t-copulas are represented Figure 6, and their comparison delayed to subsection 4.6.

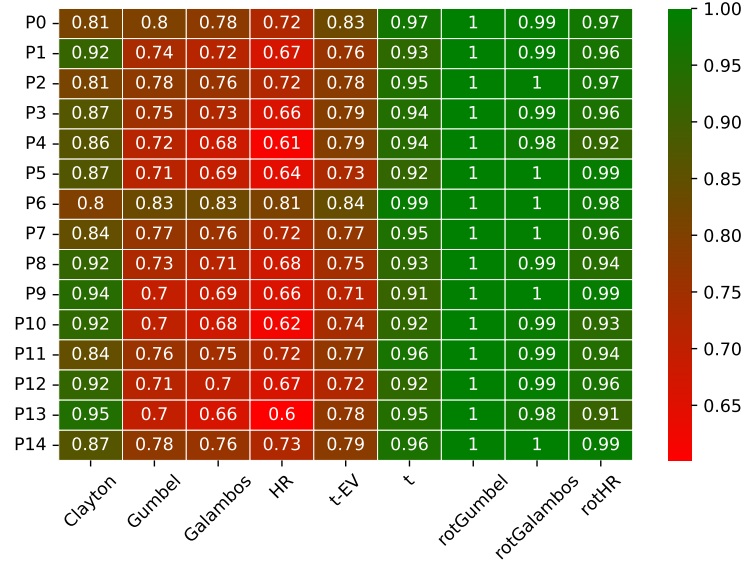


Fig. 2. Log-likelihoods for each copula and for each participant. Closest to 1 (greenest) indicated best fit. The log-likelihoods are the result of averaging the log-likelihoods of the 4 replications. They are then normalized for each participant by dividing the log-likelihood by the maximum log-likelihood over all potential copulas.

## 4 MODELING THE EFFECT OF USER STRATEGY ON ENDPOINTS

The previous section investigated the nature of the dependence between  $ID_e$  and MT for a given strategy *i.e.*, the balanced one participants naturally choose during a pointing experiment. We now investigate the effect of participant-selected strategy on that dependence.

### 4.1 The GO dataset

We investigate the effect of participant strategy on movement time and  $ID_e$ , using the Guiard-Olafsdottir (GO) dataset [19]. The pointing task used to create the GO dataset differs from a typical Fitts task as it lacks a predefined  $W$  (no nominal accuracy). Participants had to aim at a 1-pixel line, with five different strategies, from speed emphasis (allowing large under/overshoots) to precision emphasis (targeting a 1-pixel-wide line). The balanced condition represents a natural strategy, akin to a standard Fitts experiment, with speed and precision conditions falling between extremes [19].

An  $ID_e$  can be calculated retrospectively, similar to Fitts' tasks. Previous research [14, 19] has shown that this data aligns with classical Fitts' law, with mean movement time ( $\overline{MT}$ ) strongly correlating with  $ID_e$ .

## 4.2 Correlation between $ID_e$ and $\overline{MT}$

A summary of the GO dataset, shown in Figure 3, plots  $\overline{MT}$  against  $ID_e$  for all participants, grouped by strategy. Participants covered a wide range of  $ID_e$  (1 to 9 bits), depending on the strategy. The separation between strategies shows effective, yet sometimes inconsistent self-regulation across participants, indicated by considerable overlaps of clusters.  $\overline{MT}$  globally increases with  $ID_e$  across strategies, but also within, as indicated by the orientation<sup>6</sup> of the 95% prediction ellipses from the bivariate Gaussian fit to each strategy, indicating a “local” speed-accuracy tradeoff. The “Speed emphasis” condition is an exception, with  $\overline{MT}$  and  $ID_e$  uncorrelated: when participants go really fast, accuracy actually becomes independent of local changes in speed.

## 4.3 Association per strategy

We also computed association measures for the GO dataset, per strategy, or aggregated for all strategies. The figures are available in the supplementary materials, and are not reproduced here:  $r$ ,  $\rho$  and  $\tau$  have similar values to the ones of the JGP dataset (see Figure 1), and the measures of associations are close to null for the speed emphasis condition.

## 4.4 Per-strategy bivariate Gaussian model for $\overline{MT}$ and $ID_e$

We model the joint increase of  $\overline{MT}$  and  $ID_e$  with a bivariate Gaussian model, for each of the five strategies of the GO dataset. We chose the bivariate model because it is simple, easy to interpret, it makes the dependence between  $\overline{MT}$  and  $ID_e$  completely specified via Pearson’s  $r$ , and is expected to fit reasonably well given the visual appearance of the clusters. The fit for each strategy is displayed by drawing the associated 95% prediction ellipsis, see Figure 3.

$$\begin{pmatrix} ID_e \\ \overline{MT} \end{pmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \mu_i \\ \mu_t \end{bmatrix}, \begin{bmatrix} \sigma_i^2 & r\sigma_i\sigma_t \\ r\sigma_i\sigma_t & \sigma_t^2 \end{bmatrix} \right). \quad (13)$$

When the strategy emphasizes precision, the mean of the Gaussian vector increases, as well as the size of the prediction ellipsis; however, except for the Speed emphasis condition, its orientation stays relatively constant, suggesting that the correlation between the  $ID_e$  and  $\overline{MT}$  may also be.

## 4.5 Bivariate Gaussian model for $\overline{MT}$ and $ID_e$ parametrized by strategy

To represent strategy numerically rather than ordinally, we map the five instructions to equally spaced points on the interval  $[-1,1]$ <sup>7</sup>, with 0 as the balanced strategy, -1 for speed emphasis, and 1 for precision emphasis. We then assessed how the five parameters of the bivariate Gaussian model Equation 13 aligned with this numerical strategy, by fitting a linear model on each parameter. The fits are displayed Figure 4 and their summaries can be found in the supplementary materials.

We observe that:

<sup>6</sup>The covariance matrix of a Gaussian bivariate is the rightmost term of Equation 13 where  $\rho$  is the correlation between the two components. The angle of the prediction ellipsis is given by  $\frac{1}{2} \arctan \left( \frac{2\rho\sigma_i\sigma_t}{\sigma_i^2 - \sigma_t^2} \right)$ .

<sup>7</sup>The choice here may seem naïve, and we could have mapped strategy to the set of real values via  $ID_e$ . However, our choice has two advantages: the mapping can be defined before movements are performed, and does not depend on each participant.



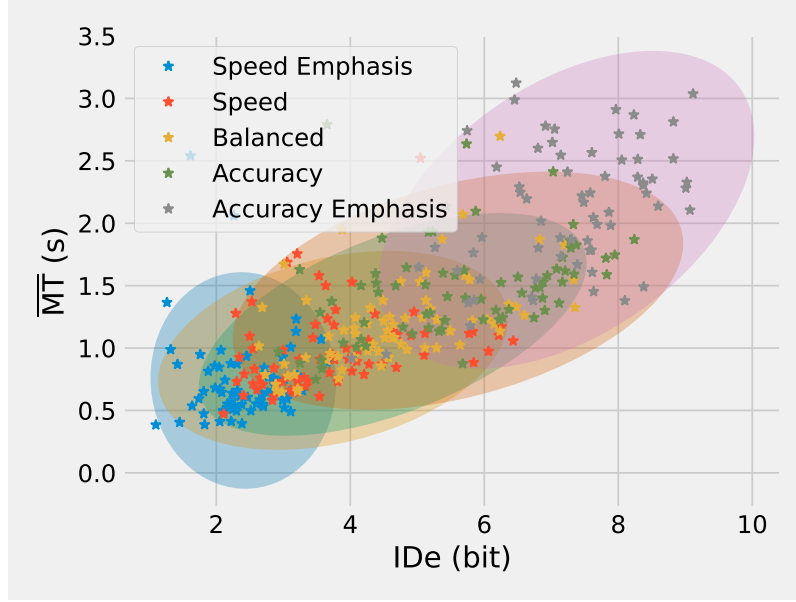


Fig. 3.  $\overline{MT}$  against  $ID_e$  for all participants and strategies for the GO dataset where the color indicates strategy. A bivariate Gaussian fit per strategy is also displayed through the associated 95% prediction ellipsis.

- The linear increase of the mean ( $\mu_i, \mu_t$ ) with numerical strategy is visually clear, and supported by high goodness of fit.
- The increase of  $\sigma_i$  and  $\sigma_t$  with numerical strategy can also reasonably be described by a linear model, although evidence is less strong than for the mean.
- The correlation between  $\overline{MT}$  and  $ID_e$  can also be modeled by a linearly increasing model. However, the slope is not statistically significant (see Supplementary materials) and it is visually clear that speed and accuracy are uncorrelated. As another model, one may also consider a null correlation for the speed emphasis condition, and a constant correlation for the other conditions; we computed the mean  $r$  for the 4 other conditions and obtained  $r = 0.44$ . We choose the latter model, since the speed emphasis condition is unlikely to be of interest in HCI applications (in contrast with the accuracy emphasis condition).

#### 4.6 Fitting copulas

Similarly to subsection 3.3, we fit copulas on the GO dataset, per strategy and per participant. The results contrast somewhat with what we found in section 3: this time the t-copula comes ahead; the rotated Gumbel does not even come as a close second. This is true when looking on the basis of participants, aggregating over strategies (left panel of Figure 5) or per strategy, aggregating over participants (right panel of Figure 5).

Contour plots of the t-copula and rotated Gumbel copula, with various marginals, are shown in Figure 6 for parameters fit on a random subset from the JGP dataset. The left panel illustrates a rotated Gumbel copula with standard normal marginals, to be compared with the middle right panel displaying the t-copula

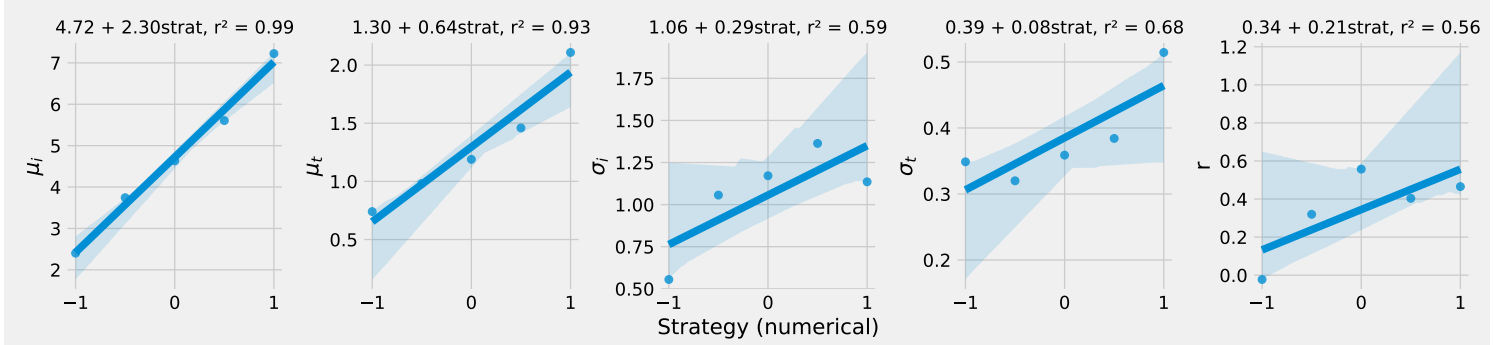


Fig. 4. The parameters of the mean and covariance of the bivariate Gaussian model defined in Equation 13 plotted against the five numerical strategies. The associated linear fits and  $r^2$  are displayed above each plot.

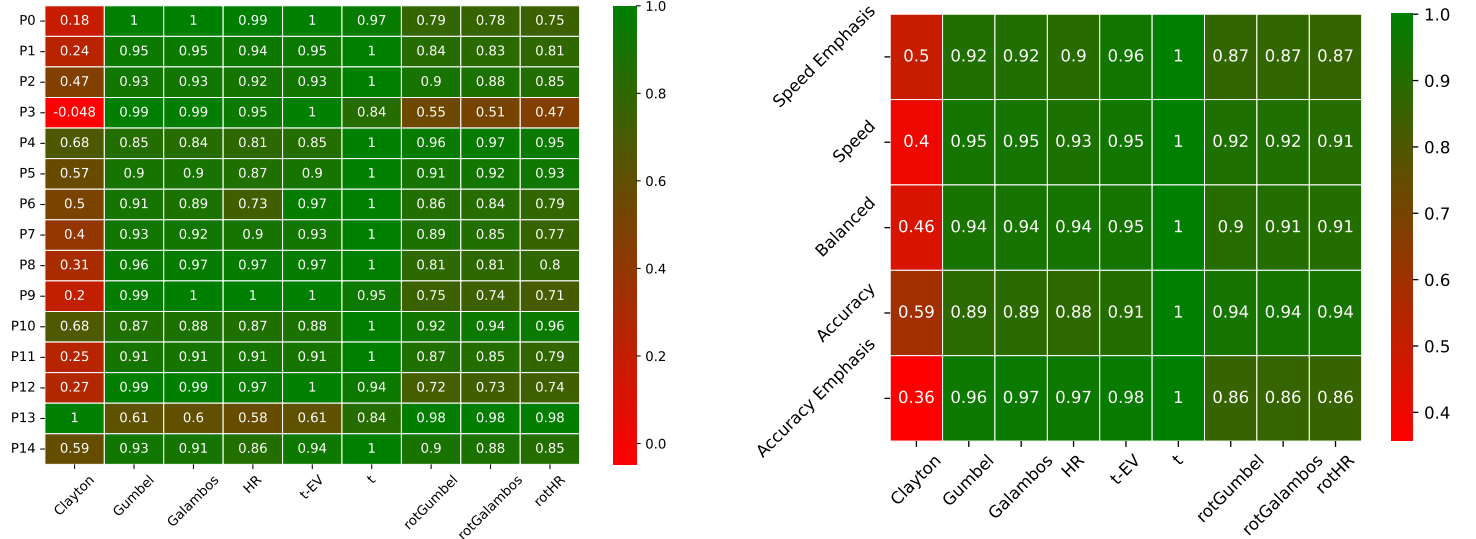


Fig. 5. Log-likelihoods of the tested copulas, normalized for each row by the best fitting copula. Left panel: Comparison on the basis of participants, when data is aggregated over all strategies. Right panel: Comparison on the basis of strategies, when data is aggregated over all participants.

with the same marginals. The t-copula is symmetric and implies equal (positive) tail dependencies, whereas the rotated Gumbel only captures lower tail dependencies<sup>8</sup>, making it initially surprising that these distinct copulas could both emerge as best candidates [5, 29].

However, examining the middle left and right panels, which show the rotated Gumbel and t-copulas for marginals estimated from the JGP dataset, clarifies this. These contour plots model the joint distribution  $p(MT, ID_e)$  and appear quite similar, with increasing conditional variance for higher  $ID_e$  values. A noticeable difference is that the t-copula's joint distribution assigns higher probabilities to MT for large  $ID_e$  values

<sup>8</sup>Lower (upper) tail dependencies refer to the correlation between only low (high) values of MT and  $ID_e$ .

compared to the rotated Gumbel. Hence, with the particular marginals of the JGP dataset, the discrepancy in upper tail dependencies vanishes.

Given the similarity in the resulting joint distributions and the t-copula’s effectiveness in modeling the JGP dataset (see Figure 2), we will use the t-copula to express the dependency between MT and  $ID_e$  for the remainder of this work.

We also evaluated if there were any participant differences or strategy difference in terms of the parameters of the estimated t-copula, using the GO dataset. In fact, we found no differences for both parameters; more details are available in the supplementary materials. On average, the two parameters of the t-copula are  $\rho_1 = 0.75$  and  $\nu \gg 100$ , indicating that for many cases, the t-copula actually reduces to a Gaussian copula.<sup>9</sup>

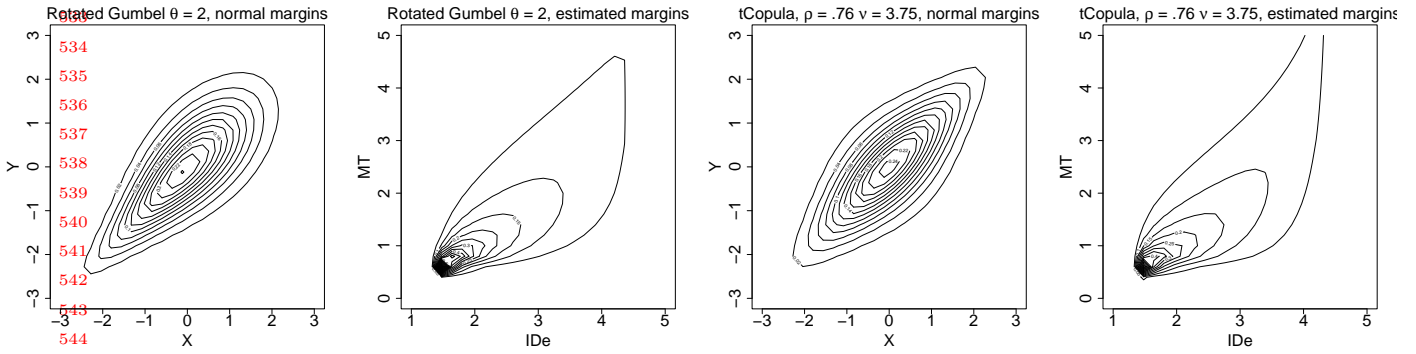


Fig. 6. Density functions for four joint distributions constructed using various copulas and marginals. Left: the rotated Gumbel copula, with standard normal margins. Middle Left: the same copula with marginals estimated from the GO dataset. This represents a statistical model of the pointing data. Middle Right: the t-copula with standard normal margins. Right: the t-copula with marginals estimated from the GO dataset. This represents a competing statistical model for pointing data.

## 5 GENERATING POINTING DATA FOR DIFFERENT SPEED-ACCURACY STRATEGIES: RATIONALE AND METHODS FOR THE DIFFERENT MODELS

We are now ready to return to our initial goal: generate pointing data that depend on user strategies, according to a parametric model. We suggest three methods; everywhere,  $\Theta$  refers to a vector of parameters which includes, but is not limited to, strategy  $\mathcal{S}$ ,  $p_{\Theta}(ID_e, MT)$  to the joint distribution between  $ID_e$  and MT parametrized by  $\Theta$ ,  $p_{\Theta}(MT|ID_e)$  to the conditional distribution of MT given  $ID_e$ .

### 5.1 Model 1: Joint distributions using Copulas.

*The model.* The first method directly defines  $p_{\Theta}(ID_e, MT)$  in terms of copulas: As explained in subsection 2.2.2, the copula  $C_{\Theta}$  links the marginals  $F_{\Theta}(ID_e)$  and  $G_{\Theta}(MT)$  into the joint  $p_{\Theta}(ID_e, MT)$

$$p_{\Theta}(ID_e, MT) = C_{\Theta}(F_{\Theta}(ID_e), G_{\Theta}(MT)). \quad (14)$$

<sup>9</sup>The parameter  $\nu$  stands for the degrees of freedom of the t-copula; on average we found  $\nu = 4056$ . Just like a t-distribution converges to a Gaussian for many degrees of freedom, so does the t-copula converge to the Gaussian copula for high values of  $\nu$ . [5, 29]. It should be noted that the estimation of the average  $\nu$  is somewhat unreliable for the current datasets, namely because the arithmetic average we computed is not necessarily meaningful when there is large spread in the data as in this case, and because the number of data points is probably too limited to catch the fine structure of dependencies. When estimated on the entire GO dataset, we find  $\nu = 16$ .

Pointing data can then be obtained by sampling directly from the joint. This method assumes access to the marginals, which can be estimated from empirical data, or given/assumed.

*Accounting for different strategies.* In the previous section, we did not find differences in the t-copula for different strategies. Hence, we only need to exploit the fact that the marginal  $ID_e$  distribution, obtained by marginalizing Equation 13, is a factor of strategy  $\mathcal{S}$  *i.e.*,

$$ID_e \sim \mathcal{N}(4.72 + 2.3\mathcal{S}, 1.06 + 0.39\mathcal{S}). \quad (15)$$

We can then sample  $ID_e$  values by drawing from the Gaussian Equation 15, and use the conditional<sup>10</sup> t-copula to sample the associated MT values.

*Inputs of the model.* To operationalize this model, one needs

- (1) The parameters of the t-copula. We found  $\rho_1 = .69$ ,  $\nu = 16.9$  by estimating the t-copula on the entire GO dataset. Without access to routines to estimate copula parameters,  $\rho_1$  can be deduced from Kendall's  $\tau$  using  $\rho_1 = \sin(\tau \frac{\pi}{2})$ , and  $\nu$  can be deduced from the upper and lower tail dependence coefficients using a more complex formula [5, Section 3].
- (2) The parameters of the marginal distributions for MT and  $ID_e$ . These can be simply estimated using any popular statistical package. For MT, several asymmetric distributions may be valid; we prefer the EMG distribution as explained subsection 2.4. For the GO dataset, we found  $\beta = 0.53$ ,  $\sigma = 0.19$ , and  $\lambda = .75$ . Note that we use the “pure” EMG distribution here, and not the one with linear conditional mean Equation 11.
- (3) The levels and number of repetitions may also be given to generate repeated measures.

*Pro's and con's.* As explained subsection 2.2.2, using copulas allows specifying directly the level of association between variables independently of the marginals. Hence, one can couple arbitrary marginals for  $ID_e$  (*e.g.*, to fit a particular task) and MT (*e.g.*, using another asymmetric distribution for MT) with the same t-copula once it has been fitted. This first model based on copulas thus provides a principled way of identifying dependencies between  $ID_e$  and MT, and enough flexibility in exploiting that dependency for generating data, including repeated measurements.

However, two negative points have to be highlighted. First, copulas have a good track record in modeling tail dependencies, but for pointing data the lower tail of the MT distribution should associate strongly with the entire  $ID_e$  distribution [16] —not only its tail—. Second, the tools to work with copulas remain somewhat underdeveloped, which may add barriers for some researchers. As an example, in writing this paper, we have used an R library for copulas, whereas the rest of the work was performed using Python.

## 5.2 Model 2: Conditional distribution using EMG.

*The model.* The second method directly leverages the EMG model of subsection 2.4. For each point, a value of  $ID_e$  is selected by sampling from the marginal  $F_\theta(ID_e)$ ; then a value of MT is selected by sampling from the conditional distribution

$$p_\Theta(MT|ID_e) = \text{EMG}(MT; ID_e, \Theta). \quad (16)$$

<sup>10</sup>The closed form formula of the conditional copula can be computed from any copula, see [29] and Appendix A.

*Accounting for different strategies.* Similar to the previous model, we can directly draw  $ID_e$  values from the Gaussian Equation 15, and sample MT from the conditional EMG distribution.

*Inputs of the model.* To operationalize this model, one needs the parameters of the EMG distribution described Equation 7 – Equation 11. We estimated these values for the GO dataset to be  $\beta_0 = 0.08$ ,  $\beta_1 = .14$ ,  $\sigma = 0.18$ ,  $\lambda_0 = .17$ ,  $\lambda_1 = .08$ . The levels and number of repetitions may also be given to generate repeated measures.

*Pro's and con's.* This method will produce data that resembles empirical data by construction, since it relies directly on the EMG distribution. One of the main weaknesses of this method is that it does not allow controlling the association between  $ID_e$  and MT independently of the EMG parameters, as shown in Appendix B: the Pearson correlation  $r$  is entirely determined by the EMG parameters. Additionally, as explained subsection 3.1 it will produce values of  $r(\overline{MT}, ID_e)$  much closer to 1 than what we found Figure 3 and Figure 4.

### 5.3 Model 3: Joint distribution of $(ID_e, \overline{MT})$ .

*The model.* This method exploits the Gaussian bivariate model Equation 13. A couple  $(ID_e, \overline{MT})$  is first sampled from the joint. Then, because the EMG distribution parameters are related to  $\overline{MT}$  by

$$\overline{MT} = \beta \mathbf{x} + \lambda \mathbf{x} = \beta_0 + \beta_1 ID_e + \lambda_0 + \lambda_1 ID_e, \quad (17)$$

we “correct”  $\lambda_1$ , ensuring it remains positive to keep the general trend of increasing conditional mean and variance, as indicated *e.g.*, by the fitted copulas Figure 6.

$$\lambda_1 = \max(0, \frac{\overline{MT} - \beta_0 - \beta_1 ID_e}{ID_e}). \quad (18)$$

We don't correct  $\beta$  since this would alter Fitts' min law parameters, and don't alter  $\lambda_0$  to ensure a minimum variance. Movement times are then sampled from the EMG with corrected parameters. Notice in Equation 18 that we force  $\lambda_1$  positive; this is to keep an increasing variance. This comes at the cost of occasional departures from the correction.

*Accounting for different strategies.* Here, we draw samples directly from the bivariate Gaussian Equation 13 with

$$\mu_i = 4.72 + 2.3 \mathcal{S} \quad (19)$$

$$\mu_t = 1.3 + .64 \mathcal{S} \quad (20)$$

$$\sigma_i = 1.06 + 0.29 \mathcal{S} \quad (21)$$

$$\sigma_t = .39 + .08 \mathcal{S} \quad (22)$$

$$r = \begin{cases} 0 & \text{if } \mathcal{S} = -1 \\ .44 & \text{else} \end{cases} \quad (23)$$

*Inputs of the model.* Same as for Model 2.

*Pro's and con's.* Similarly to the previous method, this method will produce data according to the EMG distribution, making it realistic. Contrary to method 2, we do control the association between  $\overline{MT}$  and  $ID_e$  in this third method. However, the needed correction is sometimes not fully applied (when  $\lambda_1 = 0$  in Equation 18, part of the  $\overline{MT}$  has not been absorbed into  $\lambda_1$ ); this method also does not really control how the movement time distribution is modified beyond the ability to fit  $\overline{MT}$ .

## 6 MODEL COMPARISON

### 6.1 Consistency

We first start by investigating the consistency of the different methods *i.e.*, how close the data produced by the generative models are to the reference dataset. We consider a dataset with large sample size so that our measures don't depend much on random sampling and are interpretable from a single simulation<sup>11</sup>. The generated datasets are compared based on the following metrics:

- (1) association measures:  $r(MT, ID_e)$ ,  $\rho(MT, ID_e)$ ,  $\tau(MT, ID_e)$ ,
- (2) Fitts' law measures: parameters of the "mean" and the "min" law,
- (3) estimation of the joint probability  $p(ID_e, \overline{MT})$  through  $r(\overline{MT}, ID_e)$ ,
- (4) computation of ISO-throughput.

We use ISO-throughput [32], also known as the "mean-of-means" throughput, since it is used in practice by some HCI researchers as a one dimensional score to evaluate overall pointing performance. Hence, it gives a concise way of measuring differences between the GO dataset and the generated ones, which are expected to be interpretable and meaningful to HCI researchers. It should be noted that, for maximum consistency, we consider the same  $ID_e$  values as those from the original dataset for the generated data.<sup>12</sup>

The results are summarized Figure 7. Overall, the differences between the models are minimal, as all models produce data that closely match the metrics of the original dataset. When considering measures of association, all models align well with the original dataset. All models lead to slightly underestimated ISO throughput, with a maximum difference of 0.25 bit/s for Model 1. In terms of Fitts' law, Model 1 maintains the mean law (OLS) but distorts the minimum law (EMG), resulting in a higher intercept and a smaller slope. Model 2, on the other hand, preserves the minimum law but slightly skews the mean law, yielding a higher intercept and lower slope. Model 3 comes closest to replicating the original dataset, with both laws nearly aligned; only the intercept in the EMG fit is slightly underestimated. All models lead to inflated values of  $\bar{r} = r(\overline{MT}, ID_e)$ . For Model 2, this is expected due to its construction based on a linear conditional expectation. For Model 1, the high values of  $\bar{r}$  are more surprising, but likely due to the symmetric nature of the t-copula. Model 3 is the only one that reduces  $\bar{r}$ , although it still does not reach the target  $\bar{r}$ , presumably due to the enforced positive  $\lambda_1$ .<sup>13</sup>

<sup>11</sup>The data were generated with a fixed seed (1234). Replications with two other seeds (777, 999) and no seeds are available in the supplementary materials. The conclusions for these replications remain the same as for the results reported in the paper, indicating the sample size of the dataset was indeed large enough

<sup>12</sup>In other words there is no  $ID_e$  variability, only MT variability in these generated datasets. All 3 models can accommodate this: for Model 1 by sampling from the conditional copula, for Model 3 by sampling from the conditional Gaussian bivariate, and for Model 2 trivially.

<sup>13</sup>Indeed, the supplementary materials demonstrate a  $\beta_0$  correction instead of a  $\lambda_1$  correction that is able to achieve the same  $\bar{r}$  as the original dataset, but this comes at the expense of significantly poorer performance in other metrics.

A final observation concerns the lower tail of the joint distribution. For low  $ID_e$  values, all models tend to generate MT values that are too low. Models 2 and 3 are particularly problematic in this regard, as they even predict some negative MT values.

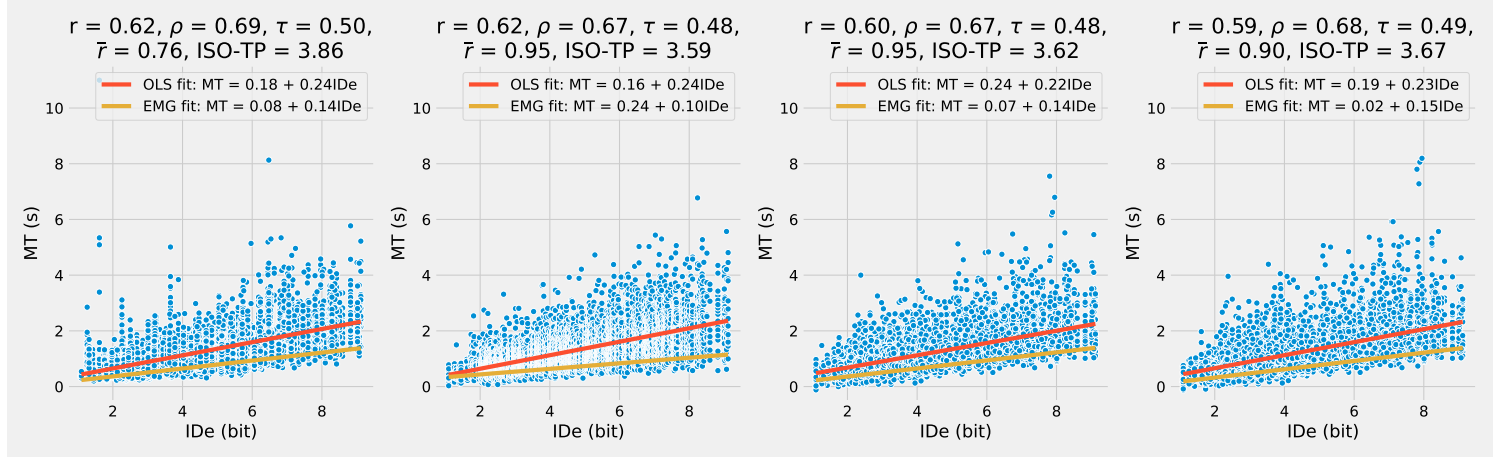


Fig. 7. Consistency as evaluated by the criterion of subsection 6.1. The original dataset (left panel) is compared to data produced by Model 1 (middle left panel), Model 2 (middle right panel) and Model 3 (right panel).

## 6.2 Comparisons of model parametrized by strategy

We have evaluated the consistency of the models for copying an existing dataset, which for all models was good. Hence, one approach that will work to generate pointing data per strategy is to fit a model per strategy, and work with 5 different models. However, as stated in our initial goal, we want a single model parametrized by strategy. We thus used the method explained in each *Accounting for different strategies* paragraph in section 5 to generate datasets from for each strategy parameter. These five datasets were then aggregated to evaluate the fit of the linear models described Figure 4. The resulting fits are shown Figure 8. Model 1 (top panel) fits well on the  $ID_e$  dimension ( $\mu_i, \sigma_i$ ), but fails on the  $\overline{MT}$  dimension ( $\mu_t, \sigma_t$ ), and the Pearson  $r$ . In contrast, Model 2 and 3 (middle and bottom panels) are a lot closer to Figure 4, with Model 3 on the whole superior in about every metric.

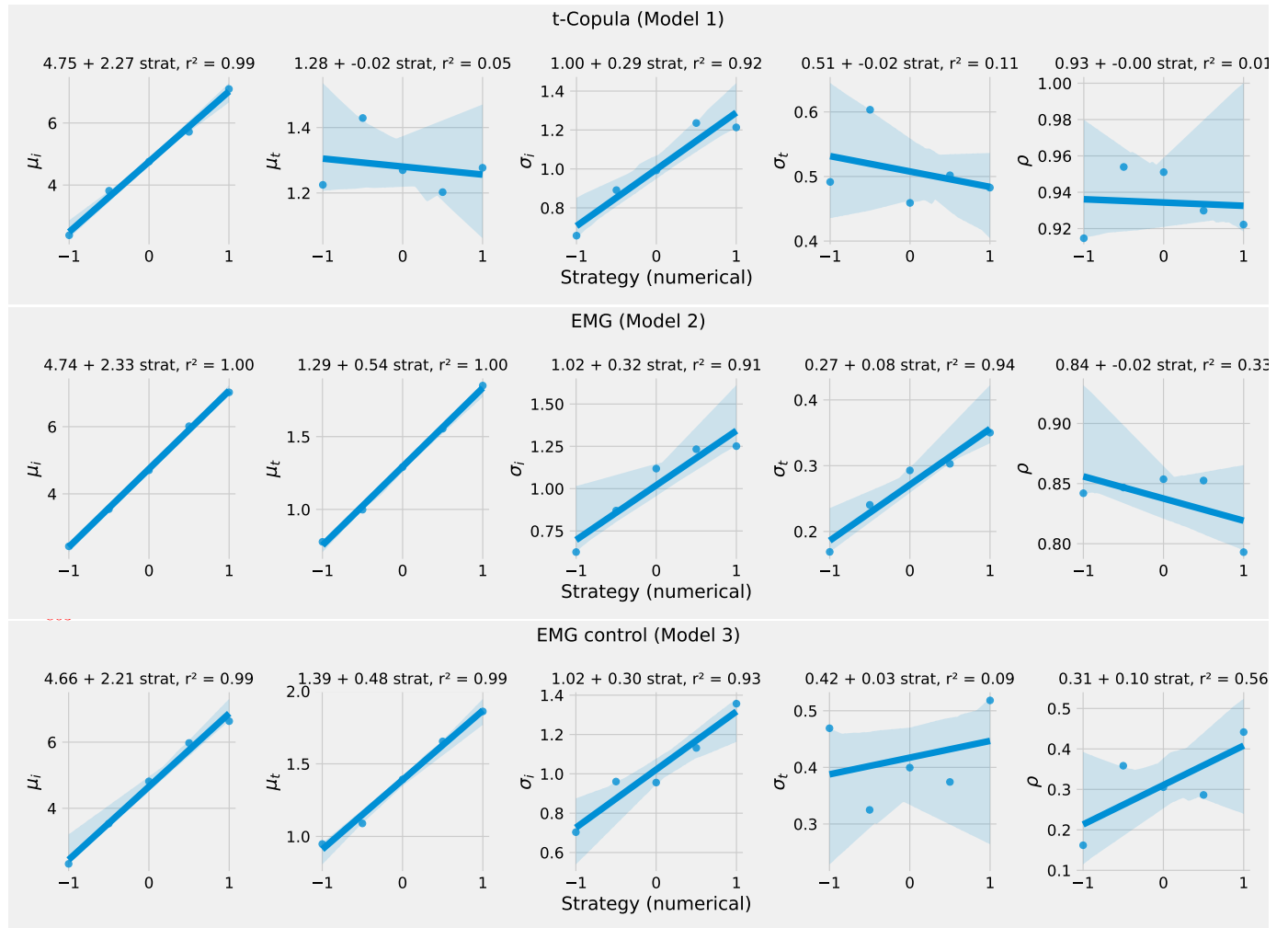


Fig. 8. An evaluation of the linear parametrization of the bivariate Gaussian model of Equation 13. Each subpanel reads like those of Figure 4. Top panel: evaluation for Model 1. Middle panel: evaluation for Model 2. Bottom panel: evaluation for Model 3.

## 7 CONCLUSION AND FUTURE WORK

We presented a comprehensive analysis of user pointing behaviors under different speed-accuracy strategies, leveraging two existing datasets. Based on this analysis, we proposed and compared three generative models that can be parametrized by strategy.

Our contributions are:

- **Introduction of Copulas to HCI:** We introduced the statistical tool of copulas to the field of HCI and demonstrated how it can be used to describe dependencies in pointing data. Given their flexibility,



copulas have the potential to model dependencies parametrically across various types of interaction data.

- Advancing Knowledge of Pointing Models: We enhanced the understanding of pointing models by describing the joint distributions of the Index of Difficulty ( $ID_e$ ) and Movement Time (MT) for various strategies, including with a new theoretical result. We also demonstrated that mapping strategies onto the interval  $[-1,1]$  allows for straightforward linear parametrization.
- Proposing and Evaluating Generative Models: We proposed and evaluated three models for generating realistic pointing data using criteria relevant to HCI researchers. Our findings suggest that no single model should be considered universally superior. Model 3 is optimal for replicating empirical data when average values (such as  $ID_e$  and  $\overline{MT}$ ) are significant. Model 2 is preferable for applications focusing on the “min” Fitts’ law. Model 1 is most suitable when low  $ID_e$  values are of interest.

To ensure a better dissemination of these ideas, we have highlighted the parameters used to generate realistic data available in the paper that could be reused directly by other researchers; the code used for this paper is available at the following anonymized [repository](#) and will be ported to a library if the paper goes to next stage.

For future research, we recommend exploring hierarchical pointing models as proposed in [40]. Notably, our supplementary materials reveal considerable inter-participant differences in  $ID_e$  distributions across strategies: some participants display highly consistent distributions, indicating strong internal strategy adherence, while others perform faster conditions more precisely than accurate ones. Similarly, there are notable differences among participants in the variability of MTs.

## REFERENCES

- [1] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2022. Effective Throughput Analysis of Different Task Execution Strategies for Mid-Air Fitts’ Tasks in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 28, 11 (2022), 3939–3947.
- [2] Stuart K Card, William K English, and Betty 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.
- [3] Olivier Chapuis, Renaud Blanch, and Michel Beaudouin-Lafon. 2007. Fitts’ law in the wild: A field study of aimed movements. (2007).
- [4] Andy Cockburn, Carl Gutwin, and Saul Greenberg. 2007. A predictive model of menu performance. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 627–636.
- [5] Stefano Demarta and Alexander J McNeil. 2005. The t copula and related copulas. *International statistical review* 73, 1 (2005), 111–129.
- [6] Heiko Drewes. 2013. A lecture on Fitts’ law. Retrieved January 5 (2013), 2017.
- [7] Florian Fischer, Mirosław Bachinski, Markus Klar, Arthur Fleig, and Jörg Müller. 2021. Reinforcement learning control of a biomechanical model of the upper extremity. *Scientific Reports* 11, 1 (2021), 14445.
- [8] Paul 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.
- [9] Christian Genest and Johanna G Nešlehová. 2017. When Gumbel met Galambos. In *Copulas and Dependence Models with Applications: Contributions in Honor of Roger B. Nelsen*. Springer, 83–93.
- [10] Julien Gori. 2018. *Modeling the speed-accuracy tradeoff using the tools of information theory*. Ph.D. Dissertation. Université Paris-Saclay (ComUE).
- [11] Julien Gori. 2024. emgregs. <https://pypi.org/project/emgregs/>. Python package.
- [12] Julien Gori and Quentin Bellut. 2023. Positional Variance Profiles (PVPs): A New Take on the Speed-Accuracy Trade-off. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [13] Julien Gori and Olivier Rioul. 2019. Regression to a linear lower bound with outliers: An exponentially modified Gaussian noise model. In *2019 27th European Signal Processing Conference (EUSIPCO)*. IEEE, 1–5.

- [14] 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.
- [15] Julien Gori, Olivier Rioul, and Yves Guiard. 2017. To Miss is Human: Information-Theoretic Rationale for Target Misses in Fitts' Law. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 260–264.
- [16] Julien Gori, Olivier Rioul, and Yves Guiard. 2018. Speed-accuracy tradeoff: A formal information-theoretic transmission scheme (fitts). *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 5 (2018), 1–33.
- [17] Julien Gori, Olivier Rioul, Yves Guiard, and Michel Beaudouin-Lafon. 2017. One fitts' law, two metrics. In *IFIP Conference on Human-Computer Interaction*. Springer, 525–533.
- [18] Julien Gori, Olivier Rioul, Yves Guiard, and Michel Beaudouin-Lafon. 2018. The perils of confounding factors: How Fitts' law experiments can lead to false conclusions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–10.
- [19] Yves Guiard and Halla B Olafsdottir. 2011. On the measurement of movement difficulty in the standard approach to Fitts' law. *PLoS one* 6, 10 (2011), e24389.
- [20] 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*. 1619–1628.
- [21] Yves Guiard and Olivier Rioul. 2015. A mathematical description of the speed/accuracy trade-off of aimed movement. In *Proceedings of the 2015 British HCI conference*. 91–100.
- [22] Jussi Jokinen, Aditya Acharya, Mohammad Uzair, Xinhui Jiang, and Antti Oulasvirta. 2021. Touchscreen typing as optimal supervisory control. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–14.
- [23] Alvin Jude, Darren Guinness, and G Michael Poor. 2016. Reporting and Visualizing Fitts's Law: Dataset, Tools and Methodologies. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2519–2525.
- [24] Maurice George Kendall. 1948. Rank correlation methods. (1948).
- [25] Thomas Langerak, Sammy Christen, Mert Albaba, Christoph Gebhardt, and Otmar Hilliges. 2022. MARLUI: Multi-Agent Reinforcement Learning for Adaptive UIs. *arXiv preprint arXiv:2209.12660* (2022).
- [26] Joseph Lee Rodgers and W Alan Nicewander. 1988. Thirteen ways to look at the correlation coefficient. *The American Statistician* 42, 1 (1988), 59–66.
- [27] Yanxi Li, Derek S Young, Julien Gori, and Olivier Rioul. 2024. A novel mixture model for characterizing human aiming performance data. *Statistical Modelling* (2024), 1471082X241234139.
- [28] I Scott MacKenzie and Poika Isokoski. 2008. Fitts' throughput and the speed-accuracy tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1633–1636.
- [29] Roger B Nelsen. 2006. *An introduction to copulas*. Springer.
- [30] Karin Nieuwenhuizen and Jean-Bernard Martens. 2016. Advanced modeling of selection and steering data: beyond Fitts' law. *International Journal of Human-Computer Studies* 94 (2016), 35–52.
- [31] Matteo Saveriano, Fares J Abu-Dakka, Aljaž Kramberger, and Luka Peternel. 2023. Dynamic movement primitives in robotics: A tutorial survey. *The International Journal of Robotics Research* 42, 13 (2023), 1133–1184.
- [32] 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.
- [33] Emanuel Todorov and Michael I Jordan. 2002. Optimal feedback control as a theory of motor coordination. *Nature neuroscience* 5, 11 (2002), 1226–1235.
- [34] Eric-Jan Wagenmakers and Scott Brown. 2007. On the linear relation between the mean and the standard deviation of a response time distribution. *Psychological review* 114, 3 (2007), 830.
- [35] Michael Wang, Hang Zhao, Xiaolei Zhou, Xiangshi Ren, and Xiaojun Bi. 2021. Variance and Distribution Models for Steering Tasks. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. 1122–1143.
- [36] Jacob O Wobbrock, Kristen Shinohara, and Alex 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*. 1639–1648.
- [37] Shota Yamanaka. 2024. Relative Merits of Nominal and Effective Indexes of Difficulty of Fitts' Law: Effects of Sample Size and the Number of Repetitions on Model Fit. *International Journal of Human-Computer Interaction* (2024), 1–18.
- [38] Jun Yan. 2007. Enjoy the joy of copulas: with a package copula. *Journal of statistical software* 21 (2007), 1–21.
- [39] Shumin Zhai, Jing Kong, and Xiangshi Ren. 2004. Speed-accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International journal of human-computer studies* 61, 6 (2004), 823–856.
- [40] Hang Zhao, Sophia Gu, Chun Yu, and Xiaojun Bi. 2022. Bayesian hierarchical pointing models. In *Proceedings of the 35th annual acm symposium on user interface software and technology*. 1–13.

## A FITTING COPULAS

We considered copulas from the most widely recognized families: elliptical (*e.g.*, the  $t$ -copula), Archimedean (*e.g.*, Clayton, Gumbel), and extreme value (*e.g.*, HR, Galambos,  $t$ -EV), as well as their rotated variants<sup>14</sup>. We utilized the R `copula` package [38] to fit these candidate copulas; copulas are estimated based on maximum likelihood estimation, and the marginals are estimated with *empirical* maximum likelihood estimation. For generating repeated measures at prescribed  $ID_e$  levels, we wrote custom code that leverages the conditional copula.

## B PROOF OF PROPOSITION 3.1

In this proof, we compute correlation coefficients between MT and  $ID_e$ , and  $\overline{MT}$  and  $ID_e$ . Interestingly, we show any distribution with a linear conditional expectation leads to  $r^2(\overline{MT}, ID_e) = 1$  whatever its conditional variance. Note that in this proof, we use the mathematical expectation (*i.e.*, population averages) rather than sample averages — in practice this means these results are asymptotic (the more precise the larger the sample size).

We define

- $Y = ID_e$  and  $X = MT$
- $\mu_X = \mathbb{E}[X]$  the mean of the random variable  $X$
- $\sigma_X = \sqrt{\text{Var}(X)}$  the standard deviation of the random variable  $X$

The Pearson correlation coefficient is defined as a normalized covariance

$$r(X, Y) = \frac{\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]}{\sigma_X \sigma_Y}. \quad (24)$$

We write  $\mathbb{E}[X|Y] = f(Y)$ ,  $\text{Var}[X|Y] = g(Y)$ . We have that

$$\mathbb{E}[XY] = \mathbb{E}[Y\mathbb{E}[X|Y]] = \mathbb{E}[Yf(Y)] \quad (25)$$

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[f(Y)] \quad (26)$$

$$\text{Var}(X) = \mathbb{E}[\text{Var}(X|Y)] + \text{Var}(\mathbb{E}[X|Y]) \quad (27)$$

$$= \mathbb{E}[g(Y)] + \text{Var}(f(Y)) \quad (28)$$

The general practice of computing block averages described subsection 2.3 essentially considers  $\overline{X} = \mathbb{E}[X|Y]$  instead of  $X$ :

$$r(\overline{X}, Y) = \frac{\mathbb{E}[\overline{X}Y] - \mathbb{E}[\overline{X}]\mathbb{E}[Y]}{\sigma_{\overline{X}} \sigma_Y}. \quad (29)$$

One notices from Equation 25 and Equation 26 that interestingly, the covariance between  $X$  and  $Y$  equals the covariance between  $\overline{X}$  and  $Y$ . We then have

$$\sigma_X^2 = \text{Var}(\mathbb{E}[X|Y]) = \text{Var}(f(Y)). \quad (30)$$

<sup>14</sup>Copulas can exhibit dependencies in the lower tails (low values) or upper tails (high values) of distributions. Given that pointing data in our dataset displays high variance at higher ID levels, copulas with upper tail dependence, such as the Gumbel copula, tend to perform poorly. By rotating the copula, we switch the dependence from the upper tail to the lower tail, providing additional copula candidates for consideration.

When compared with Equation 27, one thus sees that the only difference between  $r(X, Y)$  and  $r(\bar{X}, Y)$  is that  $\mathbb{E}[\text{Var}(X|Y)]$  is not present in the denominator. This shows  $r(\bar{X}, Y)$  is always larger than  $r(X, Y)$ , and explains the practice of block averages.

*Linear conditional expectation and standard deviation leads to perfect correlation.* It is known that pointing data reaches very high values of  $r(\bar{X}, Y)$ , often above .9, and sometimes even above .99. Here, we show that a linear model of conditional expectation and standard deviation will reach  $r(\bar{X}, Y) = 1$ . Linear conditional expectation and standard deviation read

$$\mathbb{E}[X|Y] = a + bY = f(Y) \quad (31)$$

$$\text{Var}[X|Y] = (c + dY)^2 = g(Y). \quad (32)$$

This implies

$$r(X, Y) = \frac{\mathbb{E}[Y(a + bY)] - \mathbb{E}[a + bY]\mathbb{E}[Y]}{\sqrt{(\mathbb{E}[(c + dY)^2] + \text{Var}(a + bY))\sigma_Y}} \quad (33)$$

$$r(\bar{X}, Y) = \frac{\mathbb{E}[Y(a + bY)] - \mathbb{E}[a + bY]\mathbb{E}[Y]}{\sqrt{\text{Var}(a + bY)\sigma_Y}}. \quad (34)$$

The covariance part simplifies as

$$\mathbb{E}[Y(a + bY)] - \mathbb{E}[a + bY]\mathbb{E}[Y] = a\mu_Y + b\mathbb{E}[Y^2] - a\mu_Y + b\mathbb{E}[Y]^2 \quad (35)$$

$$= b\sigma_Y^2. \quad (36)$$

Then

$$\text{Var}(a + bY) = b^2\text{Var}(Y) = b^2\sigma_Y^2, \quad (37)$$

which leads to  $r(\bar{X}, Y) = 1$ . Note that  $g(Y)$  does not play a role in this, as the variance of  $X$  for a given  $Y$  is nullified when considering  $\bar{X}$ .

The extra term in  $r(X, Y)$  can also be evaluated as

$$\mathbb{E}[(c + dY)^2] = \mathbb{E}[c^2 + 2cdY + d^2Y^2] \quad (38)$$

$$= c^2 + 2cd\mu_Y + d^2(\sigma_Y^2 + \mu_Y^2) \quad (39)$$

and some further algebra shows

$$\sqrt{(\mathbb{E}[(c + dY)^2] + \text{Var}(a + bY))} = \sqrt{c^2 + \mu_Y(2cd + d^2) + (d^2 + b^2)\text{Var}(Y)}, \quad (40)$$

which offers no further simplification for  $r(X, Y)$ . This implies that the dependence between  $\text{ID}_e$  and MT can not be specified independently of the EMG parameters, even though all possible values of  $r(X, Y)$  are theoretically attainable:

- If  $b^2\sigma_Y^2 \gg c^2 + \mu_Y(2cd + d^2)$ , then  $\sigma_X \rightarrow \sqrt{b^2 + d^2}\sigma_Y$  and  $r(X, Y) \rightarrow \frac{b}{\sqrt{b^2 + d^2}}$ . If, in addition,  $b \gg d$ , this reduces to  $r(X, Y) = \text{sign}(b)$  i.e.,  $\pm 1$ . In practice, this can occur if  $b$  is very large, which is unusual.
- If  $c^2 + \mu_Y(2cd + d^2) \gg b^2\sigma_Y^2$ , then  $\sigma_Y \ll \sigma_X$  and  $r(X, Y) \rightarrow 0$ .

## C THE WHO MODEL

The Who model by Guiard and Rioul [21] is an axiomatic model that conforms to the following 5 axioms:

- a1 There is a minimum movement time  $MT_0$  achievable
- a2 If we define relative precision as  $y = \sigma/D$  with  $\sigma$  the standard deviation of endpoints, then there is a minimum relative precision  $y_0$  achievable.
- a3 The function  $f$  that links MT to  $y$ ,  $y = f(MT)$  is decreasing and convex
- a4 For a given participant effort  $k$ , there is an operation such that  $MT \odot y = k$
- a5 Effort invested is never total

The WWho model

$$(y - y_0)^{1-\alpha} (MT - MT_0)^\alpha = k \quad (41)$$

satisfies a(1) – a(4), and is fit on the convex hull of the scatterplot in the  $(MT, y)$  space to satisfy a(5).

The WWho model is *not* a description of distribution of movement times, instead it characterizes the “best” samples (axioms 2 and 5), which is operationalized by fitting not on all data points but only on the convex hull of the dataset. Guiard and Rioul never described what the distribution of movement times could look like within their WWho model.