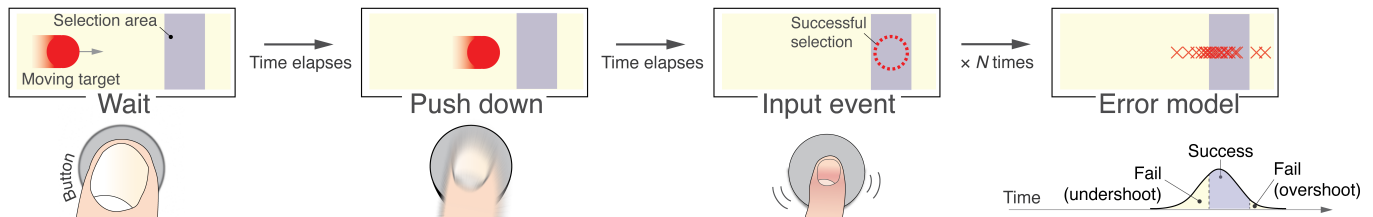


# Modelling Error Rates in Temporal Pointing

Byungjoo Lee  
Aalto University, Finland  
bjlee1985@gmail.com

Antti Oulasvirta  
Aalto University, Finland  
antti.oulasvirta@aalto.fi



**Figure 1.** This paper presents a novel predictive model for error rates in temporal pointing. In temporal pointing, the task is to select a target, which appears for selection within a limited time window. This task is common in interactions that require temporal precision, synchrony, or rhythm. Our model predicts the mean and spread of the response distribution as a function of the distance (amplitude or duration-to-target) and width of this time window (precision or duration-of-target). Error rates are computed from this distribution.

## ABSTRACT

We present a novel model to predict error rates in temporal pointing. With temporal pointing, a target is about to appear within a limited time window for selection. Unlike in spatial pointing, there is no movement to control in the temporal domain; the user can only determine when to launch the response. Although this task is common in interactions requiring temporal precision, rhythm, or synchrony, no previous HCI model predicts error rates as a function of task properties. Our model assumes that users have an implicit point of aim but their ability to elicit the input event at that time is hampered by variability in three processes: 1) an internal time-keeping process, 2) a response-execution stage, and 3) input processing in the computer. We derive a mathematical model with two parameters from these assumptions. High fit is shown for user performance with two task types, including a rapidly paced game. The model can explain previous findings showing that touchscreens are much worse in temporal pointing than physical input devices. It also has novel implications for design that extend beyond the conventional wisdom of minimising latency.

## Author Keywords

Temporal pointing; Human performance models; Fitts' law

## ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

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).

CHI'16, May 07–12, 2016, San Jose, CA, USA  
Copyright © 2016 ACM. ISBN 978-1-4503-3362-7/16/05\$15.00  
DOI: <http://dx.doi.org/10.1145/2858036.2858143>

## INTRODUCTION

This paper presents novel data, analysis, and a mathematical model for *user performance in temporal pointing tasks* in human–computer interaction (HCI). In temporal pointing, the user's goal is to make a discrete selection of a target about to appear for a bounded time window. Figure 1 shows an example. Imagine trying to shoot a target that is crossing the screen at some velocity. Your finger resting on a button, you intend to press it down when the enemy is under the selection area. If the input event occurs too early, or too late, you miss the target. Temporal pointing is not limited to gaming; it occurs in applications requiring temporal precision, synchrony, or rhythm. In this paper, we focus particularly on temporal pointing that involves practically no spatial pointing.

When we started working on this topic, we were surprised to learn that temporal pointing is not addressed by predictive models in HCI. Fitts' law and its variants [8, 9, 33, 40, 10, 24] are limited to the spatial domain. From Fitts' law's perspective, our scenario is trivial: as the finger is already on (or above) the button, travel distance is minimal and movement time constant (i.e.,  $MT \approx a$ ). As we argue later, the case is covered neither by models of reaction time nor by sensorimotor synchronisation studies in psychology.

Our objective was to account for several effects that are of obvious interest to HCI but that, on the other hand, complicate the modelling task. Firstly, to be relevant to design, the model should address the properties of the temporal target window, including its distance (amplitude or duration-to-target) and width (precision or duration-of-target) in time. It is easier to shoot when the enemy is moving slowly or the set of crosshairs is large. We use temporal distance  $D$  and temporal width  $W$  to refer to these properties in this paper. Secondly, visual cues affect how predictable the target is. In the worst case, no cue is available other than that the target has appeared in the past. In a better case, the target can be perceived moving on the display. Thirdly, the input device ob-

viously has an effect on performance, but it is not known how. Indeed, a few studies in HCI looking at temporal performance in games have found large differences among input devices. In particular, they have found that gaming performance can be up to 150% worse in playing with a touchscreen device rather than a physical controller [5, 42, 43]. We argue later in this paper that these differences are not trivially attributable to latencies or tactile feedback. Fourthly, users can also adopt different response strategies. In the scenario above, would you try to anticipate and start pressing down slightly *before* the enemy reaches the crosshairs, or would you rather wait until it is there and press down then?

This paper contributes by 1) defining temporal pointing in terms of the concept of temporal target and 2) deriving a novel mathematical model to predict error rates. The model assumes that users have an *implicit temporal point of aim* within the target window, the point in time at which they *intend* the input event to be registered. However, the model further assumes that their ability to elicit the input event at that time is hampered by variability accumulating over three processes: 1) an internal time-keeping process, 2) a response-execution stage, and 3) input processing in the computer. Two empirical parameters capture the changes in the point of aim and variability. The model yields a task-specific response distribution, a Gaussian in time, from which error rate  $E$  can be calculated. The mathematical description captures a transition between two response strategies found in previous studies: anticipation and reaction [7, 31]. Users tend to rely more on anticipation (their internal clock) than on reaction when the task becomes more difficult. Although simple, the model adequately addresses two quite different cases:

- Approaching target: The target is visually moving toward a selection region, as in the shooter game.
- Blinking target: The target is not moving but is appearing with some predictable interval, such as that in playing a musical instrument.

Our work offers some insights into design that we believe are novel. The concept of temporal target is easy to understand and can be analysed, even for dynamic tasks. When it is combined with the model, one can predict how changes in the interactive task affect user performance. For example, one can use it in designing game levels for the degree of difficulty. We also discuss design factors in input methods affecting the implicit target point and variability in responses. These go beyond the standard wisdom that minimising latency is important.

In the rest of the paper, we firstly argue that the task is not covered by previous models used in HCI. We then explain the theory and mathematics behind the model. Finally, we report results from two studies: one with a blinking target and the other with a dynamically changing visual target.

## PRIOR ART IN SPATIAL AND TEMPORAL POINTING

Figure 2 delineates existing models by way of two dimensions: 1) temporal and 2) spatial task requirements. Temporal pointing, as defined here, can be thought of as the temporal counterpart of the familiar Fitts’ task. As does the Fitts’ task,

		Spatial Task Requirements			
		Goal	Constraints		
		‘As short as possible’	Distance	Distance & width	
Temporal Task Requirements	Goal	‘As quickly as possible’	Ratcliff’s model [26, 37]	Gan & Hoffman’s model [9]	Fitts’ law [8] (output: time)
	Constraints	Distance (duration-to-target)	Wing & Kristofferson model [38, 39]	Schmidt’s law [33]	Wobbrock’s model [40]
		Distance & Width (duration-of-target)	<b>Our model</b> (output: error rate)		

**Figure 2. The task studied in this paper is defined as the temporal counterpart to Fitts’ law: The task is to hit a target appearing after some duration (distance) for a limited time (width). However, movement in space is minimal, as the finger is resting on/above a button. Fitts’ law predicts movement time with constant error rate in space (4%). By contrast, our model predicts error rates in time with constant task completion time (duration-to-target). Previous models do not address this task.**

it assumes limited precision (width  $W$ ) and amplitude (distance  $D$ ) in temporal selection. Unlike in spatial pointing, however, there is virtually no movement in space. The user can only determine *when* to launch the response. The figure also shows that models of response times in psychology do not cover the case with  $D$  and  $W$ . We now expand on this argument.

**Note:** In the following discussion, to indicate when a variable refers to the spatial domain, we use a subscript  $s$ . For variables in the temporal domain, we use a subscript  $t$ .

## Spatial Pointing

Although the familiar models of spatial pointing in HCI look at temporal aspects of performance, such as movement time, they do not predict user performance in the task of hitting a temporally constrained target.

Fitts’ law [8] predicts minimum movement time  $MT$  with two spatial requirements: movement distance  $D_s$  and target width  $W_s$ . In other words, the user is asked to move as quickly as possible to hit the target. Fitts introduced the concept of index of difficulty:  $MT = a + bID_s = a + b \log_2(2D_s/W_s)$ . Wobbrock et al. [40] analysed Fitts’ law to derive a model of error rates with three task requirements mixed in the temporal and spatial domain: target size  $D_s$ , target width  $W_s$ , and temporal target distance  $D_t$ . As this paper does, they predicted error rates, because there was an explicit requirement to maintain a temporal distance to the target  $D_t$ . Besides, when the spatial requirement  $W_s$  is relaxed from Fitts’ task, targeting becomes more open-loop-like. Gan and Hoffman [9] proposed a model to predict  $MT$  when the goal is to hit a large target at distance  $D_s$ . Hoffman’s law originally predicted  $MT = a + b\sqrt{D_s}$ . With additional assumptions, it can be expressed as the constant ratio between  $D_s$  and the standard deviation (SD) of end points [12].

Schmidt’s law [33] adds one temporal requirement to Gan and Hoffman’s model, time to target  $D_t$ . This changes the task from time-minimisation to spread-minimisation at a known

distance in time. The law predicts that SD (in the spatial domain) is linearly related to movement velocity. If time to the target  $D_t$  is controlled, SD increases linearly with spatial distance  $D_s$ . In this case, the model is the same as that of Hoffman et al. [9, 12]. Guiard et al. [10] recently presented the most comprehensive model of spatial aiming. The *WHO* model explains different aspects of spatial movement by reference to six axioms of human movement. That model subsumes Fitts' law and Schmidt's law by reference to a trade-off phenomenon among the axioms. The model is not depicted in Figure 2 since our analysis is based only on task instructions rather than actual user response. In addition, models inspired by control theory [6, 15, 20, 19, 25] have looked at speed-accuracy trade-offs in spatial pointing tasks.

### Reaction

The Ratcliff model [26, 37, 30] addresses *simple reaction*: a response to the onset of an abrupt stimulus or event. Reaction-time distributions are often Gaussian or skewed Gaussians. The model assumes that reaction time  $RT$  has two sources of variation: decision time  $T_d$  and non-decision time  $T_{er}$ , which is further broken down into the two sub-components  $x$  and  $y$ :  $RT = T_d + T_{er} = x + T_d + y$ . The non-decision term has two components. The first is the perceptual encoding of the stimulus that lasts for some duration denoted with  $x$ . After perception of the stimulus, a stochastic decision process starts. During this period, evidence is accumulated (diffused) in the brain that the stimulus should be responded to. After sufficient evidence for passing threshold  $a$ , the decision process stops and the motor response process begins, with duration  $y$ . The duration of the decision process  $T_d$  is user- and task-dependent and varies across trials. The drift rate (or the accumulation of evidence) is assumed to be normally distributed with mean  $v$  and SD  $\eta$ . The two non-decision components  $x$  and  $y$  are summed to  $T_{er}$  and treated together in the model because they are difficult to disentangle in real data.

*Choice reaction* refers to the task of choosing from among  $N$  alternatives one that matches a given target. In a similarity to spatial pointing, a speed-accuracy trade-off has been noted and explained by reference to limited information capacity [30]. Faster reaction results in a higher error rate. A user can decrease errors by thinking more. When this speed-accuracy trade-off function is assumed to be linear, we get the well-known Hick-Hyman law [14].

### Synchronisation

The Wing and Kristofferson (WK) model [39, 38] predicts performance in the *synchronisation* task, in which one has to synchronise, for example, tapping or clapping with some repeating stimulus, such as a beep sound. The beep is regarded as an impulse signal, and the user has no explicit requirement of temporal preciseness  $W_t$ . In other words, only  $D_t$  (inter-onset interval, IOI) is controlled. The model divides the cognitive process in synchronisation tasks into two parts: an internal time-keeping process is followed by a motor response with time delay. Both time-keeping precision and motor delay are assumed to be stochastic and separately represented by additive random noise with a mean and variance. Noise

components can be extracted from a tapping experiment by examining autocorrelation in consecutive asynchronies [39]. The variance of the time-keeping process linearly increases with IOI [27, 29]. There is a similarity here to the model of ballistic movements in the spatial domain, in which the SD of end location linearly increases with target amplitude  $D_s$ . We will use this linearity as a primary assumption in our model. By contrast, variance in motor responses remained relatively constant and lower than time-keeping variance. This implies that variance in motor responses dominates in fast tapping (IOI sub-250 ms).

As with reaction times, the distribution of asynchronies is approximated well with a Gaussian [18, 21]. Research has shown that stimulus modality and differences between individuals affect performance. Auditory stimuli (such as beep sounds) are better for synchronisation, especially with higher tempos (100–125 ms IOI limit), than visual stimuli [28]. Furthermore, large individual-to-individual differences have been observed, for example, between musicians and non-musicians [28, 27].

While the WK model explains two important variances, it does not explain the offset between stimulus and tapping, which also changes with IOI [27, 29]. A Bayesian model of human timing [2, 41, 1] provides an explanation to the varying aim point in synchronisation. It explains this as statistical self-correction occurring between taps when a person is trying to maximise performance for a given loss function.

### Interception

*Intercepting* a moving target demands considering *both* spatial and temporal amplitude to the target, but, depending on the target's speed and size, precision might also be required [4]. To our knowledge, there is no model of target interception that would consider the case of temporal pointing wherein the spatial task requirement is negligible. Anticipation-coincidence studies [17, 3] and the Newell group's work [22, 11, 23, 16, 13] have offered a mixed view, wherein temporal and spatial requirements co-exist. Two distinct categories of existing models can be recognised: 1) models assuming preprogrammed responses and 2) on-line control models. There has been some criticism of these models [35]. One critical issue is that several strategies are possible. For example, when one is capturing a target, four distinct strategies have been identified (pursuit, head-on, receding, and perpendicular) [35]. The strategy is affected by how the target moves, its identity, whether it is to be captured or hit, and how it is approached. One replicated finding is that when moving targets require more temporal precision (smaller  $W_t$ ), people increase the movement velocity [36, 4, 34]. However, faster movement also decreases spatial precision, as Schmidt's law predicts. This has been explained by reference to a process that maintains optimal strategy between spatial and temporal precision [32, 41]. In summary, although interception is a generalisation of both spatial and temporal pointing, at present no model is able to predict performance in our case – i.e., when spatial task requirements are virtually non-existent.

## Summary

Our model builds on several insights from prior work. Firstly, we assume – and later show – that response distributions in time are Gaussian. Secondly, we assume 1) a time-keeping process and 2) a response process. Both processes are unreliable, and variability is modulated by task and experience. Unlike previous work, we assume an implicit point of aim for a given temporal selection window, which can be understood as an optimal strategy of the user, chosen in terms of temporal speed–accuracy trade-off.

## OVERVIEW AND DERIVATION OF THE MODEL

Our model predicts the distribution of users’ responses in time, from which we compute error rate  $E$  [0, 1]. It is computed by subtracting areas occurring inside the target from the total area in the distribution (see Figure 1). The model takes two inputs defining *the temporal target*: width (or the duration in time for which it remains selectable)  $W_t$  and distance  $D_t$ , or time until target onset. The response distribution is represented as a Gaussian with a mean of  $c_\mu \times W_t$  and standard deviation of  $c_\sigma \times D_t$ . Two empirical parameters,  $c_\mu$  and  $c_\sigma$ , determine the mean and spread of the distribution. They have the following explanations:

1.  $c_\mu$  is the implicit temporal aim point of the user.
2.  $c_\sigma$  captures aggregated noise and uncertainty across the three stages to pointing.

We now discuss theoretical assumptions of the model. We then derive the model and, finally, discuss its use.

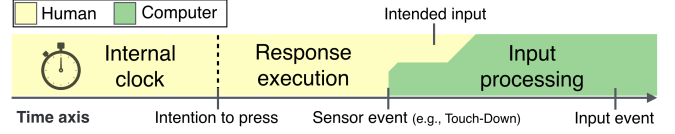
### Theoretical Assumptions

The model assumes three stages preceding the final input event in the application. Figure 3 gives an overview.

**1. Internal clock:** As in the WK model [39, 38], we assume that users rely on an internal clock to decide when to launch the response. The time-keeping is similar to a ballistic motor movement: it is based on an internal estimate and uncorrectable once executed. The estimate is affected by the temporal target – the larger  $D_t$ , the worse the estimate (if you ask someone to count out 10 seconds with eyes closed, the error will be much larger than in counting out, say, three seconds). Moreover, the estimate is noisy. If the target is perceived only intermittently, or its velocity is hard to estimate, the estimate is worse.

**2. Response execution:** The response stage starts with launching the movement of the finger. Models of spatial targeting, such as Fitts’ law, explain performance here only if the movement trajectory involved is large. In our case, the trajectory involved is minimal and the target (button) very large. Variability in this stage is caused by several sources, including users’ estimate of when the input is registered, variability in finger travel distance (e.g., if the finger hovers over a virtual button), and variability in muscle activation during motion.

**3. Input processing:** Sensor data are registered and processed further to determine the input event. The duration of this stage depends on the input device and input method. For



**Figure 3.** The three stages preceding the final input event: 1) internal clock, 2) response execution, and 3) input processing in the computer.

example, a touch event on a touchpad can be registered as input at Touch-Down or at Touch-Release. For simplicity, we do not manipulate input-processing duration in our study.

*Shift in response strategy:* We also assume (and show) that response strategy changes with temporal target. When  $W_t$  is large, users do not need to anticipate the target; they simply respond when it appears (reactive strategy). If  $W_t$  is small, they must anticipate (anticipatory strategy). The model captures this shift as we show below.

### Mathematical Derivation

The response distribution  $R$  as a function of time  $t$  centred at  $\mu$  from target onset is assumed to be Gaussian:

$$R(t|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}, \text{ where} \quad (1)$$

$$\sigma = c_\sigma \times D_t \text{ and } \mu = c_\mu \times W_t \quad (2)$$

The linear relationship between  $\sigma$  and  $D_t$  was inherited from sensorimotor synchronisation studies. More intuitively,  $c_\sigma$  describes the ballistic time estimation of the internal clock as similar to that of spatial ballistic movements [12, 9]. The mean  $\mu$  of the distribution represents the implicit aim point of the user for a given temporal target. It assumes that users will consistently aim at a fixed ratio  $c_\mu$  for the temporal width  $W_t$  of the target. Parameter  $c_\mu$  is a positive constant, and the largest permissible value is 1.0.

Plugging  $\sigma$  in Equation 2 into Equation 1, we obtain a function for response distribution:

$$R(t|W_t, D_t) = \frac{1}{c_\sigma D_t \sqrt{2\pi}} e^{-\frac{t^2}{2(c_\sigma \times D_t)^2}} \quad (3)$$

We can now calculate error rate  $E$  ( $\in [0, 1]$ ) by subtracting the hit area from the total area in the distribution:

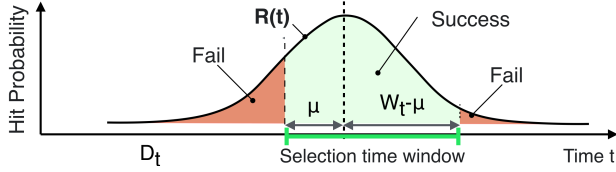
$$\begin{aligned} E(D_t, W_t) &= 1 - \int_0^{W_t - \mu} R(t) dt - \int_0^\mu R(t) dt \\ &= 1 - \frac{1}{2} \left[ \operatorname{erf}\left(\frac{W_t - \mu}{\sigma\sqrt{2}}\right) + \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) \right] \\ &= 1 - \frac{1}{2} \left[ \operatorname{erf}\left(\frac{(1 - c_\mu)}{c_\sigma\sqrt{2}} \cdot \frac{W_t}{D_t}\right) + \operatorname{erf}\left(\frac{c_\mu}{c_\sigma\sqrt{2}} \cdot \frac{W_t}{D_t}\right) \right] \end{aligned} \quad (4)$$

where  $\operatorname{erf}(x)$  is the (known) error function encountered in integrating the normal distribution:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (5)$$

Figure 4 illustrates this. Note: Error function  $\operatorname{erf}(x)$  should not be confused with error rate function  $E$  given above.





**Figure 4.** Error rate is computed by integrating a Gaussian within the range defined by failed trials.

### The Final Model

The model predicts errors as a function of how difficult the target is. The further away it is (the larger the  $D_t$  value) or the smaller the target width ( $W_t$ ), the harder the task. Task difficulty is measured by  $ID_t$  (unit: bits), in a similar way to that of Fitts' law (see Figure 5):

$$ID_t = \log_2(D_t/W_t) \quad (6)$$

We compute  $E$  from Equation 4 by using index of difficulty  $ID_t$ . The model collapses to a single dimensionless quantity:

$$E(ID_t) = 1 - \frac{1}{2} \left[ \operatorname{erf}\left(\frac{1 - c_\mu}{c_\sigma 2^{(ID_t + 0.5)}}\right) + \operatorname{erf}\left(\frac{c_\mu}{c_\sigma 2^{(ID_t + 0.5)}}\right) \right] \quad (7)$$

### Effects of Varying Noise and Anticipation on Error Rate

By varying  $c_\mu$  and  $c_\sigma$ , we can observe their effects on error rate. In Figure 6, we vary one of these parameters at a time while fixing the other parameter. The  $x$ -axis represents index of difficulty ( $ID_t$ ) as defined above. Note that  $ID_t=0$  can result in a non-zero error rate, because the process is stochastic in consequence of the ballistic property of the internal clock. If the variability in aiming is much greater than  $W_t$ , there will always be some amount of undershooting or overshooting.

We make the following observations:

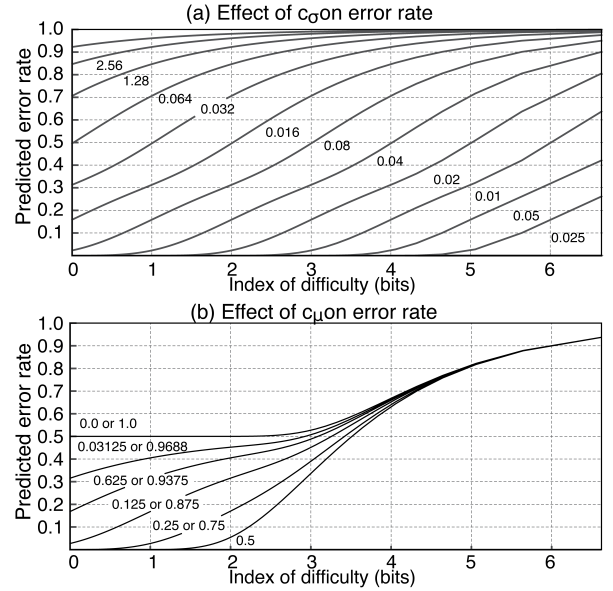
1. After around  $3 \leq ID_t \leq 4$ , the error rate converges to one increasing trend. While the detail range will change with the  $c_\sigma$  we use for the plot, it could basically represent a cross-over point where the anticipative strategy is not effective anymore.
2. Ideally, the best aim point is at the centre of the temporal target window  $c_\mu = 0.5$ .
3. Increasing  $c_\sigma$  shows virtually no effect after  $c_\sigma = 5$ .

### Limitations

The model subscribes to the assumption that the relationship between  $\mu$  and  $W_t$  is linear. This is practical, because the error rate can be expressed as a function of  $D_t/W_t$ , or index of difficulty. However, this linearity would not hold for

ID (bits)	Distance/Width	Task
2	4	Time  Temporal target
1	2	Temporal target
2	4	Temporal target

**Figure 5.** Temporal target and task difficulty: The further away the target is (the larger its  $D_t$ ) or the smaller its width ( $W_t$ ), the harder the pointing task. Difficulty is measured by *index of difficulty*  $ID_t$  (unit: bits). We show two hard tasks (2 bits) and an easy task (1 bit).



**Figure 6.** Effect of model parameters  $c_\sigma$  and  $c_\mu$  on error rate with varying index of difficulty, where (a)  $c_\mu$  was fixed at 0.16 and (b)  $c_\sigma$  was fixed at 0.065.

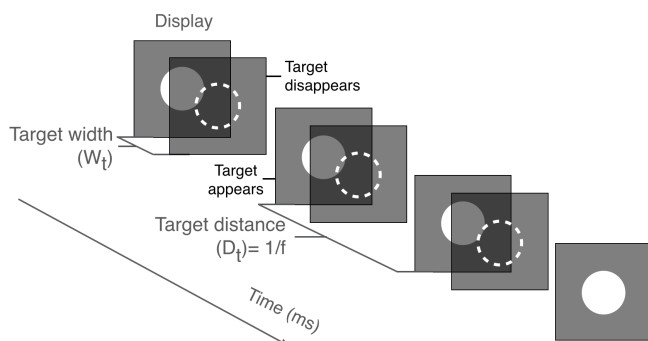
$W_t$  greater than human reaction time ( $\sim 300$  ms). If so, a user could choose to be purely reactive rather than apply anticipation. However, the reaction task is not covered in our model. There are similar limitations with overly short  $D_t$  due to the rate limits in internal clocks [28] as discussed in the ‘Synchronisation’ section. Also, the scope with the model is limited to tasks that involve very little spatial movement. If the user must aim at a button, for example, another stage is introduced. This is not covered in the model.

### Obtaining Parameters Empirically

In a similar way to Fitts' law models, the two parameters can be determined in a simple empirical study wherein  $W_t$  and  $D_t$  are varied. We generally recommend limiting  $ID_t$  to between 1 and 6. In the experiments described below, we used IDs of 9 and 12 to test the model, and it was found to be sufficient. For a visually cued task, we recommend  $D_t \geq 700$  ms. In an auditorily cued task,  $D_t$  can be less, even as brief as 250 ms [28]. Width  $W_t$  should be between 0 ms and 200 ms. If temporal targets are repeating,  $W_t$  should be shorter than  $D_t$ , to avoid overlap between targets. In a blinking study, it is straightforward to set  $D_t$  and  $W_t$ . In contrast, in studies using visually approaching targets, spatial regions must be converted into temporal ones. This is done by considering the associated velocities and sizes of regions. An example is provided with Experiment 2 (Flappy Bird).

### EXP. 1: BLINKING TASK WITH FIVE INPUT METHODS

Experiment 1 involved the blinking task, the simplest form of temporal pointing. Users had solely the internal clock to rely on. In our task, a filled circle is blinking with a fixed onset interval, as shown in Figure 7. The circle is visible for  $W_t$ . Participants must ‘eliminate’ it by producing an input event just when the target is visible. Target distance  $D_t$  is the inverse of blinking frequency:  $1/f$ .



**Figure 7.** Experiment 1 uses a blinking task. The user must create an input event by pressing a button during the duration of a circle’s visibility.

We also compare five input methods. The hypothesis is that they differ in implicit aim point  $c_\mu$  and in variability  $c_\sigma$ . We selected the methods such that the results have a bearing in relation to the previously reported difference between physical and touchscreen buttons in gaming performance:

- **Touch-Down:** Evoked when the finger is first registered as contacting the sensor surface.
- **Touch-Maximum<sup>1</sup>:** Evoked when the contact area of the finger is maximised.
- **Touch-Release:** Evoked when the finger has lifted off from the sensor surface.
- **Key-Press:** Evoked when the tactile switch first contacts the electrode.
- **Key-Release:** Evoked when the electrodes first become separated (when contact between them is lost).

## Method

**Participants:** Twenty participants (8 males, 11 females, 1 other) were recruited from a local university. The mean age was 28 (SD=8.3). Five wore eyeglasses. Ten had experience with touchscreen games or musical instruments. We did not control the number of musicians in the sample; six of our participants reported having a background as a musician.

**Experimental design:** The experiment followed a within-subject design with three independent variables: *target width*  $W_t$ , *target distance*  $D_t$ , and *input method*. The levels were the following:

- $W_t$ : 50, 116.7, and 200 ms
- $D_t$ : 600, 900, and 1200 ms
- $ID_t$ : 1.56, 2.17, 2.36, 2.59, 2.95, 3.37, 3.59, 4.17, and 4.59 bits
- Input methods: Touch-Down, Touch-Maximum, Touch-Release, Key-Press, and Key-Release

Participants completed 100 selections per  $ID_t$  level. The order of  $ID_t$  was randomised within an input method. The order of input methods was also randomised.

**Task:** The participants were instructed to select a blinking circle without skipping any of them. This is to keep target dis-

<sup>1</sup>According to the law of conservation of energy, this indicates the timing of maximum kinetic energy transferred to the touch surface.

tance within the range permissible in the model. They were told to select the target as quickly and as accurately as possible *from its onset*. They were informed that quicker selection may require anticipating when the target appears since the target duration would be shorter than typical human reaction time.

**Materials:** The target was a white-filled circle presented on a black background (diameter: 23 mm). After its appearance, the opacity of the circle was linearly decreased. It disappeared completely at  $W_t$ . When selection was successful, the target disappeared immediately. Upon a miss, no feedback was given.

**Procedure:** Participants sat on a regular office chair and used a laptop placed on a desk. After adjustment of the chair, a pre-study questionnaire was administered, asking about previous experience in gaming and music. Next, the subjects were introduced to the task and given several examples of successful and failed selections. Training included making 10 selections per  $ID_t$  condition. In the experiment, a short break was provided after each  $ID_t$  condition. An experiment took about one hour to finish.

**Apparatus:** The software was implemented as a Java application running on a laptop (MacBook Pro late-2012, 2.2 GHz Intel Core i7 with Intel Iris Pro 1536 MB graphics card). Its touchpad and keyboard were used as the input devices. The size of the display was 35.8 cm × 24.7 cm. In the touch input conditions, participants used the touchpad with the preferred finger (index or middle). For the two physical-key conditions, they used the ENTER key. The refresh rate of the application was 60 fps.

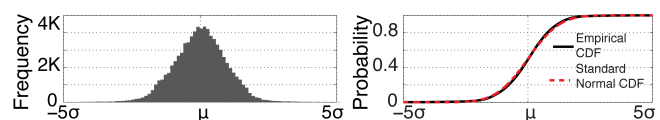
## Results

Before the modelling results, we show that the response distribution is Gaussian. For statistical testing, we use repeated-measure ANOVA with an alpha level of 0.05. Greenhouse–Geisser correction was used for the violation of sphericity.

### Normality Test for the Response Distribution

The assumption about normality holds for the large majority of the data. A normalised histogram of 81,000 hits and a CDF plot of both normal and empirical data are shown in Figure 8. For testing normality, we applied one-sample Kolmogorov–Smirnov for each block of a hundred repeats of  $ID_t$ +method conditions. An alpha of 0.05 was used.

In the 810 distributions obtained (one per condition per user), 79.88% were statistically a normal distribution. Only two users had a large quantity of non-normal trials (33% and 44%). We excluded these two from the analysis. The rest had fewer non-normal trials (M=6.78% SD=7.75%).



**Figure 8.** Response distributions in Experiment 1 (blinking task). On the right, a CDF comparison between the normal and empirical spread of hits. On the left, a histogram of normalised hits.

### Testing Assumptions about $D_t$ and $W_t$

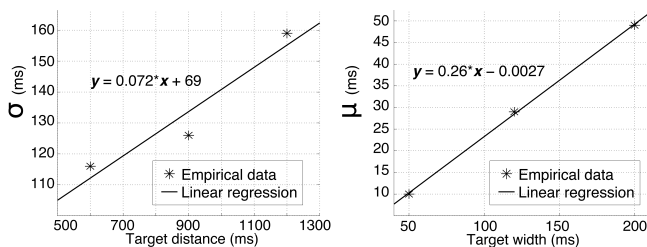
Although the number used for both  $D_t$  and  $W_t$  in the experiment was limited to 3, the assumptions put forward in Equation 2 are supported by the data. Firstly, the effect of  $W_t$  on the centre of the response distribution ( $\mu$ ) was significant:  $F(1.23, 20.98) = 31.84$ ,  $p < 0.001$ . A linear polynomial contrast statistic further supports the model's assumption that this relationship is linear:  $F(1, 17) = 35.33$ ,  $p < 0.001$ ,  $R^2 = 0.99$ . The slope is presented in Figure 9. From the slope, we can estimate  $c_\mu$  to be 0.26 for the trials overall. In other words, users were targeting the midpoint between the onset and the target centre.

Secondly, as expected, target distance  $D_t$  had a significant effect on the spread of the response distribution ( $\sigma$ ):  $F(1.31, 22.30) = 14.28$ ,  $p < 0.001$ . The fit, at  $R^2 = 0.91$ , was lower than for  $W_t$ . Pairwise comparisons showed a significant difference between 900 ms and 1200 ms ( $p < 0.001$ ). However, the difference was not significant between 600 ms and 900 ms ( $p = 0.18$ ). Consequently, error rate was affected by  $D_t$ , too:  $F(1.43, 24.30) = 10.27$ ,  $p = 0.002$ . The condition  $D_t = 600$  ms showed a higher error rate ( $M = 0.59$ ,  $SD = 0.24$ ) than the 900 ms ( $M = 0.51$ ,  $SD = 0.24$ ) and 1200 ms ( $M = 0.57$ ,  $SD = 0.23$ ) ones. Pairwise comparison showed a significant difference in error rates between 600 ms and 900 ms ( $p = 0.001$ ) and between 900 ms and 1200 ms ( $p < 0.001$ ).

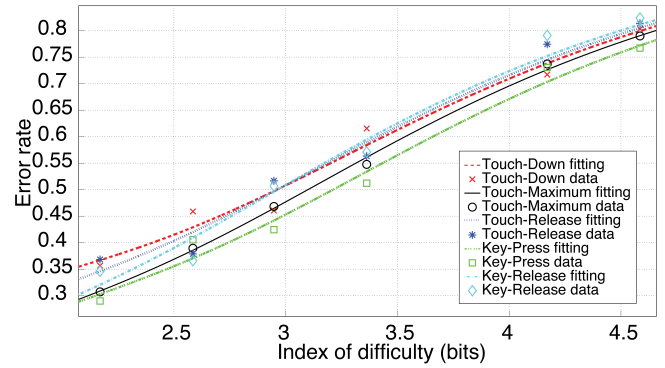
### Effect of Input Method

As predicted, input method showed a significant effect on error rate:  $F(2.38, 40.51) = 5.74$ ,  $p = 0.004$ . Pairwise comparisons show that Touch-Down ( $M = 57.4\%$ ,  $SD = 2.9\%$ ) had a 4.5% higher error rate ( $p = 0.005$ ) than Key-Press ( $M = 52.9\%$ ,  $SD = 3.1\%$ ). Error rates for Touch-Maximum ( $M = 55.6\%$ ,  $SD = 3.1\%$ ) were 4.3% lower ( $p = 0.01$ ) than those for Touch-Release ( $M = 59.9\%$ ,  $SD = 3.2\%$ ), and 4.8% lower than Key-Release's ( $M = 60.4\%$ ,  $SD = 3.1\%$ ) with a borderline effect ( $p = 0.053$ ). The Key-Press method showed a 7% and 7.5% lower error rate than Touch-Release ( $p < 0.001$ ) and Key-Release ( $p = 0.003$ ), respectively.

Input method had a significant effect on distribution centre ( $\mu$ ):  $F(2.43, 41.38) = 10.39$ ,  $p < 0.001$ . As expected, Touch-Down showed an earlier response ( $M = 12.4$  ms,  $SD = 38.93$  ms) than the others ( $p < 0.001$ ). Touch-Maximum ( $M = 27.7$  ms,  $SD = 42.94$  ms) and Key-Press ( $M = 23.7$  ms,  $SD = 37.23$  ms) were similar in this respect. Touch-Release ( $M = 36$  ms,  $SD = 47.63$  ms) and Key-Release ( $M = 43.5$  ms,  $SD = 45.53$  ms)



**Figure 9.** Target width  $W_t$  versus  $\mu$  (right) and target distance  $D_t$  versus  $\sigma$  (left) in Experiment 1. As the model predicts, a linear trend fits the data.



**Figure 10.** Predicted and observed error rates for the five input methods in Experiment 1 (excluding  $D_t = 600$  ms).

showed the latest response, about 10 to 20 ms later than the others ( $p < 0.012$ ).

We also analysed the interaction effect between input method and  $ID_t$  (see Figure 10) on error rate. Pairwise comparison showed that for lower IDs (2.17, 2.59, and 3.37 bits), the fast input methods generated more errors. Touch-Down showed a 4.9–10.3% higher error rate ( $M = 6.81\%$ ) than Touch-Maximum ( $p < 0.039$ ) or Key-Press ( $p < 0.02$ ). However, for higher IDs (2.95, 3.37, 4.17, and 4.59 bits), slow inputs had a higher error rate than faster inputs. Touch-Release and Key-Release showed a 4.4–9.2% higher error rate ( $M = 5.95\%$ ) than Touch-Maximum ( $p = 0.046$ ), Touch-Down ( $p < 0.022$ ), or Key-Press ( $p < 0.045$ ).

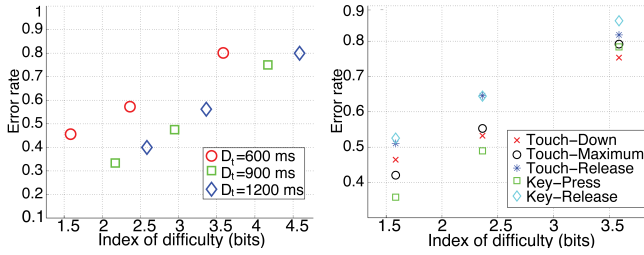
### Model Fitting

Practically we do not need to measure  $\sigma$  and  $\mu$  to obtain  $c_\sigma$  and  $c_\mu$ . Instead, the final model in Equation 7 requires only IDs (known) and corresponding error rates (empirical) for finding the parameters. For this part, the model was fitted against average error rates in all nine ID conditions. When fitting all data, including  $D_t = 600$  ms, we obtained  $R^2 = 0.744$ . For this model,  $c_\sigma$  is 0.0852 ( $p = 0.001$ ) and  $c_\mu$  is 0.0634 ( $p = 0.27$ ).

However, without that condition ( $D_t = 600$  ms), model fit was much higher, at  $R^2 = 0.997$ . The input methods were covered well by the model, with Touch-Maximum, Key-Press, Touch-Down, Touch-Release, and Key-Release at  $R^2 = 0.998$ , 0.984, 0.974, 0.977, and 0.977, respectively. The estimates obtained for  $c_\mu$  and  $c_\sigma$  are given in Table 1.

Input method	$c_\mu$	$c_\sigma$	$R^2$
Touch-Down	0.131	0.0805	0.974
Touch-Maximum	0.183	0.0773	0.998
Touch-Release	0.161	0.0833	0.977
Key-Press	0.168	0.0705	0.984
Key-Release	0.204	0.0864	0.977
Overall	0.169	0.0795	0.997
Including 600 ms	0.0634	0.0852	0.744

**Table 1.** Model parameters and fitness scores by input method in Experiment 1. Key-Press showed the least input noise of all the methods. Touch-Maximum was the best among the touch-based inputs.



**Figure 11. Error rate increases rapidly with the fastest tapping condition ( $D_t=600$  ms, left). An input device with slower sensor events suffered more from the problem at  $D_t=600$  ms (right).**

### Shortest Target Distance

We examined more closely why the model fit is much higher when the shortest distance condition (600 ms) is excluded. It reveals a lower limit to the range of  $D_t$  conditions explainable by our model.

Plotting error rate versus  $ID_t$  shows an anomaly in our data at  $D_t = 600$  ms (left, Figure 11). The spread of responses increases rapidly. In fact, this effect was predicted by the WK model [38] and by rate-limit studies of sensorimotor synchronisation [28]. These have shown that the decrease in the spread of responses stops below 600 ms for a visually presented stimulus. Two factors explain this: 1) noise from the response-execution stage becomes greater with faster tapping and 2) there is lack of preparation time for the internal clock.

The hypothesis gains more support when one compares the input methods. An input device with slower sensor events should suffer more from the problem at  $D_t = 600$  ms, because it will have more latency. Indeed, we found that the interaction effect between  $ID_t$  and input method was significant:  $F(9.15,155.51)=2.24$ ,  $p=0.022$  (right, Figure 11). If we consider only the three  $ID_t$ s that include the 600 ms target distance, pairwise comparisons show consistently higher error rates for the input methods with the latest response (Touch-Release and Key-Release). Touch-Release and Key-Release showed a 6.4–16.7% higher error rate ( $M=11.87\%$ ) than Touch-Down ( $p<0.038$ ), Touch-Maximum ( $p<0.039$ ), or Key-Press ( $p<0.004$ ).

### Musicians' Greater Precision

We can also report that participants with musical experience benefited more from the Touch-Maximum method in relative terms. For musicians,  $c_\sigma$  for this method was 0.0578 ( $p<0.001$ ) and  $c_\sigma$  for Key-Press was 0.0556 ( $p<0.001$ ). This indicates near-perfect aiming in this condition. Musicians also showed lower noise for Touch-Release ( $c_\sigma=0.0655$ ,  $p<0.001$ ) than Touch-Down ( $c_\sigma=0.0703$ ,  $p=0.0015$ ). By contrast, non-musicians showed higher noise for Touch-Maximum ( $c_\sigma=0.0921$ ,  $p<0.001$ ) than Key-Press ( $c_\sigma=0.0803$ ,  $p<0.001$ ).

### Summary

In summary, the basic assumptions are supported by the data, on the assumption when we concentrate on the range  $D_t > 600$  ms. In that range, our model predicts error rates in temporal pointing up to 0.997 of  $R^2$ . The model also captures differences between popular input methods and provides

a novel explanation for them. It explains the differences in terms of implicit aim point and variability. We return to this finding in the 'Discussion' section.

## EXP. 2: VISUALLY CUED TEMPORAL POINTING

The second study looked at a more realistic, dynamically changing task. Our case is a popular mobile game called *Flappy Bird*, illustrated in Figure 12. At the late January 2014, the game suddenly became the most downloaded game in the iOS store. The gamer's goal is to pass as many 'pillars' as possible by flying between them. The game involves no other controls than tapping the display, but it has been criticised for its extreme difficulty. Comparing this to the blinking experiment, we expect a lower value of  $c_\sigma$  (that is, less spread), because visual cues reduce uncertainty in the internal clock.

First, we analyse game performance in terms of temporal pointing, dividing the game into two, different temporal tasks.

### The Flying Dynamics of the Bird

To determine the temporal targets, we firstly need to understand how the bird moves. Each tap makes the bird 'jump' in accordance with a pre-defined trajectory. The shape of the trajectory is governed by three motion parameters: horizontal velocity of the bird  $v_x$  (pixels/frame), downward gravity  $g_y$  (pixels<sup>2</sup>/frame), and the initial vertical velocity of the bird  $v_{y0}$  after a jump. In jumping, the vertical velocity of the bird is updated by  $v_{y0}$  in the upward direction. With Newtonian kinematics, the position of the bird ( $x_b, y_b$ ) after a jump can be expressed as a function of time  $t$ :

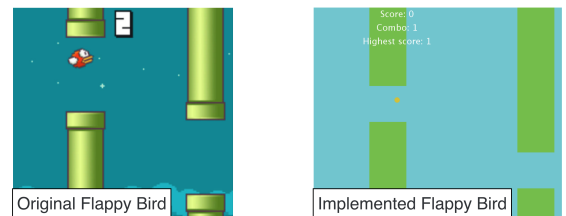
$$\begin{aligned} x_b &= x_0 + v_x \times t \\ y_b &= y_0 - g_y/2t(t - D_t) \quad \text{where } D_t = 2v_{y0}/g_y \end{aligned} \quad (8)$$

Note that  $t$  represents a timeframe in the game.  $x_0$  and  $y_0$  represent the global location of the bird at jump initiation. After  $D_t$  timeframes have passed, the bird will return to initial height  $y_0$  because of the gravity.

### Analysing Temporal Targets in Flappy Bird

We now analyse the timing of jumps in the pillar landscape as a temporal pointing task. We convert the spatial relationship between the bird and a pillar to a temporal target.

Firstly, we analyse temporal width  $W_t$ . When there is no pillar near the bird, a target can be assumed to allow very low temporal precision. We exclude these very easy jumps as irrelevant.



**Figure 12. The Flappy Bird game (left) and our experimental replication of it (right). The game requires manoeuvring a bird with jumps through a pillar landscape.**



Figure 13 shows two distinct regions, A and B. The regions are defined by boundaries 1 and 2, marked in the figure. As region A allows wider space for the bird to fly without collision, the temporal target has a larger  $W_t$ . We can relate both boundaries to the bird's motion as is shown below:

$$\text{Boundary 1: } y = h - gD_t^2/8$$

$$\text{Boundary 2: } y = \begin{cases} -g/2v_x^2(x - v_x D_t/2)^2 & : x \in [0, v_x D_t/2] \\ 0 & : x > v_x D_t/2 \end{cases} \quad (9)$$

Note that  $h$  represents the height of the tunnel minus the bird's diameter. Now by substituting  $y_b$  in Equation 8 into  $y$  of Equation 9, we can calculate the time  $T_1$  and  $T_2$  when the bird crosses each boundary 1, 2 after a jump. Time gap ( $T_2 - T_1$ ) was averaged over  $(x_0, y_0)$  to obtain averaged  $W_t$  for a region as shown below:

$$W_t = \frac{1}{nm} \sum_{x_0}^n \sum_{y_0}^m (T_2 - T_1) \quad (10)$$

Next,  $D_t$  can be determined by means of Equation 8. We assume that a player tries to maintain a flying altitude by repeatedly executing jumps. The distance between those jumps then becomes the target distance ( $D_t$  in Equation 8).

## Method

### Participants

We recruited 10 participants (5 males, 5 females) from a local university. Mean age was 29 (SD=8.5). Seven reported familiarity with touch-gaming or music.

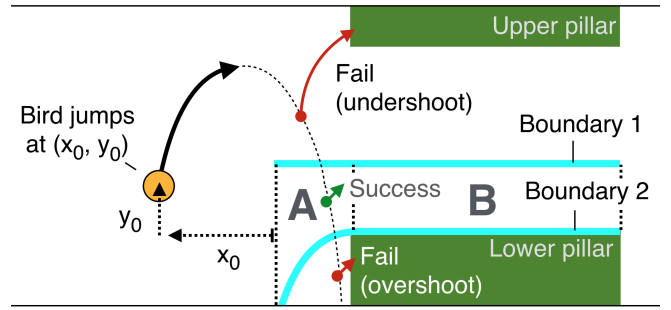
### Task and Materials

We implemented a Java version of Flappy Bird, emulating the original game. In our application, the bird is represented with an orange circle. The objective is to cross as many pillars as possible. If the bird hits a pillar, the game restarts from the beginning with a score of 0. Collision with a pillar is determined from the edge of the circle, as in the original game. The number of pillars crossed is shown on the screen, together with a high score. The frame rate of the game was maintained at 60 fps, and all target distances and target widths are converted to the time measure at this rate.

### Experimental Design

Two temporal target distances and three temporal target widths were used to create levels in the experiment. Target distance  $D_t$  was determined from Equation 8, with two separate gravities used, 0.22 and 0.11 pixels<sup>2</sup>/frame, while  $v_{y0}$  was fixed at 5 pixels/frame. We chose to change the gravity to maintain a consistent feeling of control for the bird when the user was initiating a jump. The resulting target distances are 757.6 and 1515.2 ms. Order of levels was randomised.

Target width  $W_t$  was determined from Equation 10. We fixed the diameter of the bird as 20 pixels (4.6 mm) and changed the distance between pillars to control  $W_t$ . The heights of the pillar openings were 132, 94, and 84 pixels for 0.22 gravity and 190, 150, and 140 pixels for 0.11 gravity. The resulting  $W_t$ s for region A were 206.9 ms, 83.7 ms, and 54.7 ms for the 0.22 gravity condition and 238.9 ms, 119.8 ms, and 91.2 ms for 0.11 gravity. The resulting widths for region B were 189.4



**Figure 13. Analysis of temporal targets in the game in Experiment 2.** Regions A and B are safe. When a jump is made before region A or B, the bird will collide with the upper pillar, undershooting the temporal target. When a jump is made after the A or B region, the bird will collide with the lower pillar, overshooting the temporal target. Combining this with an analysis of the jumping trajectory, we can calculate temporal targets.

ms, 54.6 ms, and 21.2 ms for the 0.22 gravity condition and 189.4 ms, 54.6 ms, and 21.2 ms for 0.11 gravity. Horizontal speed of the bird  $v_x$  was maintained at 3 pixels/frame, and the width of the tunnel was fixed at  $D_t \times v_x$  to ensure that both region A and region B could be encountered uniformly. The exact values are given also in the appendix.

### Procedure

The participants were asked to play the game to the best of their ability. They were encouraged to beat the other participants' high score. The beginning location of the bird was randomly perturbed within a pillar width. The distance between pillars was fixed at three times the tunnel width. The vertical locations of tunnels were perturbed for every game within 400 pixels from the centre of the display. Playing time was fixed at six minutes for the higher gravity and 12 minutes for the lower gravity level. Total play time was one hour.

Participants used the ENTER key to jump, using the same keyboard as in Experiment 1. We recorded the score of the participants and the location of the bird for all jumps. Collisions were also recorded, with collisions with the upper pillar distinguished from those with the lower pillar. We can obtain the error rate for each region as is shown below:

$$E_A = \frac{\text{Deaths in region A}}{\text{Trials in region A}} \quad E_B = \frac{\text{Deaths in region B}}{\text{Trials in region B}}$$

### Model Fitting

The two parameters  $c_\mu$  and  $c_\sigma$  were used for both regions, A and B. This assumes that only the temporal target changes in those regions, not the implicit aim point or variability.

The overall fit was high:  $R^2 = 0.87$ ; see Figure 14. As expected,  $c_\sigma$  was low, at 0.00802 ( $p=0.012$ ). This is only 11.38% of that observed in the blinking study. Theoretically, it indicates lower noise in the internal clock arising from visual motion cues (see Figure 15).  $c_\mu$  was estimated to be 0.0741 ( $p=0.0182$ ). This is 44.2% of the value in Experiment 1. In other words, players tried to jump the bird closer to boundary 1, rather than boundary 2. Boundary 1 is not visible as boundary 2 is. Perhaps users tried to avoid the lower pillar.

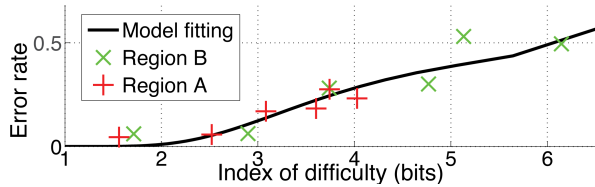


Figure 14. Experiment 2 modelling results: The model shows high correlation ( $R^2 = 0.87$ ) with empirical data from game regions A and B.

### Cross-Validation

We performed a repeated twofold cross-validation to test the generalisability of our model. The model parameters were obtained over nine randomly chosen levels in the game and tested on the rest (3). We use mean absolute error (MAE) between model prediction and the other three game conditions as the fitness score. Over 100 iterations, we obtain a mean MAE of 3.1% (SD=0.22). In other words, after the parameters are obtained, the model can be used to predict the user's performance with high accuracy.

## DISCUSSION

This paper has presented a novel mathematical model for understanding error rates in temporal pointing. The task is common in HCI, but no previous model deals with the case involving a bounded target. The data collected here lend considerable credence to the model.

The model provides a novel explanation to differences among input devices in temporal aspects of performance. It predicts that users aim for temporal targets differently, with variation by input method, and it points out three stages that affect variability in responses and thereby error rates. This explains a hidden aspect of the finding that touchscreen gaming is associated with much lower performance than gaming with physical keys [5, 42, 43]. This difference is not trivially due to latency, as one could easily think. The model exposes two factors. Firstly, targeting with the touchscreen is hard because the timing of the sensor event is uncertain. We showed that coinciding of the input event and the moment when the finger has maximum impact on the surface aids in targeting. The Touch-Maximum event allows a user to learn a mapping that is more predictable and therefore improves accuracy. With Touch-Maximum, error rates were decreased by up to 9.0% from the empirical observation. Secondly, Touch-Maximum still suffered from greater noise ( $c_\sigma = 0.0773$ ) than Key-Press ( $c_\sigma = 0.0705$ ). Users cannot precisely control how high they hold their finger, and the variation in distance between the finger and surface may have increased variability.

We also noted that two participants in Experiment 1 did not exhibit Gaussian response distributions, and we excluded

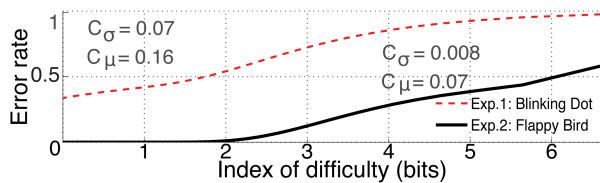


Figure 15. Comparison between data from Exp. 1 and Exp. 2: Despite its apparent difficulty, temporal pointing is easier with visual cueing.

them from the modelling. We suspect that their non-Gaussian patterns were due to intermittent change in strategy. To make the model more universally applicable, strategies not covered by reactive and proactive approaches deserve more attention.

Experiment 2 shows that performance in some complex, dynamic targeting tasks can be modelled successfully. Dividing the game into two distinct regions associated with two types of temporal targets, we obtained high fit for the model.

Our findings point out that the scope of the model is limited to  $D_t > 600$  ms. Besides determining the scope of the model also for other parameters, challenges for future work include considering how to combine temporal with spatial pointing. This would allow dealing with complex responses in, for example, touch games. To improve the model, we also need to understand how changes in task, user, or conditions affect  $\mu$  and  $\sigma$ . Our studies assumed relatively simple and predictable temporal targets, but in many real-world interactions the targets come and go, each with properties that have to be inferred on the fly.

We envision some possibilities in using the model to design better interactions. Firstly, the model could be used analytically to tune the difficulty of temporal pointing in interactive tasks. For example, one could use the model to design game levels or dynamic interaction techniques. This can be done by following the same procedure we used in Experiment 2 to determine temporal targets. Secondly, dynamic adaptation of temporal targets could be tried. In other words, the difficulty of a task could be changed dynamically to fit the level of the user. If we can obtain an estimate of the model parameters for a user, perhaps from logging data, we can subtly tune the size of the target. This might, for example, allow a novice gamer to enjoy gaming with experts without a noticeable change in appearance. Alternatively, such a personalised model could be used to determine when to provide additional cues to facilitate temporal performance. Finally, the model informs the development of input methods. It makes the prediction that of equal importance with minimal latency is that the device allow accurate and reliable implicit aiming. This means designing the input event in such a way that it provides a physically clear input timing to the user.

## ACKNOWLEDGMENTS

This project has received funding from the Academy of Finland project COMPUTED and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 637991).

## APPENDIX: TEMPORAL TARGETS IN EXPERIMENT 2

	ID(bits)	Tunnel height(px)	Gravity(px <sup>2</sup> /fr)	$D_t$ (ms)	$W_t$ (ms)
Region A	1.56	132	0.22	757.6	256.8
	3.08	94	0.22	757.6	89.4
	3.74	84	0.22	757.6	56.6
	2.52	190	0.11	1515.2	263.9
	3.60	150	0.11	1515.2	124.7
	4.03	140	0.11	1515.2	92.9
Region B	1.71	132	0.22	757.6	231.2
	3.74	94	0.22	757.6	56.7
	5.13	84	0.22	757.6	21.6
	2.90	190	0.11	1515.2	203.4
	4.77	150	0.11	1515.2	55.6
	6.15	140	0.11	1515.2	21.4

## REFERENCES

1. Luigi Acerbi, Daniel M Wolpert, and Sethu Vijayakumar. 2012. Internal representations of temporal statistics and feedback calibrate motor-sensory interval timing. (2012).
2. Rasmus Bååth. 2015. Estimating the distribution of sensorimotor synchronization data: A Bayesian hierarchical modeling approach. *Behavior research methods* (2015), 1–12.
3. James J Belisle. 1963. Accuracy, reliability, and refractoriness in a coincidence-anticipation task. *Research Quarterly. American Association for Health, Physical Education and Recreation* 34, 3 (1963), 271–281.
4. Anne-Marie Brouwer, Jeroen BJ Smeets, and Eli Brenner. 2005. Hitting moving targets: effects of target speed and dimensions on movement time. *Experimental Brain Research* 165, 1 (2005), 28–36.
5. Kimberly Chu and Chui Yin Wong. 2011. Mobile input devices for gaming experience. In *User Science and Engineering (i-USEr), 2011 International Conference on*. IEEE, 83–88.
6. ERFW Crossman and PJ Goodeve. 1983. Feedback control of hand-movement and Fitts' law. *The Quarterly Journal of Experimental Psychology* 35, 2 (1983), 251–278.
7. David A Engström, JA Scott Kelso, and Tom Holroyd. 1996. Reaction-anticipation transitions in human perception-action patterns. *Human Movement Science* 15, 6 (1996), 809–832.
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. Khai-chung Gan and Errol R Hoffmann. 1988. Geometrical conditions for ballistic and visually controlled movements. *Ergonomics* 31, 5 (1988), 829–839.
10. 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*. ACM, 91–100.
11. Peter A Hancock and Karl M Newell. 1985. The movement speed-accuracy relationship in space-time. In *Motor Behavior*. Springer, 153–188.
12. Errol R Hoffmann. 2013. Which version/variation of Fitts law? A critique of information-theory models. *Journal of motor behavior* 45, 3 (2013), 205–215.
13. Tsung-Yu Hsieh, Yeou-Teh Liu, Gottfried Mayer-Kress, and Karl M Newell. 2013. The movement speed-accuracy relation in space-time. *Human movement science* 32, 1 (2013), 257–269.
14. Ray Hyman. 1953. Stimulus information as a determinant of reaction time. *Journal of experimental psychology* 45, 3 (1953), 188.
15. Steven W Keele. 1968. Movement control in skilled motor performance. *Psychological bulletin* 70, 6p1 (1968), 387.
16. Seonjin Kim, Les G Carlton, Yeou-Teh Liu, and Karl M Newell. 1999. Impulse and Movement SpaceTime Variability. *Journal of motor behavior* 31, 4 (1999), 341–357.
17. David N Lee and others. 1976. A theory of visual control of braking based on information about time-to-collision. *Perception* 5, 4 (1976), 437–459.
18. Jiří Mates, Ulrike Müller, Tomáš Radil, and Ernst Pöppel. 1994. Temporal integration in sensorimotor synchronization. *Journal of cognitive neuroscience* 6, 4 (1994), 332–340.
19. David E Meyer, Richard A Abrams, Sylvan Kornblum, Charles E Wright, and JE Keith Smith. 1988. Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological review* 95, 3 (1988), 340.
20. David E Meyer, JE Smith, and Charles E Wright. 1982. Models for the speed and accuracy of aimed movements. *Psychological Review* 89, 5 (1982), 449.
21. George P Moore and Jessie Chen. 2010. Timings and interactions of skilled musicians. *Biological cybernetics* 103, 5 (2010), 401–414.
22. KM Newell. 1980. 30 The Speed-Accuracy Paradox in Movement Control: Errors of Time and space. *Advances in Psychology* 1 (1980), 501–510.
23. KM Newell, Les G Carlton, and Seonjin Kim. 1994. Time and space-time movement accuracy. *Human Performance* 7, 1 (1994), 1–21.
24. Réjean Plamondon and Adel M Alimi. 1997. Speed/accuracy trade-offs in target-directed movements. *Behavioral and Brain Sciences* 20, 02 (1997), 279–303.
25. Jeff Pressing. 1999. The referential dynamics of cognition and action. *Psychological Review* 106, 4 (1999), 714.
26. Roger Ratcliff and Hans PA Van Dongen. 2011. Diffusion model for one-choice reaction-time tasks and the cognitive effects of sleep deprivation. *Proceedings of the National Academy of Sciences* 108, 27 (2011), 11285–11290.
27. Bruno H Repp. 2005. Sensorimotor synchronization: a review of the tapping literature. *Psychonomic bulletin & review* 12, 6 (2005), 969–992.
28. Bruno H Repp. 2006. Rate limits of sensorimotor synchronization. *Advances in Cognitive Psychology* 2, 2-3 (2006), 163–181.

29. Bruno H Repp and Yi-Huang Su. 2013. Sensorimotor synchronization: a review of recent research (2006–2012). *Psychonomic Bulletin & Review* 20, 3 (2013), 403–452.
30. Andries F Sanders and Andries Sanders. 2013. *Elements of human performance: Reaction processes and attention in human skill*. Psychology Press.
31. Richard A Schmidt. 1968. Anticipation and timing in human motor performance. *Psychological Bulletin* 70, 6p1 (1968), 631.
32. Richard A Schmidt and David E Sherwood. 1982. An inverted—U relation between spatial error and force requirements in rapid limb movements: Further evidence for the impulse-variability model. *Journal of Experimental Psychology: Human Perception and Performance* 8, 1 (1982), 158.
33. Richard A Schmidt, Howard Zelaznik, Brian Hawkins, James S Frank, and John T Quinn Jr. 1979. Motor-output variability: a theory for the accuracy of rapid motor acts. *Psychological review* 86, 5 (1979), 415.
34. J Tresilian, J Oliver, and T Carroll. 2003. Temporal precision of interceptive action: differential effects of target size and speed. *Experimental brain research* 148, 4 (2003), 425–438.
35. James R Tresilian. 2005. Hitting a moving target: perception and action in the timing of rapid interceptions. *Perception & Psychophysics* 67, 1 (2005), 129–149.
36. James R Tresilian and Andrew Lonergan. 2002. Intercepting a moving target: effects of temporal precision constraints and movement amplitude. *Experimental Brain Research* 142, 2 (2002), 193–207.
37. Eric-Jan Wagenmakers, Han LJ Van Der Maas, and Raoul PPP Grasman. 2007. An EZ-diffusion model for response time and accuracy. *Psychonomic bulletin & review* 14, 1 (2007), 3–22.
38. Alan M Wing and AB Kristofferson. 1973a. The timing of interresponse intervals. *Perception & Psychophysics* 13, 3 (1973), 455–460.
39. Alan M Wing and Alfred B Kristofferson. 1973b. Response delays and the timing of discrete motor responses. *Perception & Psychophysics* 14, 1 (1973), 5–12.
40. Jacob O Wobbrock, Edward Cutrell, Susumu Harada, and I Scott MacKenzie. 2008. An error model for pointing based on Fitts' law. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1613–1622.
41. Daniel M Wolpert and Michael S Landy. 2012. Motor control is decision-making. *Current opinion in neurobiology* 22, 6 (2012), 996–1003.
42. Loutfouz Zaman and I Scott MacKenzie. 2013. Evaluation of nano-stick, foam buttons, and other input methods for gameplay on touchscreen phones. In *International Conference on Multimedia and Human-Computer Interaction-MHCI 2013*. 69–1.
43. Loutfouz Zaman, Daniel Natapov, and Robert J Teather. 2010. Touchscreens vs. traditional controllers in handheld gaming. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology*. ACM, 183–190.