How Relevant is Hick's Law for HCI?

Wanyu Liu, Julien Gori, Olivier Rioul, Michel Beaudouin-Lafon, Yves Guiard

HAL Id: hal-02867301

https://hal.archives-ouvertes.fr/hal-02867301

Submitted on 14 Jun 2020

How Relevant is Hick's Law for HCI?

Wanyu Liu[1,2,3]   **Julien Gori**[3,2]   **Olivier Rioul**[2]   **Michel Beaudouin-Lafon**[3]   **Yves Guiard**[3,2]

[1]STMS, IRCAM, CNRS, Sorbonne Université, F-75004, Paris, France
[2]LTCI, Télécom Paris, Institut Polytechnique de Paris, F-91120, Palaiseau, France
[3]Université Paris-Saclay, CNRS, Inria, Laboratoire de Recherche en Informatique, F-91400, Orsay, France
aliu@ircam.fr, {jgori,mbl}@lri.fr, {olivier.rioul,yves.guiard}@telecom-paristech.fr

ABSTRACT

Hick's law is a key quantitative law in Psychology that relates reaction time to the logarithm of the number of stimulus-response alternatives in a task. Its application to HCI is controversial: Some believe that the law does not apply to HCI tasks, others regard it as the cornerstone of interface design. The law, however, is often misunderstood. We review the choice-reaction time literature and argue that: (1) Hick's law speaks against, not for, the popular principle that 'less is better'; (2) logarithmic growth of observed temporal data is not necessarily interpretable in terms of Hick's law; (3) the stimulus-response paradigm is rarely relevant to HCI tasks, where choice-reaction time can often be assumed to be constant; and (4) for user interface design, a detailed examination of the effects on choice-reaction time of psychological processes such as visual search and decision making is more fruitful than a mere reference to Hick's law.

Author Keywords

Hick's law, the Hick-Hyman law, stimulus-response, choice reaction time, information, uncertainty, logarithm, convexity

CCS Concepts

•**Human-centered computing → HCI theory, concepts and models;**

INTRODUCTION

Together with Fitts' law [13], Hick's law, or the Hick-Hyman law [20, 21] was first introduced to human-computer interaction (HCI) by Newell, Card and colleagues [6, 33] in the early 1980s as a contribution of psychology to the design of human-computer interfaces. In Chapter 2 of *The Human Information-Processor* [6, p.27], they describe Hick's law as the time T a person takes to make a decision among a set of n choices:

$$T = a + b \times H \tag{1}$$

where a and b are empirically determined constants and H is the uncertainty captured by the information-theoretic notion

of entropy, defined[1] by

$$H = \log_2 n \tag{2}$$

when the n choices are equally probable or by

$$H = \sum_{i=1}^{n} p_i \log_2(1/p_i). \tag{3}$$

when the n choices have different probabilities of occurrence p_i. Thus increasing the stimulus uncertainty increases decision time.

The rationale behind the use of entropy is to consider the human mind, in light of Shannon's theory, as a communication system that transmits information from input (sensory and perceptual processes) to output (overt responses) [44]. Information theory was popular among experimental psychologists in the 1950s as a tool to investigate human performance in various tasks, notably Hick and Hyman's choice reaction time [20, 21] and Fitts' speed-accuracy trade-off when executing aimed movement [13].

The choice-reaction paradigm is a classic topic in the study of human performance that continues to be relevant to contemporary psychology. Many models, among Hick's law, have been developed to describe the time it takes to respond to a given stimulus, including accumulation [37] and memory-based models [2]. Proctor and Schneider [36] provide a comprehensive review of these accounts. In this paper, we focus on the use of Hick's law and the choice-reaction paradigm in HCI given its interesting position in the community:

- While the law is ignored in many HCI textbooks [38, 45], it is taught in HCI classes as one of the few quantitative laws in psychology. We interviewed a few HCI professors on their understanding of Hick's law and received replies such as *"it's about response time", "decision making time", "I have two slides about the law, adapted from the slides by someone else", "I teach it but I don't feel 100% comfortable about talking about it"*.

- Conceptually, it is often regarded as a fundamental law of interface design. It appears in design books, e.g. Lidwell et al. [25], and numerous on-line articles discussing how understanding Hick's law can improve interface design, e.g. Nikolov [34] or Sauro [41]. It has been claimed to

---

[1]Note that in the book [6, p.27], the formulas are presented as $\log_2(n+1)$ and $\sum_{i=1}^{n} p_i \log_2(1/p_i + 1)$ respectively. A more detailed discussion about n *vs.* $n+1$ can be found in [48].

apply to a large number of contexts, including menu design, device settings and road signs. Essentially, when faced with a set of choices, designers are guided by this "Hick-based" design principle with the concept *less is better*, i.e. it is better to split the set of choices into smaller categories, instead of overwhelming users with all choices at once.

- In practice, the law has not seen many successful applications. Only a few HCI studies have capitalized on Hick's law, e.g. Soukoreff & Mackenzie [49]. In 2005, Seow [43] compared Hick's law and Fitts' law, two major information-theoretic laws, and examined the possible reasons, such as the complexity of computing information measures, for the failure of Hick's law to gain momentum in the field. Nevertheless, few studies have incorporated Hick's law into their work since then, e.g. Cockburn et al. [9].

The lack of comprehensive understanding of Hick's law may explain why many HCI researchers have not ventured to apply it to interaction tasks. Furthermore, there seems to be different definitions of Hick's law. While for psychologists the law exclusively has to do with the choice-reaction paradigm, HCI researchers seem to apply it whenever choices are presented to the user, including for visual search time, e.g. [23, 28, 50], decision time, e.g. [9] or reaction time, e.g. [39]. Does the law really apply to these settings?

The goal of this article is to dispel some misunderstandings about Hick's law in HCI studies by providing a clearer picture of the choice-reaction paradigm. First, we revisit the historical context of the choice-reaction paradigm in psychology and re-examine HCI studies that have used Hick's law. We demonstrate that a number of logarithmic-time phenomena observed in HCI are not explained by Hick's law; conversely, we argue that choice-reaction time does not always scale logarithmically with the number of choices – it can often be considered constant. We then further examine the so-called "Hick-based" design principles using a mathematical analysis and show that they cannot be justified by Hick's law. We conclude with the practical implications for HCI of this new look at Hick's law.

CHOICE-REACTION TIME IN PSYCHOLOGY
There is a long tradition in psychology of studying the choice-reaction paradigm. In this section, we review these studies and attempt to clarify the definition of Hick's law.

Before Information Theory
The first results on reaction time (RT) are due to Helmholtz [18], the famous physician/physicist of the nineteenth century. He determined that signals travel the nervous system at about 60 m/s. That the observed RT was longer than the calculated signal traveling time suggests that humans were not simply hard-wired and that time was required for additional perceptual and cognitive processes ("perceiving" and "willing"). By the end of the nineteenth century and early twentieth century, three other important results were known:

- Donders [12] introduced the three-class taxonomy of reaction time that is still in use today: *simple reaction time* is the time it takes to react, with a predetermined response, to the onset of a stimulus whose identity is known in advance

but whose time of occurrence is uncertain; *choice-reaction time* is the time it takes to react to the onset of one of several possible stimuli, following a given stimulus-response mapping rule; and *go-no-go reaction time* is the time it takes to respond to a stimulus that may or may not occur at a predetermined point in time. Donders showed that *simple RT* was the shortest and *choice RT* the longest.

- Merkel [29] performed an experiment showing for the fist time that *choice reaction time* varies as a smooth monotonic function of the number of alternatives: the larger the set of stimuli, the longer the RT.

- Many psychologists have attributed the "RT phenomenon" to such causes as the division of attention or a reduction in the effective intensity of the stimulus without providing a quantitative theory. Blank [3] was the first to postulate a logarithmic relationship between reaction time and the number of alternatives but did not give further explanation.

The Information Analogy: Hick & Hyman
Hick's work was strongly inspired by Merkel's results. In fact one could say that Hick did nothing but replicated Merkel's experiment and used a logarithmic scale for the x-axis rather than Merkel's linear scale. This would miss an important point: Hick's contribution is conceptual rather than experimental. Using the information-theoretical rationale that was popular at the time, Hick interpreted this logarithmic curve by considering the human as a channel of information transmission [20]. Accordingly, reaction time is seen as resulting from the uncertainty of the stimulus, which can only be processed at some maximum rate. The information rationale had an immediate effect: if the "information" – in layman's sense – mattered, then all the ways in which information could be varied should matter (e.g. the frequency of each stimuli should matter). This introduced new ways of testing the relationship between choice reaction time and the "information" provided by the stimulus.

Hick [20] conducted three experiments with the same apparatus: ten lamps were arranged in an irregular circle and connected to a device that was coded to light up a random lamp every 5 seconds [19]. Each of the participants' fingers was connected to a Morse key[2] corresponding to a lamp. The relationship between the keys and lamps was predetermined. The participants' task was to press the key corresponding to the lamp that lit up. Both stimulus presentation and response were recorded in binary code by moving paper. A uniform distribution of stimuli was used.

The first experiment was carried out to confirm the fitness of $\log(n+1)$ to reaction time *RT*, rather than $\log n$, which is the entropy formula accounting for equally probable choices in information theory [44]. Hick himself served as the only participant in the first experiment, varying the number of stimuli from 2 to 10. He trained himself over 8,000 trials before the experiment and removed incorrect reactions, therefore the transmitted information equals the entropy of stimuli. As

---

[2]Also known as telegraph key: a switching device used primarily to send Morse code.

noted, he found that $a + b \times \log(n+1)$ gave a better fit than $a + b \times \log n$.

The second experiment, which included the author and another participant, investigated reaction time in the case of incomplete information transmission. Both participants were well trained before the experiment and were instructed to make errors, which were included in the analysis. Hick estimated the joint probability distribution between the stimuli and the actual reactions and computed the transmitted information[3]. Then he introduced the notion of *degree of choice* n_e where $I(X;Y) = \log n_e$, and showed that in the case of partially transmitted information, $\log(n_e + 1)$ also provided a better fit than $\log n_e$. He therefore concluded that *"the amount of information extracted R is proportional to the time taken to extract it, on the average"*, where R is defined as

$$R = \log(n+1) \quad \text{or} \quad R = \log(n_e + 1) \qquad (5)$$

with $n_e = n$ if no errors are made and all information gets transmitted. The term "+1" accounts for the participants' uncertainty about the "no stimulus" condition. If RT is proportional to R, then the rate of information gain is constant.

Hick plotted his data as a function of the number of alternatives (n or n_e) and did not explicitly postulate a linear relationship between choice reaction time RT and the transmitted information R. In 1953, Crossman [10] used a card-sorting task and plotted data as a function of R. The same year, Hyman [21] also plotted data as a function of R and suggested that a linear function of stimulus information within the range of 0 to 3 bits could be considered.

Hyman [21] varied the entropy of the stimuli in three ways:

- by changing the total number n of stimuli (Hick [20], Merkel [29]). In this case, uncertainty increases with the number of stimuli;

- by changing the probability that each stimulus is indeed activated. The more similar the probabilities, the higher the uncertainty. In the limit case, when all the stimuli are activated with equal probability, uncertainty reaches its maximum value of $\log n$;

- by establishing "grammar rules", i.e. introducing conditional probabilities between successive stimuli. For example, if stimuli B is activated, then it is certain that stimuli D will be activated next.

Choice Reaction Time: Results
After Hick's and Hyman's respective experiments, a number of studies measuring reaction time in a choice-reaction task were reported. Here we summarize the main results.

---

[3]The mutual information between two discrete random variables X and Y is defined as

$$I(X;Y) = \sum_y \sum_x p(x,y) \log \frac{p(x,y)}{p(x)p(y)} = H(X) - H(X|Y) \quad (4)$$

where $p(x,y)$ is the joint probability distribution of X and Y, and $p(x)$ and $p(y)$ are the marginal probability distributions of X and Y, respectively.

Choice Reaction Time: Context
Most studies were conducted with the goal of measuring choice reaction time on very simple tasks, such as responding to bulbs lighting up by pressing keys (Hick [20]) or through speech (Hyman [21]). It is not clear how well the logarithmic relationship between time and information holds for more complex tasks that potentially require a lot of decision making. One exception is given by Crossman [11], who investigated card sorting and found results consistent with the rest of the literature.

Measuring Information: Entropy or Mutual Information?
It has consistently been found that for a range between 1 and 3 to 4 bits, choice reaction time increases linearly with stimulus entropy, irrespective of whether the number of stimuli n, the probability associated with each stimulus, or the sequential dependency between consecutive stimuli is being manipulated. However, whenever the number of choices becomes larger, it seems that reaction time is consistently over-estimated. In fact, Fitts & Posner [16] indicated that whatever the number of possible stimuli, reaction time will seldom exceed 1s. Fitts et al. [15] showed that the response to very low probability alternatives is faster than predicted by the law. On the other hand, Pollack [35] found that the linear relationship extends to about 10 bits in a task where words had to be named. The actual range where the relationship holds is thus very dependent on the actual task.

Reaction time is underestimated by Hick's law when many mistakes are made by the participant. Fitts [14] reported that beyond 0.6 bits of equivocation[4], the loss of information resulting from errors increases faster than the temporal gain allowed by increased response speed. From now on we use the term *stimulus uncertainty* rather than the vague term of "information" to characterize the stimuli.

Effect of Stimulus-Response Compatibility
Stimulus-Response (S-R) compatibility was introduced in psychology to characterize the fact that it is easier to respond to a stimulus using certain responses than others. If the stimulus is coded in terms of digits appearing on a screen, it is for example much easier for someone to call out the corresponding digit than, say, to call a letter or another digit out. It has repeatedly been found that the better the S-R compatibility, the shallower the slope relating reaction time to stimulus uncertainty. In fact, whereas Hick's light and key experiment reports rates of about 200 ms/bit, there are many cases where the slope approaches 0 ms/bit (Fig. 1). This is the case for Leonard's experiment [24] in which the subject rested his fingers upon vibrators and pressed the vibrator that was activated, or Mowbray's experiment [30] where subjects reacted to the visual presentation of Arabic numerals by merely reading them aloud. In such cases, which correspond to extremely high levels of S-R compatibility, there is virtually no effect of stimulus uncertainty on reaction time.

---

[4]Equivocation $H(X|Y)$ (see Equation 4) quantifies the uncertainty due to the presence of selection mistakes only. It can be computed as the difference between stimulus entropy and mutual information.
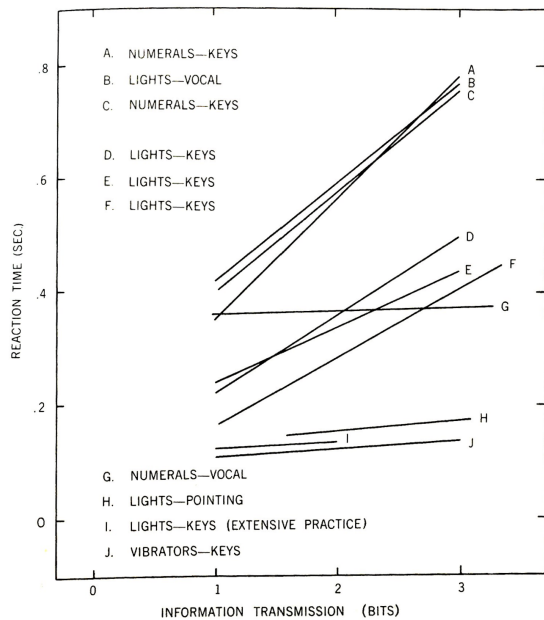
Figure 1. Reaction time as a function of stimulus uncertainty in different tasks. Constant reaction time: J [24], G [30], I [31]; Larger slope: F [20], B [4]. Taken from Fitts & Posner [16, p.105].

Indeed, as Fitts and Posner pointed out [16], anything that decreases the spatial or energy correspondence between input and output, therefore reducing compatibility, increases the slope. This principle is of obvious relevance to interface design; designers should strive to maximize S-R compatibility.

Effect of Learning
The effect of learning is very similar to that of S-R compatibility. When participants are heavily trained, the effects of the uncertainty of the stimulus and even of S-R compatibility can be reduced so that reaction time is almost constant, regardless of the number of items. Mowbray's experiment [31] showed that reaction time for choices among up to 10 possibilities could be reduced to that of a two choices alternative when a subject practiced a key-press task for a period of 6 months. Although this is somewhat questioned by Welford [48], it is clear that practice will significantly reduce the slope: Knight & Dagnall [22] reported slopes dropping from 73 ms/bit to 23 ms/bit after two months of practice. Seibel [42] reported that there is almost no difference in reaction time between responses to 31 (5 bits) or 1,023 (10 bits) stimuli after more than 75,000 trials.

Clarifying Hick's law
It is clear at this point that Hick's law is much more complicated than Card et al.'s description [6] would suggest. It is not just that learning, S-R compatibility and stimulus uncertainty all affect choice reaction time – these factors strongly interact with one another. For example, if one wishes to modify uncertainty by changing the probabilities of activation of each stimulus, then the subject has to go through an extensive learning phase, as discussed by Hyman [21]. Yet, she will inevitably improve her skill in the matter of the experiment, leading to a reduction of the slope.

Similarly, good S-R compatibility is usually desirable, otherwise one faces a poor design. However, this makes the effect of Hick's law harder to grasp, as the influence of stimulus uncertainty is then highly reduced.

Finally, learning does not affect all experiments in the same way. The highest rates are usually found with experiments using words. This is not necessarily because S-R compatibility is particularly good, but rather because reading and remembering words is a highly over-learned task, which we train daily. What can we say about Hick's light-key association task? Is this a completely new task, or are we somewhat familiar with it?

Instead of simply using Card et al.'s description [6] that Hick's law relates decision time to a set of choices, we propose the following, a more comprehensive version of Hick's law:

The choice reaction time for users performing a simple task grows linearly with the stimulus uncertainty, measured by entropy, in the range of 1 to 4 bits. The higher the S-R compatibility and the heavier the training, the shallower the slope. With appropriate learning, the effects of S-R compatibility and stimulus uncertainty can be reduced to almost zero.

THE CHOICE-REACTION PARADIGM AND HCI STUDIES
In this section, we revisit the HCI applications that have used Hick's law (see Table 1), comparing them with the choice-reaction paradigm in psychology and outlining the discrepancies in the use of the law. We argue that (a) Hick's law has often been misused in HCI and (b) it is in fact largely irrelevant for HCI tasks.
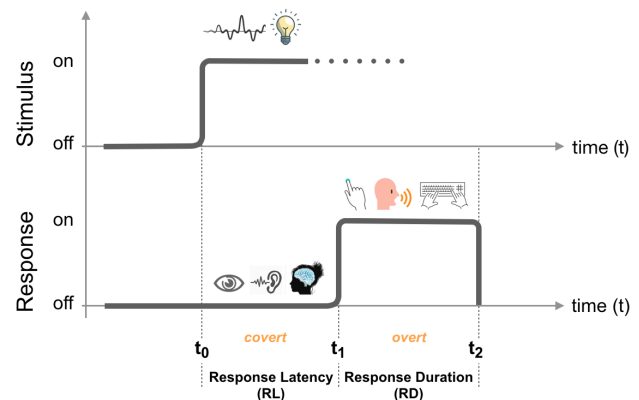

Figure 2. Time segmentation in the classic stimulus-response paradigm of psychology. t_0 **denotes stimulus onset, and** t_1 **and** t_2 **denote the response onset and termination, respectively. These three time markers make it possible to distinguish the response latency** (*RL*)**, the response duration** (*RD*)**, and the task completion time** ($TCT = RL + RD$)**. While most reaction-time experiments are designed so that** *RD* **occupies a negligible portion of** TCT**, in more realistic settings (see Table 1) more often than not it is the** *RL* **that is relatively short.**

	Hick [20]	Hyman [21]	Landauer & Nachbar [23]	Cockburn et al. [9]	Soukoreff & Mackenzie [49]	Mackenzie et al. [27, 28]	Wobbrock & Myers [50]	Ruiz et al. [40]
Task	Reaction	Reaction	VS	Decision	VS	VS	VS	Mode switch
Stimuli	Random	Random	Ordered	Random	Keyboard	Keyboard	Random	Random
Participants	Trained	Trained	All	Users from block 2	Novices	Novices	All	All
Distribution	Uniform	Non-uniform	Uniform	Zipfian	Uniform	Uniform	Uniform	Uniform
Stimulus Uncertainty	Up to 3.32 bits	Up to 2.81 bits	Up to 4 bits	Up to 3.58 bits	4.75 bits	4.75 bits	Up to 2 bits	Up to 3 bits
Measure	MI	Entropy	Entropy	Entropy	Entropy	Entropy	Entropy	Entropy
Formula	$\log(n+1)$ or $\log(n_e+1)$	$\log n$ or $-\sum_{i=1}^{n} p_i \log_2 p_i$	$\log n$	$\log n$	$\log n$	$\log n$	$\log n$	$\log n$
Results	Log	Log	Log	Log	Log	Not log	Log	Log

Table 1. A comparison of Hick 1952 [20], Hyman 1953 [21] and HCI studies that used Hick's law (MI: Mutual information; VS: Visual Search).

Decomposition of Time Measures

The first source of confusion is the notion of *reaction time*. As we have seen earlier, there are three types of reaction time in psychology and the Hick's paradigm falls into the second category: *choice-reaction time*, where response follows a certain stimulus-response mapping rule that requires learning. However, this notion has often been misunderstood and misused in HCI to interpret the time one takes to react to *anything*.

If we consider the stimulus-response paradigm from a measurement perspective, given a certain task, we typically face three time marks (Fig. 2): Stimulus-onset time t_0, Response-onset time t_1, and Response-termination time t_2, allowing the calculation of three relevant time durations: *RL* (Response Latency) $= t_1 - t_0$, *RD* (Response Duration) $= t_2 - t_1$, and the total task completion time $TCT = t_2 - t_0$. *RL* is the time it takes to complete a sequence of covert, unobservable mental processes, whereas *RD* is the time it takes to complete an overt response such as a hand movement or a vocal utterance. The previous section established that Hick along with other psychologists measured the choice reaction time tasks where *RD* was extremely short (time to push down a key) compared with *RL* (time to make the decision of which key to press down). In contrast, all the HCI studies in Table 1 assume a stimulus-response (S-R) paradigm where movement time MT (i.e. RD) contributes a relatively large portion of task completion time. Despite the fact that *RL* involves different mental processes in these studies, such as stimulus identification, visual search (e.g. [23, 49]), memory search, or decision to respond (e.g. [9]), the authors of these studies attributed these phenomena to Hick's law.

In fact, some of these phenomena can indeed be explained by a logarithmic function of the number of stimuli, but attributing them to Hick's law is a very unfortunate conflation between the formula and the law. For example, visual search in a hierarchical structure is logarithmic, as shown by Landauer and Nachbar [23]. Anything that involves a *divide and conquer* strategy can be logarithmic (e.g. Cockburn and Gutwin [8] show that scrolling time in an alphabetically organized menu is logarithmic). In many cases, this has nothing to do with Hick's law.

Effect Size of Hick's Law

The second source of confusion is the above-mentioned *stimulus-response compatibility*. As stated in the previous section, the slope in Hick's law depends on S-R compatibility. What is the S-R compatibility of the tasks we usually ask participants to perform? And how familiar are the participants in HCI experiments with the tasks we have them perform? More generally, what is the effect size of choice reaction time in HCI studies?

Arguably, the answer to the first question is that every HCI task has an extremely good S-R compatibility. We always ask participants to respond to the exact corresponding stimulus. For instance, selecting item 4 when item 4 is lit up in a menu or typing the letter a when it is time to type the letter a. We rarely ask participants to select item 10 when 4 is lit up or type the letter m when a is the stimulus. In a way, the S-R compatibility in psychology studies is never an issue in HCI. As we have seen from the literature, we can expect a very shallow slope of Hick's law.

To illustrate this point, we reanalyzed data from Roy et al. [39] and Liu et al. [26] where a simple command selection task was carried out. In Roy et al. [39], participants had to select a highlighted command by touching the screen of a tablet with a predetermined finger in the *Glass* condition vs. one of several fingers in the *Glass+Skin* condition. In Liu et al. [26], participants were instructed to move the mouse pointer to hit a highlighted stimulus among several commands. The finger or the pointer was rested on a designated area at the beginning of each trial. In both cases, reaction time was measured between highlight onset and the first move of the finger or pointer. Fig. 3 shows the plot of reaction time as a function of stimulus uncertainty. The slope of Hick's law is very small: 32ms/bit in the *Glass* condition, 8ms/bit in the *Glass+Skin* condition and 4ms/bit for command selection [26].

A recent study by Wright et al. [51] also demonstrates that when stimuli are spatial and linked to the responses in an intuitively simple layout, reaction time minimally depends on the number of alternatives. In a mode-switching study using the non-preferred hand, Ruiz et al. [40] report a logarithmic relation between the number of modes and reaction time but do
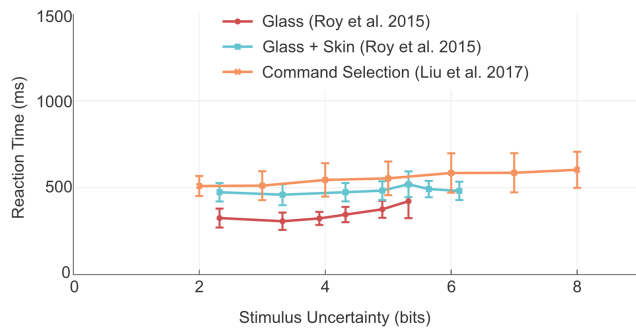
Figure 3. Reanalysis of data from Roy et al [39] and Liu et al [26]: reaction time as a function of stimulus uncertainty.

not find a significant effect of the number of modes on reaction time. This supports our claim that Hick's law is indeed not particularly relevant to HCI design. In practice, reaction time *RT* can usually be treated as a constant.

Differences between Novices and Experts

The third confusion has to do with the notion of *learning*. In HCI studies, an important issue is whether Hick's law applies to novice users, e.g. [28], expert users, e.g. [9], or both, e.g. [50]. Seow points out that Hick's law appears suited in predicting novice performance only [43]. As argued before, we believe that, once again, it depends on the task, S-R compatibility and practice.

In his experiments, Hick trained participants with more than 8000 practice trials whereas in Hyman's experiments, more then 15,000 trials were registered, yet there was still a slope in choice reaction time. In a way, the participants in Hick's and Hyman's experiments were trained enough to internalize the mapping between the lamp and the key (Hick) and between the lamp and the word (Hyman), but not trained enough to completely eliminate the effect of uncertainty on RT. Indeed, as shown in the previous section, later studies confirmed that an extremely well-trained participant can react in short and almost constant time regardless of stimulus uncertainty [31]. There are also tasks that we have become experts at, such as responding to Arabic numerals by reading them aloud [30], and tasks that require very little information processing, such as pressing the key that is felt to vibrate under one's fingers [24].

In HCI studies, the learning effect often comes from a new interaction technique and/or task, but does not, most of the time, come from the mapping between stimulus and response due to the extremely good S-R compatibility. Therefore, all users can be considered as experts in the choice-reaction paradigm.

Summary

Comparing the historical context of the choice-reaction paradigm and the use of Hick's law in HCI studies, we show that Hick's law is of little relevance to most HCI applications and choice-reaction time can be assumed to be constant. Many phenomena can lead to a logarithmic function, such as visual search in a hierarchical menu [23] or any *divide and conquer* strategy. Therefore observing a logarithmic relationship is not sufficient to attribute the underlying phenomenon to Hick's law.

HICK'S LAW AND DESIGN

In this section, we further examine the use of Hick's law, or rather of a logarithmic function, in the design community to justify the need to display as few items as possible. Unlike controlled experiments where the stimulus is explicit, in this scenario the stimulus is implicitly defined by the user's goal of selecting one item among a set. Attaining this goal might involve user behaviors different from those typically needed for a simple reaction, e.g. visual search for a single item, or finding matching pairs of items.

Based on a mathematical analysis, we first show that the "Hick-based" principle put forward in the design community is unduly justified by Hick's law. Then, we build on the previous discussions and suggest that, in spite of the complexity of the psychological processes behind response latency *RL* (Fig. 2), we can advocate a simple design strategy, on the premise that *RL* is either convex or concave. This leads us to a general design principle for Response Latency *RL* that goes beyond, and therefore includes, Hick's law and simple reaction.

The Hick-based Design Principles

In the design community, Hick's law is interpreted as a general design guideline, which we refer to as the "Hick-based design principle" in this article. In the book *Universal Principles of Design* [25], Lidwell et al. state that *"Designers can improve the efficiency of design by understanding the implications of Hick's law"* (p. 120). Similarly, in a Web entry titled *"Hick's law: Making the choice easier for users"*, Soegaard [46] writes that *"Understanding Hick's law means you can design so that more users will visit and stay on your website"*.

Wang [47] states that *"Essentially, Hick's law provides a general guideline for the design and use of hierarchical menu structures. This is consistent with the study [23] showing that users do not consider each choice one by one. What they normally do is to subdivide the choices into categories, and choices in each category are further divided. The resulted structure will be a tree, which can help users to make a quicker decision."*

Ali & Liem [1] claim that *"Within the context of design, Hick's law promotes the use of design methods to simplify decision-making in situations where designers are presented with multiple options. In practice, it has fundamentally proven to be effective in the design of menus, control display, way finding layout."*

Hick's law is also invoked in guidelines for designing applications for mobile devices [32], visualizations [17] and spreadsheets [7]. It seems that Hick's law is a magical formula in the design community and is widely used to rationalize two principles: (a) Minimize the number of choices; and (b) Categorize choices, instead of overwhelming users with all choices at once. In the next subsection, we use a simple car website scenario to see whether these Hick-based design principles hold.
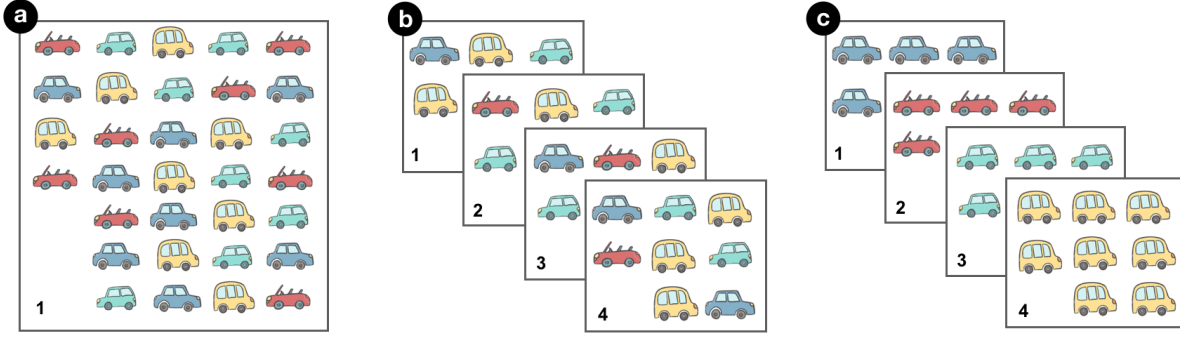
Figure 4. A designer (a) displays all 32 items on the same page; (b) splits 32 items randomly on 4 pages and (c) splits 32 items on 4 pages according to their colors.

The Concavity of the Logarithm Contradicts the Hick-based Design Principle: Car Website Scenario

We consider the following problem, where a designer has to display all $N = 32$ items on some website (Fig. 4). We compare the RL predicted by Hick's law for three different strategies of the designer

1. The designer displays the N items on the same page. Following a direct application of Hick's law, RL is given by

$$RL = a + b \log_2(32) = a + 5 \times b. \tag{6}$$

2. The designer splits the N items randomly on, say, 4 pages. Choosing one item from one page takes

$$rl = a + b \log_2(8) = a + 3 \times b. \tag{7}$$

This is the time to choose one item, or no item at all for one page. The person choosing the item will not always find the item in the first page; on average he has to go through 2.5 pages before finding the right item. Hence RL is given by

$$RL = 2.5 \times rl = 2.5 \times a + 7.5 \times b. \tag{8}$$

3. The designer splits the N items on, say, 4 pages, according to 4 exclusive categories, i.e. if an item belongs to one category, it cannot belong to another. In that case, the person choosing the items will first choose a category, then an item within that category. Direct application of Hick's law gives

$$RL = a + b \log_2(4) + a + b \log_2(8) = 2 \times a + 5 \times b. \tag{9}$$

As a result, the optimal strategy according to Hick's law consists of displaying all the items at once on the same page, contrary to the design principle that choices should be categorized.

The following general result holds. When there are N items to be displayed that can be separated into k subgroups, applying Hick's law leads to the following:

- It is never advantageous to split elements into uncategorized subgroups of equal sizes. The average RL for selecting an item from N items, equally categorized into k subgroups is

given by

$$RL = \frac{k+1}{2} \left[a + b \log(N/k) \right], \tag{10}$$

which generalizes Eq. (8). Since $\frac{k+1}{2}$ is obviously larger than 1, the question thus becomes whether the inequality

$$\frac{k+1}{2} \log(N/k) \geq \log(N) \tag{11}$$

is true or not. It can be shown[5] that for $N \geq k^{\frac{k+1}{k-1}}$, Ineq. (11) is true.

- It is not advantageous, or at best useless, to split elements into categorized subgroups of equal size. Indeed, since $N = k \times \frac{N}{k}$, $\log N = \log k + \log N/k$, so that

$$RT = a + b \log(k) + a + b \log(N/k) \tag{12}$$
$$= a + (a + b \log N). \tag{13}$$

This generalizes Equation (9). Assuming there are m nested categories, we must ultimately pay the price of an additional $(m-1)a$ seconds on RL.

The conclusion of this small example is that Hick's law–which, for the argument, we assume here to apply to the design problem under consideration–actually suggests displaying as many items as possible, which is contrary to common sense and our general experience. Hence, the design principle cannot be justified by Hick's law, nor by any other logarithmic RL function.

There are many different phenomena taking place when a user is investigating the items, which are far more complex than simple stimulus response. For instance, if one were to examine visually one item after the other before initiating a response, RL would increase linearly with N [48]. If one were to compare a pair of items among a set of size N, time would likely increase quadratically, since there are $N(N-1)/2$ possible pairs.

---

[5]This is valid for all practical cases, since as k grows large, this condition becomes $N \geq k$. With our previous example, $k = 4$ and we get $N \geq 10.08$.

RL Design Principle: A Matter of Convexity

We previously suggested several non logarithmic *RL* functions. Despite the complexity of the psychological process behind *RL* and the many potential *RL* functions, a simple design strategy can be advocated, on the premise that *RL* is either convex or concave.

As in the previous example, we consider two different situations: one in which items can be categorized, the other in which they cannot. This leads to two different convexity results. Let f be the function that relates *RL* to N

$$\text{RL} = f(N), \tag{14}$$

Case 1: Items cannot be categorized
If the items are displayed in two sets of sizes x and y ($x + y = N$), then $\text{RL} = f(x) + f(y)$, whereas if the items are displayed all at once, $\text{RL} = f(x+y)$. Therefore, determining whether or not we should split the items boils down to whether $f(x+y)$ is greater or lower than $f(x) + f(y)$

$$f(x+y) \geq f(x) + f(y), \tag{15}$$

If Equation (15) holds, then f is said to be superadditive; else f is subadditive. It can be shown [5] that if $f(0) = 0$:

$$f \text{ convex implies } f \text{ superadditive}; \tag{16}$$
$$f \text{ concave implies } f \text{ subadditive}. \tag{17}$$

An important information is thus whether f is a convex or concave function. This leads to our first *RL* design principle:

> *If items are not categorized and f is convex, e.g. quadratic, then grouping items will reduce RL. If f is concave, e.g. logarithmic, then it is better to display all items at once[6].*

Case 2: Items are categorized
For this case, we assume for simplicity that the items are categorized in x categories with y items in each. Therefore $N = xy$. From the example of the previous subsection, it is clear that the relevant question is whether $f(x) + f(y)$ is greater or not than $f(xy)$. It is easily shown through the previous result that

$$f(\exp(.)) \text{ convex implies } f(x) + f(y) \leq f(xy); \tag{18}$$
$$f(\exp(.)) \text{ concave implies } f(x) + f(y) \geq f(xy). \tag{19}$$

In Hick's paradigm reaction time grows logarithmically, in which case $f(\exp(x))$ is linear; this is the limit case where $f(x) + f(y) = f(xy)$, i.e. the two strategies are equivalent. As all functions relating selection time to the number of items realistically grow faster than the logarithm (any practical HCI setup is likely more time consuming than Hick's task), we can state the second *RL* design principle as follows:

> *If items are categorized, then they should always be split.*

Fig. 5 summarizes the two *RL* design principles.

---

[6] Note that we use the convention that a choice from a null set takes 0 seconds, which explains the extra conditions needed in (10) as $\log(0) \to -\infty$. Also note that using the formulation of (5) conveniently solves this issue by adding 1 to N.
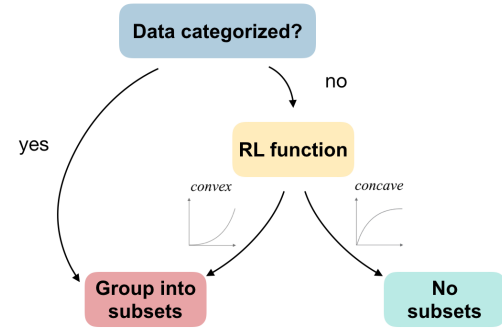

Figure 5. Design guideline based on data categorization and *RL* function.

CONCLUSION

We have shown that Hick's law is not always relevant to the tasks studied in HCI. Psychologists have successfully investigated the limits and conditions of application of the law, but this knowledge is rarely applied in HCI.

To summarize, Hick's law is often misused in HCI and the psychology of choice-reaction time is of little relevance to most HCI applications because high S-R compatibility and over-learned tasks result in very short and almost constant reaction times. However, Hick's law can inform good design by encouraging high S-R compatibility and learnability. Rather than wondering whether or not Hick's law applies, we therefore encourage to better observe and measure these characteristics.

We show that logarithmic growth of observed temporal data is not necessarily interpretable in terms of Hick's law. Many mechanisms can lead to a logarithm function, such as visual search in a hierarchical menu and scrolling in an alphabetically organized menu. These cases have nothing to do with the choice-reaction paradigm. Rather than the current dichotomous interrogation of whether Hick's law is the appropriate model to summarize a given dataset or not, we advocate for further work to better understand the actual mechanisms driving interaction.

Using a simple car scenario, we also demonstrate that Hick's law, or rather a logarithmic function, cannot justify the "Hick-based" design principle for organizing a set of choices. For user interface design, a detailed examination of the cognitive processes involved in response latency, such as visual search and decision making, is more fruitful than a mere reference to Hick's law.

Looking forward, creating an empirical taxonomy for *RL* analogous to the theoretical notion of computational complexity might be useful for HCI researchers who strive to model human behavior. Since sub-additivity is the main issue, an index that quantifies this property would probably be useful. $S = \frac{f(x)+f(y)}{f(x+y)}$ is a natural candidate. More empirical and theoretical work should determine whether S is indeed useful.

REFERENCES

1. Abu Ali and Andre Liem. 2014. The use of formal aesthetic principles as a tool for design conceptualisation and detailing. *DS 81: Proceedings of NordDesign 2014, Espoo, Finland 27-29th August 2014* (2014), 490–499.

2. John R Anderson. 1996. ACT: A simple theory of complex cognition. *American psychologist* 51, 4 (1996). DOI:http://dx.doi.org/10.1037/0003-066X.51.4.355

3. G Blank. 1934. Brauchbarkeit optischer Reaktionsmessungen (Usability of optical reaction measurements). *Indust. Psychotech.* 11 (1934), 140–150.

4. Robert W Brainard, Thomas S Irby, Paul M Fitts, and Earl A Alluisi. 1962. Some variables influencing the rate of gain of information. *Journal of experimental psychology* 63, 2 (1962), 105. DOI:http://dx.doi.org/10.1037/h0045304

5. Andrew M Bruckner, E Ostrow, and others. 1962. Some function classes related to the class of convex functions. *Pacific J. Math.* 12, 4 (1962), 1203–1215. https://projecteuclid.org/euclid.pjm/1103036122

6. Stuart K. Card, Allen Newell, and Thomas P. Moran. 1983. *The Psychology of Human-Computer Interaction*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA.

7. Hock Chuan Chan, Suparna Goswami, and Hee-Woong Kim. 2012. An alternative fit through problem representation in cognitive fit theory. *Journal of Database Management (JDM)* 23, 2 (2012), 22–43. DOI:http://dx.doi.org/10.4018/jdm.2012040102

8. Andy Cockburn and Carl Gutwin. 2009. A predictive model of human performance with scrolling and hierarchical lists. *Human–Computer Interaction* 24, 3 (2009), 273–314. DOI:http://dx.doi.org/10.1080/07370020902990402

9. 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 (CHI '07)*. ACM, New York, NY, USA, 627–636. DOI:http://dx.doi.org/10.1145/1240624.1240723

10. ERFW Crossman. 1953. Entropy and choice time: The effect of frequency unbalance on choice-response. *Quarterly Journal of Experimental Psychology* 5, 2 (1953), 41–51. DOI:http://dx.doi.org/10.1080/17470215308416625

11. ERFW Crossman. 1955. The measurement of discriminability. *Quarterly Journal of Experimental Psychology* 7, 4 (1955), 176–195. DOI:http://dx.doi.org/10.1080/17470215508416692

12. Franciscus Cornelis Donders. 1868. Die schnelligkeit psychischer processe (On the speed of mental processes). (1868).

13. 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. DOI:http://dx.doi.org/10.1037/h0055392

14. Paul M Fitts. 1966. Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of experimental psychology* 71, 6 (1966), 849. DOI:http://dx.doi.org/10.1037/h0023232

15. Paul M Fitts, James R Peterson, and Gerson Wolpe. 1963. Cognitive aspects of information processing: II. Adjustments to stimulus redundancy. *Journal of Experimental Psychology* 65, 5 (1963), 423. DOI:http://dx.doi.org/10.1037/h0047993

16. Paul M Fitts and Michael I Posner. 1967. Human performance. (1967).

17. Peter Hall, Claude Heath, and Lizzie Coles-Kemp. 2015. Critical visualization: a case for rethinking how we visualize risk and security. *Journal of cybersecurity* 1, 1 (2015), 93–108. DOI:http://dx.doi.org/10.1093/cybsec/tyv004

18. Hermann Helmholtz. 1850. Ueber die Fortpflanzungsgeschwindigkeit der Nervenreizung. *Annalen der Physik* 155, 2 (1850), 329–330.

19. WE Hick. 1951. A simple stimulus generator. *Quarterly Journal of Experimental Psychology* 3, 2 (1951), 94–95. DOI:http://dx.doi.org/10.1080/17470215108416779

20. William E Hick. 1952. On the rate of gain of information. *Quarterly Journal of Experimental Psychology* 4, 1 (1952), 11–26. DOI:http://dx.doi.org/10.1080/17470215208416600

21. Ray Hyman. 1953. Stimulus information as a determinant of reaction time. *Journal of experimental psychology* 45, 3 (1953), 188. DOI:http://dx.doi.org/10.1037/h0056940

22. AA Knight and PR Dagnall. 1967. Precision in movements. *Ergonomics* 10, 3 (1967), 321–330. DOI:http://dx.doi.org/10.1080/00140136708930874

23. T. K. Landauer and D. W. Nachbar. 1985. Selection from Alphabetic and Numeric Menu Trees Using a Touch Screen: Breadth, Depth, and Width. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '85)*. ACM, New York, NY, USA, 73–78. DOI:http://dx.doi.org/10.1145/317456.317470

24. J Alfred Leonard. 1959. Tactual choice reactions: I. *Quarterly Journal of Experimental Psychology* 11, 2 (1959), 76–83. DOI:http://dx.doi.org/10.1080/17470215908416294

25. William Lidwell, Kritina Holden, and Jill Butler. 2010. *Universal principles of design, revised and updated: 125 ways to enhance usability, influence perception, increase appeal, make better design decisions, and teach through design*. Rockport Pub.

26. Wanyu Liu, Olivier Rioul, Michel Beaudouin-Lafon, and Yves Guiard. 2017. Information-Theoretic Analysis of Human Performance for Command Selection. In *IFIP Conference on Human-Computer Interaction*. Springer, 515–524. DOI:http://dx.doi.org/10.1007/978-3-319-67687-6_35

27. I Scott MacKenzie and Shawn X Zhang. 2001. An empirical investigation of the novice experience with soft keyboards. *Behaviour & Information Technology* 20, 6 (2001), 411–418. DOI: http://dx.doi.org/10.1080/01449290110089561

28. I Scott MacKenzie, Shawn X Zhang, and R William Soukoreff. 1999. Text entry using soft keyboards. *Behaviour & information technology* 18, 4 (1999), 235–244. DOI: http://dx.doi.org/10.1080/014492999118995

29. Julius Merkel. 1883. *Die zeitlichen verhältnisse der willensthätigkeit (The Temporal Relations of the Actions of Will, or The Timing of Voluntary Action)*. W. Engelmann.

30. GH Mowbray. 1960. Choice reaction times for skilled responses. *Quarterly Journal of Experimental Psychology* 12, 4 (1960), 193–202. DOI: http://dx.doi.org/10.1080/17470216008416726

31. GH Mowbray and MV Rhoades. 1959. On the reduction of choice reaction times with practice. *Quarterly Journal of Experimental Psychology* 11, 1 (1959), 16–23. DOI: http://dx.doi.org/10.1080/17470215908416282

32. Fatih Nayebi, Jean-Marc Desharnais, and Alain Abran. 2013. An expert-based framework for evaluating ios application usability. In *Software Measurement and the 2013 Eighth International Conference on Software Process and Product Measurement (IWSM-MENSURA), 2013 Joint Conference of the 23rd International Workshop on*. IEEE, 147–155. DOI: http://dx.doi.org/10.1109/IWSM-Mensura.2013.30

33. Allen Newell and Stuart K Card. 1985. The prospects for psychological science in human-computer interaction. *Human-computer interaction* 1, 3 (1985), 209–242. DOI: http://dx.doi.org/10.1207/s15327051hci0103_1

34. Anton Nikolov. 2017. Design principle: Hick's Law - quick decision making. (2017). https://uxplanet.org/design-principles-hicks-law -quick-decision-making-3dcc1b1a0632.

35. Irwin Pollack and JM Pickett. 1963. The intelligibility of excerpts from conversation. *Language and Speech* 6, 3 (1963), 165–171. DOI: http://dx.doi.org/10.1177/002383096300600305

36. Robert W Proctor and Darryl W Schneider. 2018. Hick's law for choice reaction time: A review. *Quarterly Journal of Experimental Psychology* 71, 6 (2018), 1281–1299. DOI:http://dx.doi.org/10.1080/17470218.2017.1322622

37. Roger Ratcliff and Philip L Smith. 2004. A comparison of sequential sampling models for two-choice reaction time. *Psychological review* 111, 2 (2004), 333. DOI: http://dx.doi.org/10.1037/0033-295X.111.2.333

38. Yvonne Rogers, Helen Sharp, and Jenny Preece. 2015. *Interaction Design, 4th edition*. Wiley.

39. Quentin Roy, Yves Guiard, Gilles Bailly, Éric Lecolinet, and Olivier Rioul. 2015. Glass+ skin: An empirical evaluation of the added value of finger identification to basic single-touch interaction on touch screens. In *Human-Computer Interaction*. Springer, 55–71. DOI: http://dx.doi.org/10.1007/978-3-319-22723-8_5

40. Jaime Ruiz, Andrea Bunt, and Edward Lank. 2008. A Model of Non-preferred Hand Mode Switching. In *Proceedings of Graphics Interface 2008 (GI '08)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 49–56. http://dl.acm.org/citation.cfm?id=1375714.1375724

41. Jeff Sauro. 2013. Five HCI Laws For User Experience Design. (2013). https://measuringu.com/hci-laws/

42. Robert Seibel. 1963. Discrimination reaction time for a 1,023-alternative task. *Journal of Experimental Psychology* 66, 3 (1963), 215. DOI: http://dx.doi.org/10.1037/h0048914

43. Steven C Seow. 2005. Information theoretic models of HCI: a comparison of the Hick-Hyman law and Fitts' law. *Human-Computer Interaction* 20, 3 (2005), 315–352. DOI:http://dx.doi.org/10.1207/s15327051hci2003_3

44. Claude E Shannon. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27 (1948), 379–423, 623–656. DOI: http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x

45. Ben Shneiderman, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholas Diakopoulos. 2017. *Designing the User Interface, 6th Edition*. Pearson.

46. Mads Soegaard. 2016. Hick's Law: Making the choice easier for users. (2016). https://www.interaction-design.org/literature/article/ hick-s-law-making-the-choice-easier-for-users

47. Harris Wang. 2014. A guide to assistive technology for teachers in special education. In *Assistive technologies: Concepts, methodologies, tools, and applications*. IGI Global, 12–25. DOI: http://dx.doi.org/10.4018/978-1-4666-4422-9.ch002

48. Alan Traviss Welford. 1968. Fundamentals of skill. (1968).

49. R William Soukoreff and I Scott Mackenzie. 1995. Theoretical upper and lower bounds on typing speed using a stylus and a soft keyboard. *Behaviour & Information Technology* 14, 6 (1995), 370–379. DOI: http://dx.doi.org/10.1080/01449299508914656

50. Jacob O. Wobbrock and Brad A. Myers. 2006. From Letters to Words: Efficient Stroke-based Word Completion for Trackball Text Entry. In *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '06)*. ACM, New York, NY, USA, 2–9. DOI: http://dx.doi.org/10.1145/1168987.1168990

51. Charles E Wright, Valerie F Marino, Charles Chubb, and Daniel Mann. 2019. A model of the uncertainty effects in choice reaction time that includes a major contribution from effector selection. *Psychological review* 126, 4 (2019), 550. DOI:http://dx.doi.org/10.1037/rev0000146