Original Article

QJEP

Quarterly Journal of Experimental Psychology
2018, Vol. 71(6) 1281–1299

DOI: 10.1080/17470218.2017.1322622
qjep.sagepub.com
SAGE

Hick's law for choice reaction time: A review

Robert W Proctor and Darryl W Schneider

Abstract

In 1952, W. E. Hick published an article in the *Quarterly Journal of Experimental Psychology*, "On the rate of gain of information." It played a seminal role in the cognitive revolution and established one of the few widely acknowledged laws in psychology, relating choice reaction time to the number of stimulus–response alternatives (or amount of uncertainty) in a task. We review the historical context in which Hick conducted his study and describe his experiments and theoretical analyses. We discuss the article's immediate impact on researchers, as well as challenges to and shortcomings of Hick's law and his analysis, including effects of stimulus–response compatibility, practice, very large set sizes and sequential dependencies. Contemporary modeling developments are also described in detail. Perhaps most impressive about Hick's law is that it continues to spawn research efforts to the present and that it is regarded as a fundamental law of interface design for human–computer interaction using technologies that did not exist at the time of Hick's research.

Keywords

Hick's law; choice reaction time; set-size effects; information theory; cognitive models

Date received: 21 January 2017; revised: 2 April 2017; accepted: 3 April 2017

Introduction

In the early 1950s, rapid advances in communication technology and the development of digital computers were being made. Concepts about information and its transmission emerging from electronic communications theory (Shannon & Weaver, 1949) and cybernetics (Wiener, 1948) seemed applicable to theorizing and research in psychology. Although information theory was developed for communication from person to person by, for example, the telephone, the human mind could be viewed as a communication system that processed information from input (sensory and perceptual processes) to output (overt responses). Although the advent of the "cognitive revolution" in psychology is sometimes dated to 1956 (Newell & Simon, 1972), Miller (2003) noted that 1956 "was only slightly better than the years just preceding and following. Many were riding the waves that began during World War II: those of servo theory, information theory, signal-detection theory, computer theory and computers themselves" (p. 142).

Research in the area of attention and performance was already experiencing the revolution and indeed was at the forefront as it occurred (Posner, 1986). Somewhat ironically, much of this research that was to revolutionize basic experimental psychology was conducted under the auspices of applied psychology, relating mainly to military applications. In England, the research emanated from the Medical Research Council's Applied Psychology Unit, whereas in the United States it originated from the Psychology Branch of the Aero Medical Laboratory. Although it is impossible to single out any particular study as the primary source of the cognitive revolution, there can be no doubt that W. E. Hick's (1952) study was an influential landmark. In his article, Hick applied the concepts of information theory to performance of choice reaction tasks, providing a striking demonstration that information and its effect on human performance could be quantified.

Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

Corresponding authors:
Robert W Proctor, Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907, USA.
Email: proctor@psych.purdue.edu

Darryl W Schneider, Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907, USA.
Email: dws@purdue.edu

Hick's (1952) article, "On the rate of gain of information," published in the *Quarterly Journal of Experimental Psychology* (QJEP), was the first of three articles demonstrating the value of information theory for the study of human performance. The other articles were by Hyman (1953), also on choice reaction time (RT), and Fitts (1954), on the speed–accuracy trade-off in executing aimed movements (Fitts' law). Hick's article was highly influential at the time and continues to be to the present. In 1974, Teichner and Krebs stated in their review of visual choice reaction research, "Much of the current theoretical approach to choice reaction time is based upon the early efforts of Hick" (p. 85). In an editorial on classic articles published in QJEP, Brysbaert (2016) noted that Hick's article was the third most cited in the journal's history. The citation count in the Web of Science on 1 April 2017 was 1163, and impressively, the article was cited from 32 to 48 times annually from 2006 to 2016. Indeed, its number of citations in any one year during that period is greater than the highest number of citations in any prior year. This increasingly high citation rate documents that Hick's article is a classic in the study of human performance that continues to be relevant to contemporary psychology. Moreover, although established in basic research, Hick's law is often listed as one of a few laws applicable to human–computer interaction (HCI; e.g., Cao, 2010; Sauro, 2013), reflecting its grounding in applied display–control interface problems.

In this article, we begin by providing historical context for Hick's groundbreaking research. We follow that with a detailed description of his article. We describe contemporary developments relating to information theory and performance and conduct a selective review of more recent work, including contemporary models that address the influence of number of alternatives and uncertainty on choice RT. Along the way, we spell out the many significant contributions of Hick's study as well as some of the issues that it raised.

Historical context

Hick was born on 1 August 1912 and lived until 20 December 1974. He joined the Medical Research Council's Applied Psychology Unit at the Cambridge Psychological Laboratory in 1944 (Psychreg.org, 2016), when it was founded, and from 1948 to 1956, he published 10 journal articles (including a book review), 8 of which appeared in QJEP. His 1952 article establishing "Hick's law" (also sometimes called the Hick–Hyman law) is far and away the most widely cited of his works, as the *h* index shown for Hick in the Web of Science is 8.

Hick worked at the Applied Psychology Unit during a fortuitous time in which studies of human perception and performance were the subject of much research. The founding director, Kenneth Craik, whose 1943 book *The Nature of Explanation* provided an early argument for the concept of mental models, focused much of the Unit's research on display and control problems in human–machine interactions. The 1946 *Progress Report* of the Unit (MRC Applied Psychology Unit, 1946) indicates that Hick assisted him on this research until Craik's untimely death in 1945 and states,

> Dr. Hick has continued the study particularly of "control" problems. He has demonstrated the best kind of equipment to use with "jolting" conditions, has studied several problems of rifle aiming, has investigated a number of different tracking systems and has shown the effects of a varying resistance when steady winding is required. This work is actively continuing. (p. 5)

Already at this time, Hick had focused on topics of control rather than perception of displays.

The first issue of QJEP includes five articles, one of which is Hick's (1948) article, "The discontinuous functioning of the human operator in pursuit tasks." This article foreshadowed his 1952 article in several respects. In it, Hick dealt with the topic of continuity versus discontinuity in performance of tracking tasks, noting,

> . . . the continuity or discontinuity with which we are chiefly concerned is that to be found, not in the relation between output and time, but in the relation between output and input between response and stimulus, in the human operator. (p. 38)

So, his concentration was already on stimulus–response (S-R) relations. Moreover, Hick identified two types of discontinuity as important in tracking tasks, the threshold for detecting a change in the stimulus and the refractory phase, to which he devoted the majority of his article. Again, his emphasis was primarily on the cognitive control aspect of discontinuity.

Hick (1948) already incorporated the term *information*, without formal definition, in his characterization of the refractory phase when making separate responses to two successive stimuli:

> . . . the stimulus is really the signal for the taking of a decision—a decision to carry out a pre-arranged plan. Such a decision implies a certain determination to ignore subsequent information, until the planned action has been brought to completion. (p. 41)

Furthermore, he mentioned unpublished experiments by himself and published experiments by Bates (1947) using early versions of the stop-signal task (Verbruggen & Logan, 2009) in which ". . . the second 'response' was merely to arrest the first" (p. 41), concluding that the estimated period during which the response could not be stopped was similar to the duration of the refractory phase when an overt response was required to the second stimulus. Hick went on to say,

> . . . one is tempted to suspect a single process underlying both the choice-reaction delay [relative to simple RT] and the

> refractory phase. And, as a pure speculation, it may be suggested that the same process subserves what we call attention. (p. 42)

Thus, again foreshadowing his 1952 article, Hick saw multiple phenomena as having a common basis in cognitive processes. On the whole, the 1948 article provides evidence that Hick was "primed" to incorporate information theory into his work and to consider its implications for performance in choice reaction tasks.

Shannon and Weaver's book, *The Mathematical Theory of Communication*, in which information theory was developed, was published in 1949, and Hick was quick to recognize its relevance to psychology. His initial published foray into information theory was an article examining its implications for intelligence testing (Hick, 1951). The 1949-1950 Progress Report (MRC Applied Psychology Unit, 1950) includes mention of information theory in a single paragraph:

> At a recent symposium held by the Royal Society on Information Theory it has been emphasised that the concepts of this subject can develop specific theories of real predictive capacity (Hick). A comprehensive account of researches in this field has recently been published (Hick and Bates). Initial experiments have also been done on the application of information theory to problems of skill, especially those associated with the perceptual aspects of the task, such as arc involved in the selection of a response to a complex visual stimulus. From the time required to respond to complex visual material it seems that the process may be a chain of operations of classification, all basically similar and so arranged as to effect the greatest economy of information (Hick).

There are two things to note about this paragraph on information theory. First, the parenthetical references to research are all to Hick, indicating that he was the leading advocate of information theory at the Applied Psychology Unit. Second, the final two sentences of the paragraph relate to the experiments published in Hick's (1952) article, to which we now turn.

Hick's (1952) article

Hick (1952) reported three experiments. He noted a finding of Merkel (1885; described in Woodworth, 1938), for which RT was an increasing function of the number of S-R alternatives in a choice task, and proposed a hypothesis "that the rate of gain of information is, on the average, constant with respect to time" (p. 12). He remarked that this hypothesis "would have been impossible in the early days of reaction-time work, because the theoretical framework did not then exist" (p. 12). The framework to which Hick referred was that of information theory, which, as noted, had been recently developed by Shannon and Weaver (1949). This theory represents information (or entropy) in terms of uncertainty, which, for a choice reaction task, increases as a function of the number of equally likely alternatives. Because information is a logarithmic function of the number of alternatives, Hick noted that the hypothesis of a constant rate of gain of information predicts that RT should be a logarithmic function of the number of alternatives. As a first step in his empirical investigation, Hick confirmed that Merkel's data were in fact fit well by a logarithmic function.

Hick then tested himself as the sole subject in his Experiment 1 on an apparatus in which 10 lamps were arranged in a small irregular circle (to minimize eye movements and obvious groupings) and the 10 fingers rested on response keys (for a photograph of Hick and his apparatus, see Figure 4 in Reynolds, 2004). In different trial blocks, Hick performed reaction tasks with no response uncertainty (a go/no-go task) and with 2 through 10 alternative possible stimuli and responses. For all choice tasks, the analyzed data were restricted to the responses of the little and ring fingers—the two fingers used to respond in the two-choice task—to control for possible differences in RTs across fingers. His RT function was again fit well with a logarithmic function. Hick fit both Merkel's (1885) data and his own using the function $RT = b \log_{10}(n+1)$, where n is the number of alternatives, 1 is added to accommodate "no stimulus" and b is the slope of the function relating RT to information. The more widely accepted version of Hick's law for n equally likely alternatives is $RT = a + b \log_2 n$, in which a is the intercept and b the slope of the function relating RT to information. Log_2 has come to be used because of its link to binary digits (bits) of information (see later), and the equation for Hick's law is sometimes expressed in terms of the average information or uncertainty (H) in a series of trials, in the form $RT = a + bH$.

Because information transmitted is less than the logarithm of n (or $n+1$) if incorrect responses are made, in Hick's Experiment 2, he and another subject performed the task with the intent of making different proportions of errors in each run of 100 trials. Noted psychologist Richard L. Gregory later identified himself as that other subject, saying,

> I was actually the only subject. Hick's Law relating RT to the number of choices is based on my nervous system. He was the only other subject and he got bored doing it, so he never finished the experiment, which went on for months. (Reynolds & Tansey, 2001, p. 6)

Hick was able to trade accuracy for speed relatively well across a range of different error proportions. Gregory had more difficulty doing that, but both his and Hick's data were again fit relatively well by a logarithmic function. Because two 10-choice sequences were used repeatedly in Experiment 2, Hick evaluated whether learning of the sequence contributed to performance in his Experiment 3.

This evaluation was accomplished by examining performance of the 10-choice task with a new sequence, which yielded results similar to those of the 10-choice task in Experiment 2 for both subjects. From this, Hick concluded, ". . . learning of the sequences did not play a large part in determining previous findings" (p. 17). He also considered differences between particular S-R pairings, concluding, ". . . there is almost certainly some relation between the reaction time of a particular response or to a particular stimulus and the corresponding uncertainty" (p. 19). But he acknowledged that "How the relation comes about cannot be inferred from the present data" (p. 19).

More important, Hick considered several conceptual models. He discussed various template models in which recognition or identification occurred by matching a presented stimulus to templates, a process that could unfold in different ways. One type of model involved replication, by which m replicas of the stimulus (presumably, with $m=n$) were produced, and each replica was compared with the templates to find a match. Hick considered three modes of replication. First, replication could be simultaneous, such that multiple replicas could be produced in the same time as one. Of relevance to our later coverage of models of Hick's law, it is notable that Hick argued that simultaneous replication "leads to a conception of reaction time as a period of continuous accumulation of evidence" (p. 21). Second, replication could be serial, in which case replication time would be proportional to m. Third, self-replication might occur, such that each replica creates a replica of itself in a geometric progression, resulting in replication time being proportional to the logarithm of m. Another type of model considered by Hick involved searching rather than replicating, by which a search among templates is conducted. Hick distinguished between random and systematic search strategies, noting that both approaches would result in search times being linear with m. Among these possibilities, Hick deemed only the self-replication model to be promising.

Hick proposed a final model based on a hierarchical sequence of binary classification decisions. He noted that if logarithms to the base 2 are used to represent information, the resulting value can be taken as the number of dichotomizing decisions to arrive at the correct response. With eight equally likely alternatives, for example, three binary decisions would be required: The first would partition the possibilities into the correct group of four, the second into the correct group of two and the final decision would identify the correct response to be made. Choice among four equally probable alternatives would require two binary decisions, and choice among two alternatives would need only one. Hick noted that this progressive classification model would yield a logarithmic relationship between RT and the number of alternatives, and he thought it seemed more appropriate than the self-replication template model.

Hick's study was groundbreaking in several regards. He made a convincing case that RT measures could yield reliable relationships about human performance. Hick also showed that evaluation of quantitative models was possible, and that alternative possible information-processing models needed to be considered in attempting to explain the information processing underlying lawful empirical relations. His Experiment 2 demonstrated not only that people could trade off speed and accuracy but that the resulting RTs could be conceived of in terms of the amount of time over which stimulus information was allowed to accumulate. Together with Fitts' law for movement time (Fitts, 1954), which was also established in the first half of the 1950s, Hick's law provided a firm foundation on which cognitive psychology could build.

Replication of the logarithmic function and empirical restrictions

Although RT has been used as a dependent measure in psychological research since the earliest days of the field (Donders, 1969/1868; Jastrow, 1890), Hick's (1952) article led a resurgence in use of the measure. Hyman (1953) published a closely related article the next year in which he described experiments that showed similar results. This similar contribution is why Hick's law is also referred to as the Hick-Hyman law. Hyman's method differed from Hick's in using only up to eight stimulus locations to which four subjects were to respond vocally by saying the assigned name from the pseudoword set Bun, Boo, Bee, Bore, By, Bix, Bev and Bate. As Hick did initially, Hyman varied the number of equally likely alternatives in his Experiment 1, but in Experiments 2 and 3, he altered the amount of information conveyed by a stimulus through, respectively, varying the probability of occurrence of particular S-R pairs and through instituting sequential dependencies between successive trials. Hyman showed that when the data from these three experiments were plotted as a function of bits of information transmitted from stimuli to responses, each subject's results could be fit with a single function of the form $RT=a+bH$. In a study with lesser impact, Crossman (1953) showed that the entire time to perform a card sorting task also conformed to a logarithmic function.

Despite this promising start, there were cracks in the information-theoretic approach from the beginning. As noted, Hick (1952) considered several processing models that could underlie the logarithmic relation between information and RT, arriving at the conclusion that one based on a series of dichotomizing decisions was the best. However, he ultimately concluded that "the simplest scheme of operations which fits the general proposition [the progressive classification model] has been found to lead to hypotheses which other aspects of the data largely fail to confirm, although they do not definitely contradict

it" (p. 25). Leonard (1958) provided additional evidence against the progressive classification model using a task in which he precued three of six alternatives at different intervals in advance of the target stimulus to which the subject was to respond. The rationale behind the precuing approach was that the precue would externalize part of the progressive classification process that Hick proposed subjects did internally. The main finding was that the interval required for RT in the precued condition to approach that of a three-choice task was much longer than the difference between the three- and six-choice tasks. This finding implied that subjects were not making a series of binary decisions.

Additional studies revealed that the rate of gain of information varied across different S-R sets, especially as a function of S-R compatibility. Moreover, extensive practice was shown to affect the slope of the set-size function. These empirical restrictions on Hick's law are discussed in the remainder of this section.

S-R compatibility

S-R compatibility is a term that refers to differences in RT and error rate as a function of the pairings of entire stimulus and response sets, or mappings of the individual members of a stimulus set to those of a response set (Proctor & Vu, 2006). Fitts and Seeger (1953) published a classic study on S-R compatibility in which subjects performed eight-choice spatial reaction tasks with all combinations of three different stimulus arrays and three different response configurations. The tasks required subjects to move a stylus (or for one response configuration, two styli) from a center position to a location corresponding to the stimulus. Although all tasks were eight-choice and thus equivalent with respect to the amount of uncertainty, both mean RTs and error rates varied dramatically as a function of the compatibility, or correspondence, of the overall stimulus and response configurations. Dassonville, Lewis, Foster, and Ashe (1999) showed explicitly that for two-, four- and eight-choice tasks requiring movement with a joystick of a cursor to a location on the screen, the slope relating RT to $\log_2 n$ was an increasing function of the incompatibility of the mapping of stimulus locations to response locations.

Not only does the slope of the Hick's law RT function vary with S-R compatibility, it is possible for set size to have little or no effect when compatibility is high. Leonard (1959) provided an initial demonstration that an RT function that does not increase as set size is increased can be obtained when stimuli and responses are highly compatible. Using vibrotactile stimuli to the respective fingers, he found no difference in RTs between two, four and eight choices; that is, the function relating number of alternatives and RT was flat (but see Ten Hoopen, Akerboom, & Raaymakers, 1982, for qualification of this finding). Brainard, Irby, Fitts, and Alluisi (1962) showed a similar result of no significant influence of set size on visual digit stimuli that were to be named, but this could have been due in part to coding the stimulus set as the complete set of digits in all cases because the sets of two (Digits 4 and 7) and four (Digits 3, 4, 7 and 8) were not familiar subsets (Fitts & Switzer, 1962). However, Alluisi, Strain, and Thurmond (1964) obtained a null effect of set size in the digit-naming task when the subsets were familiar (Digits 1 and 2; Digits 1-4; Digits 1-8), providing evidence that lack of familiarity with the subsets could not be the whole story.

Several other situations that yield flat or even reversed RT functions were identified more recently. Kveraga, Boucher, and Hughes (2002) found no influence of number of alternatives for saccadic eye movements toward the locations where the stimuli appeared (prosaccades; see also Kveraga & Hughes, 2005; Saslow, 1967; a similar finding for smooth pursuit eye movements was reported by Berryhill, Kveraga, Boucher, & Hughes, 2004). However, Kveraga and colleagues did find set-size effects consistent with Hick's law for antisaccades (eye movements away from the stimulus location) and for manual keypress responses. Lawrence (2010; Lawrence, St John, Abrams, & Snyder, 2008) obtained a reversed set-size effect (which they called an anti-Hick's effect) for saccades toward stimulus locations but a standard set-size effect when the stimuli were central arrows pointing to locations. The anti-Hick's effect for saccadic eye movements was evident even when eye-movement responses were performed concurrently with aimed hand movements to the visual stimulus on a screen (Lawrence & Gardella, 2009). Dassonville et al. (1999; see also Berryhill, Kveraga, & Hughes, 2005) found no increase in RT from two to eight alternatives with a spatially compatible mapping in their study using joystick-controlled cursor movements, but not when the target location was signaled by letter stimuli indicating compass directions. Wright, Marino, Belovsky, and Chubb (2007) found no set-size effect for visually guided aimed movements, but they did obtain one with keypress responses.

Whereas the lack of set-size effects for vibrotactile and prosaccadic responses can be accommodated by assuming that they bypass a limited-capacity response-selection mechanism, the lack of an effect for aimed cursor and arm movements seems more problematic. Wright, Marino, Chubb, and Rose (2011) evaluated a hypothesis that the key factor is being able to make a direct attention shift to the target location, but they concluded that their findings did not support such an account. They suggested that the reason why keypress responses show a significant increase in RT as uncertainty increases, whereas aimed movements do not, is that the keypresses are made in a different response plane and require effector selection, whereas the aimed movements do not. In the case of Dassonville et al.'s (1999) study, this relation holds for the action effect (cursor movement), though not for the plane in which the physical movement is made.

Schneider and Anderson (2011) noted that many of the demonstrations of null or negligible set-size effects not conforming to Hick's law involved highly compatible situations in which subjects did not need to maintain the S-R associations in memory. For example, pressing the key that vibrates below a finger (Leonard, 1959), making a saccade toward the location of a stimulus (Kveraga et al., 2002) and using a joystick to move a cursor to a stimulus location (Dassonville et al., 1999) are all situations that do not require memory for different S-R alternatives. In contrast, situations that require remembering S-R associations, such as arbitrary mappings between stimulus locations and vocal responses (Hyman, 1953), do tend to show set-size effects consistent with Hick's law. Thus, the role of memory in the choice reaction task might be an important determinant of empirical restrictions on Hick's law.

Practice

Practice is another factor that influences the slope of the Hick's law function. Not too surprisingly, the slope reduces with practice (Davis, Moray, & Treisman, 1961; Hale, 1968; Mowbray & Rhoades, 1959; for a review, see Teichner & Krebs, 1974). For example, Hale (1968) had different groups of subjects perform a numerical choice reaction task with two, four or eight alternatives and a compatible mapping of stimuli to responses for 1000 trials in each of five sessions. In the first session, RT was close to 500 ms longer with eight choices than with two choices, but in the fifth session, this difference was reduced to a little over 300 ms. In a recent study, Wifall, Hazeltine, and Mordkoff (2016; Experiment 1) had subjects practice a choice reaction task with eight alternatives for six sessions of 640 trials per session, followed by separate sessions involving two, four or eight alternatives. Hick's law was still obtained in the later sessions after nearly 4000 trials of practice, but the difference in mean RTs between eight and two alternatives was only 149 ms.

Teichner and Krebs (1974) analyzed results from many experiments and concluded that with sufficient practice, RT might become independent of set size. They summed up their detailed study of the then-existing literature on visual choice RT as follows:

> In terms of their influence on choice reaction time, the three most important variables appear to be level of practice (N_T), number of different possible stimulus-response pairs (N_A), and the particular stimulus-response combination used. These variables operate jointly. Thus, while practice serves to reduce the slope of the curve relating choice reaction time to N_A, the slope itself is dependent on the stimulus-response code. (p. 94)

Although Hick's law provides a good description of the influence of set size and information more generally in a variety of situations, Teichner and Krebs' statement, along with the influence of S-R compatibility mentioned earlier, makes clear that the slope of the function can vary between zero and several hundreds of milliseconds.

Additional challenges to Hick's law

Teichner and Krebs' (1974) review indicated that dozens of studies were conducted in the 1950s and 1960s in which set sizes were manipulated and rates of information transmission were calculated in various contexts. As noted, in many studies, choice RTs were found to vary linearly with the logarithm of set size (or the amount of information), and this empirical regularity is one of the reasons why the term "Hick's law" came into use. However, it was inevitable that some researchers would try to "break" the law from either an empirical angle (e.g., by identifying boundary conditions in which the typical set-size effect was not observed) or a theoretical angle (e.g., by identifying weaknesses in the information-theoretic interpretation of the set-size effect). In this section, we review some of these additional challenges to Hick's law. To avoid redundancy with earlier text, we omit the previously discussed findings of null set-size effects obtained under conditions of high S-R compatibility or extensive practice, which have also been considered challenges to Hick's law.

Very large set sizes

Many studies of Hick's law involve manual keypress responses, whereby a single key is depressed by a single finger. Due to the physiological constraint of having 10 fingers, the maximum set size used in these experiments does not exceed 10 alternatives (or 3.3 bits of information). As a result, Hick's law has been investigated predominantly for choice RT tasks in which set size was manipulated in the range of 2-10 alternatives. A few studies involving set sizes beyond that range have yielded mixed results with respect to whether the data conform to Hick's law.

Conrad (1962) reported an experiment in which subjects read lists of nonsense syllables. Each list consisted of 320 syllables, and set size was manipulated from 4 to 32 across lists by varying the number of distinct syllables. Conrad found that reading time per item was linearly related to the logarithm of set size across the entire range, consistent with Hick's law (see also Davis et al., 1961; Experiment 2). In a second experiment, Conrad used separate lists of nonsense syllables and three-letter words, with set sizes of 4 and 32. Longer RTs were obtained for the larger than the smaller set of nonsense syllables, but no set-size effect was found for words. The latter finding replicates an earlier result from Pierce and Karlin (1957). In their first experiment, subjects simply read pages of words as quickly as possible, with the vocabulary set size (number of distinct words) ranging from 2 to 256. Except for a

slight advantage with a set size of 2, reading rate was relatively unchanged as set size increased from 4 to 256.

Instead of having subjects read words, Pollack (1963) did an experiment in which they classified words into predefined response categories (e.g., *goat* and *pig* were to be given an "animal" response, whereas *piano* and *drum* were to be given a "music" response). Across lists of words, the number of exemplars per response category varied from 1 to 24, and the number of response categories varied from 2 to 48. The conditions involving a single exemplar per response category are arguably the most appropriate for testing Hick's law. For these conditions, there was a general trend for longer RTs with increasing set size (number of response categories), but the data pattern was not strictly monotonic (i.e., there were cases in which a larger set size yielded a slightly shorter or equivalent RT when compared with a smaller set size).

Hilgendorf (1966) conducted an experiment in which subjects made keypress responses to different sets of symbolic stimuli (mostly numbers) ranging in size from 2 to 1000. For the largest set sizes, subjects saw a series of numbers on each trial (e.g., sampled from 000 to 999 for the set size of 1000) and pressed the corresponding sequence of keys after releasing a home key. RT was partitioned into two components, one associated with recognition (measured as the time from stimulus onset to release of the home key) and one associated with movement (measured by subtracting recognition time from the total time elapsed from stimulus onset to the last keypress). Hilgendorf found that all three RT measures (recognition, movement and total time) increased linearly with the logarithm of set size, supporting Hick's law.

However, a different outcome was obtained by Seibel (1963) in a study that involved a set size exceeding 1000 alternatives. He had subjects respond to different patterns of 10 lights by making spatially compatible patterns of simultaneous keypresses on a 10-key response device. For example, illumination of the second, third and seventh lights required pressing the second, third and seventh keys within a 100-ms interval from the first to the last keypress. Using all possible patterns of the 10 lights yielded a set size of $2^{10} - 1 = 1023$ alternatives. After extensive practice that yielded typical learning curves, Seibel (1963) found that RTs for the set size of 1023 alternatives averaged only 20-30 ms longer than RTs in sessions with a restricted set size of 31 alternatives. In a similar experiment with set sizes ranging from 5 to 31 alternatives, Seibel (1962) obtained null or inconsistent set-size effects. Based on these results, Seibel (1963) suggested that the linear relationship between RT and the logarithm of set size held in the range of two to eight alternatives, but not beyond that range.

Longstreth (1988) arrived at a similar conclusion based on experiments involving numerical stimuli and keypress responses. In his Experiment 1, set size was varied from 2 to 21 alternatives, with most sets consisting of one- and two-digit numbers (e.g., 4, 7 and 47 were the stimuli in a set of size 3) that required single- and double-keypress responses, respectively. RTs to the stimuli common to all sets (4 and 7) exhibited an approximately linear relationship with the logarithm of set size up to 10 alternatives, consistent with Hick's law, but the function flattened out at larger set sizes. This curvilinear relationship across the entire range of set sizes was replicated in two additional experiments, leading Longstreth to conclude that Hick's law did not apply beyond about eight alternatives (3 bits of information).

We think there are two important considerations to bear in mind when interpreting the results of the studies reviewed here. First, studies of word reading with varying vocabulary set sizes (e.g., Conrad, 1962; Pierce & Karlin, 1957) tap into a highly practiced skill; therefore, it is unclear whether departures from Hick's law can be attributed to extreme set sizes rather than to extensive practice. Second, for studies not involving word reading, very large set sizes often require tasks, stimuli or responses that differ markedly from those used in typical studies of Hick's law. For example, Seibel (1963) acknowledged that his findings might reflect the uniqueness of using patterns of light stimuli and simultaneous keypress responses, instead of having a single light mapped to a single keypress for each alternative. Fitts and Posner (1967) suggested that the difficulty of S-R coding might become critical for tasks involving large set sizes, especially when the S-R mappings are novel or arbitrary.

Alternative equations

Many researchers have found that choice RT data are well described by a linear equation involving either the logarithm of set size or the average information in a series (i.e., Hick's law). In general, it does not matter which base is used for the logarithm (although when the base is 2, then information is expressed in bits) or whether the logarithm's argument is simply the set size or the set size plus 1 (the latter being preferred by Hick, 1952). However, some authors have questioned the logarithmic relationship or suggested alternative equations as replacements for or extensions to Hick's law.

A notable effort in this vein was by Longstreth, El-Zahhar, and Alcorn (1985). They conducted a series of experiments in which numerical stimuli required responses of different keypress durations. For example, responding to the stimulus 2 involved pressing a key for two time units and then releasing it (in their first experiment, one time unit was equal to 30 ms). Set size was manipulated by varying the set of possible numbers from which the stimulus was selected on each trial. Longstreth and colleagues obtained null or very shallow set-size effects, leading them to argue that Hick's law is false. They suggested that set-size effects

might be attributable to attention switching rather than to information transmission. According to their theoretical account, if the relevant S-R alternative in a set of size n was in the focus of attention (occurring with a probability of $1/n$ for equiprobable alternatives), then responding was fast; otherwise, attention had to be switched to the relevant alternative (with a probability of $1-1/n$). They proposed an equation of the form $RT=a+b(1-n^{-1})$, where a is a constant and b is the time to switch attention. This power function was shown to produce set-size effects that seemed consistent with empirical data, at least for the studies selected by Longstreth and colleagues.

Welford (1987) argued that Longstreth et al.'s (1985) dismissal of Hick's law was too far-reaching. Welford summarized several data sets in which a logarithmic function relating RT to set size (i.e., Hick's law) yielded better fits than did the power function proposed by Longstreth and colleagues. Among other criticisms, he noted that some of the power function fits to data yielded negative values for the intercept parameter a, which was psychologically implausible. Finally, he discussed how variants of Hick's (1952) progressive classification model might be able to explain slight departures from a log-linear relationship. In a reply, Longstreth and Alcorn (1987) continued to favor the power function, regarding fit differences to be small and countering Welford's other criticisms. Longstreth (1988) also argued that the power function was superior to the logarithmic function in accounting for data from experiments that included very large set sizes. In a commentary on Longstreth's (1988) paper, Kvålseth (1989) reiterated some of Welford's key points, highlighted an error in the derivation of the power function and discussed how the function could be extended to accommodate unequal stimulus probabilities and repetition effects.

More recently, Mordkoff and colleagues have proposed alternative equations as extensions to Hick's law (Mordkoff, in press; Wifall et al., 2016). The work of Wifall et al. was motivated by the observation that in many studies of Hick's law, there is a 1:1 mapping of stimuli to responses. Consequently, the number of stimuli is confounded with the number of responses when set size is manipulated. This confound makes it unclear whether Hick's law reflects changes in stimulus set size, response set size or both. This issue has been partially addressed in past research using the information-reduction procedure (Posner, 1964), which involves a many-to-one mapping of stimuli to responses. By this procedure, one can either increase the number of stimuli while holding the number of responses constant (e.g., going from a 4:2 mapping to an 8:2 mapping) or increase the number of responses while holding the number of stimuli constant (e.g., going from a 4:2 mapping to a 4:4 mapping). Studies involving the information-reduction procedure have revealed a general tendency (not without exceptions) for RTs to become longer as each set-size variable increases (e.g., Fitts & Biederman, 1965; Keele, 1970; Morin, Forrin, & Archer, 1961; Pollack, 1963; Rabbitt, 1968). For example, in Rabbitt's experiment, different groups of subjects had 4:2, 8:2, 8:4 or 8:8 mappings of digit stimuli to keypress responses. A stimulus set-size effect was reflected in longer mean RTs with the 8:2 mapping than with the 4:2 mapping, whereas a response set-size effect was reflected in longer mean RTs going from the 8:2 to 8:4 to 8:8 mappings.

Wifall et al. (2016) were interested in which set-size variable contributes more to Hick's law. Their experiments involved simplified Chinese characters as stimuli and Hyman's (1953) single-syllable pseudowords as vocal responses. In Experiment 1, they used 2:2, 4:4 and 8:8 mappings of stimuli to responses and found a set-size effect consistent with Hick's law. In Experiment 2, they used 2:2, 4:2 and 8:2 mappings to manipulate only stimulus set size and found a shallow set-size effect relative to that of Experiment 1. In Experiment 3, they used 4:2, 4:4 and 8:4 mappings to partially disentangle manipulations of stimulus and response set size, and they found that the response set-size effect (comparing data from the 4:4 and 4:2 mappings) was larger than the stimulus set-size effect (comparing data from the 8:4 and 4:4 mappings). Given that response set size seemed to matter more than stimulus set size, Wifall et al. proposed that the standard information-based equation for Hick's law, $RT=a+bH$, should be amended to reflect the separate contributions of stimulus and response variables: $RT=a+b_sH_s+b_rH_r$, where the s and r subscripts denote stimulus and response parameters, respectively. It will be informative to see whether future studies involving different types of stimuli and responses yield stimulus and response set-size effects of differing magnitudes that support this extended equation for Hick's law.

Mordkoff (in press) proposed a different extension to the formulation of Hick's law. His work was motivated by the observation that standard equations for Hick's law explain mean RTs for blocks of trials rather than for individual trial types within blocks. For example, when Hyman (1953) manipulated the average information in blocks by varying the probability of occurrence of particular S-R pairs, he found that mean RTs at the block level were well described by the equation $RT=a+bH$, but that the resulting mean RT predictions derived for trial types with different information values (e.g., low- and high-probability S-R pairs) did not match the observed means. He concluded, "These analyses mean that we cannot predict, on the basis of the regression line fitted to the means of the conditions, what the mean reaction times will be to the components which make up a condition" (Hyman, 1953, p. 194).

To address this issue, Mordkoff (in press) explored different hierarchical variants of Hick's law, evaluating how well they described block- and trial-level RTs for a choice reaction task with three alternatives that varied in

probability within different blocks. In the first layer of each variant, mean RTs at the block level were regressed onto the average information in the blocks, and then residual times for each trial type were computed by subtracting predicted from observed RTs. In the second layer of each variant, the residual times were regressed onto a different predictor than average information. Mordkoff found that the secondary predictor yielding the best fit to his data was the relative probability of trial type i, $\Delta p(i)$, defined as the difference between that trial type's probability and the mean probability of all trial types in a particular block. This finding led to the following extended equation for Hick's law at the trial level: $RT(i) = a + bH + c\Delta p(i)$, where a and b are the intercept and the slope, respectively, from the first-layer application of the standard equation for Hick's law to block-level data, and c is the slope from the second-layer application to residual times. A notable feature of the extended equation is that when all trial types are equally likely, which is often the case when set size is manipulated without constraints on trial randomization, then $\Delta p(i) = 0$ and one is left with the standard equation for Hick's law, $RT = a + bH$. It will be instructive to see whether this extension to Hick's law produces good fits to future data sets.

Sequential effects

As demonstrated by Hyman (1953), the average information in a block of trials can be manipulated not only by varying set size but also by manipulating either the unconditional or the first-order conditional probabilities of different S-R alternatives for a fixed set size. Manipulating first-order conditional probabilities is generally the same as manipulating S-R repetition probabilities. Kornblum (1967, 1968, 1969) drew attention to the fact that typical manipulations of set size for equiprobable alternatives yield a confound with S-R repetition: as set size increases, S-R repetition probability decreases. For example, unconstrained randomization would yield 50% repetitions for a set size of 2, but only 25% repetitions for a set size of 4.

The confound represents a challenge to Hick's law because it is well established that there are sequential effects in choice reaction tasks, with S-R repetitions being faster than switches (Bertelson, 1961, 1963, 1965; Campbell & Proctor, 1993; Hale, 1969; Kornblum, 1967, 1975; Pashler & Baylis, 1991; Rabbitt, 1968; Smith, 1968; for reviews, see Kornblum, 1973; Luce, 1986). Given that repetitions occur less often as set size increases, it follows that repetition effects will contribute less to mean RTs for larger set sizes. The result is a putative set-size effect attributable to S-R repetition rather than to set size per se. However, it is unlikely that empirical set-size effects entirely reflect varying contributions of repetition effects to mean RTs, in part because set-size effects have been found in data restricted to specific S-R transitions (e.g., Kornblum, 1967; Rabbitt, 1968; Schneider & Anderson, 2011; Wifall et al., 2016).

Sequential effects also pose a challenge to the information-theoretic framing of Hick's law. Kornblum (1968, 1969) noted that first-order conditional probabilities can be manipulated for a fixed set size such that there is a nonmonotonic relationship between S-R repetition probability and average information in a block of trials. Consequently, it is possible to compare mean RTs for blocks that differ in repetition probability but have the same average information (e.g., for a set size of 4, blocks with repetition probabilities of 0.08 and 0.44 are both associated with an average of 1.87 bits of information; see Figure 1 in Kornblum, 1969). From an information-theoretic standpoint, if the rate of information transmission were constant (as implied by Hick's law), then one would predict equivalent mean RTs for blocks with the same average information. However, Kornblum's data were inconsistent with that prediction: For each comparison of equal-information blocks, mean RT was shorter for the block with the higher repetition probability (see also Bertelson, 1961). Not surprisingly, this difference was driven by shorter RTs for repetitions than for switches. Based on these findings, Kornblum concluded that the information-theoretic account of choice RT could be rejected.

Hyman and Umiltà (1969) argued that Kornblum's (1968) rejection of the information hypothesis was premature. They noted that his "serial" experimental design (marked by the use of a very short response–stimulus interval of 140 ms) may not have been ideal for enabling subjects to fully extract all available contextual information prior to stimulus onset. In a partial replication of Kornblum's (1968) experiment, but with a much longer response–stimulus interval of about 7.5 s, they found only very small differences between equal-information conditions. However, Kornblum's (1969) "discrete" experimental design (involving a minimum response–stimulus interval of 2.9 s) yielded a data pattern similar to that of his serial experiment, making it unclear what conditions are ideal for testing the information hypothesis.

Information theory and psychology

We briefly address a final challenge concerning the very use of information theory in psychology, which has implications for Hick's law. Laming (1966, 2001) noted an inconsistency in the analogy between the communication systems described by Shannon and Weaver's (1949) theory and the human subjects performing choice reaction tasks that motivated the theoretical perspective of Hick (1952) and others. Information theory applies to the rate at which a long, continuous series of signals is passed through a channel, which allows for accumulation of signals and detection of redundancies. In contrast, a typical choice reaction task involves serial and discrete processing of

individual stimuli (signals), such that a stimulus has to be processed completely and a response to it produced before the next stimulus in the series is presented. Laming argued that this mismatch between how information is used in communications and in psychology may underlie some of the issues associated with attempts to explain cognitive processing in terms of information transmission.

Luce (2003) echoed and expanded on this point in an article in which he reflected on why information theory faded from psychology after the 1960s. He noted that information theory in the context of communication systems deals with unstructured sets of neutral items that are interchangeable. In contrast, psychology often deals with structured sets of stimuli that are not interchangeable due to similarity, learned associations and other properties. Luce pointed out that the structural aspects of stimuli can strongly influence behavior, partly through the occurrence of sequential effects. In his opinion, the gradual realization that the structured stimuli of psychological experiments could not be treated in the same way as unstructured signals passing through a communication channel led to a diminishing role for information theory in psychology.

Contemporary models of Hick's law

In recent years, there has been considerable interest in developing and testing computational models of choice behavior that can account for RTs and errors for decisions involving more than two alternatives. Given that Hick's law characterizes the relationship between RT and the number of alternatives, it has served as a benchmark for evaluating these models at either a theoretical level or an empirical level (or both). Model evaluation at a theoretical level involves determining whether a model produces a set-size effect on RT that corresponds to the log-linear equation describing Hick's law (e.g., McMillen & Holmes, 2006; Usher & McClelland, 2001; Usher, Olami, & McClelland, 2002). Model evaluation at an empirical level involves determining whether a model produces the set-size effect on RT found in experimental data (e.g., Hawkins, Brown, Steyvers, & Wagenmakers, 2012a, 2012c; Leite & Ratcliff, 2010; Schneider & Anderson, 2011). Both approaches are reasonable, given that the equation representing Hick's law often provides an excellent descriptive account of data. In this section, we selectively review modern models of Hick's law, describing their basic features and explaining how they produce set-size effects in multi-alternative choice situations. For brevity, we omit discussion of related modeling work addressing unidimensional absolute identification (e.g., Brown & Heathcote, 2008; Lacouture & Marley, 1995) or the neural mechanisms underlying multi-alternative decisions (e.g., Albantakis & Deco, 2009; Churchland, Kiani, & Shadlen, 2008).

Evidence accumulation models

Most contemporary models of Hick's law are variants of evidence accumulation models for choice RT tasks (Luce, 1986; Ratcliff & Smith, 2004; Schweickert, 1993; Townsend & Ashby, 1983). These models work by gradually accumulating noisy evidence for different response alternatives. Each alternative typically has its own accumulator or counter, which means that the number of accumulators varies directly with the size of the set of response alternatives. Evidence accumulation is based on taking a series of noisy samples of input from the stimulus in either discrete time steps or continuously through time. The accumulator associated with the correct response receives stronger input than the other accumulators do, resulting in a higher rate of evidence accumulation for correct than for error responses. Each accumulator typically starts with zero evidence (assuming no bias toward a particular response) and sampling proceeds until there is sufficient evidence to reach either an absolute or a relative response criterion, at which point the response with the most evidence has been selected. With an absolute criterion, accumulation stops when the evidence in the highest accumulator reaches a threshold that is common to all the accumulators. With a relative criterion, accumulation stops when the evidence in the highest accumulator exceeds that of the next highest accumulator by a threshold, although other relative differences can also be used. This maximum minus next-to-maximum stopping rule approximates an optimal multi-hypothesis sequential probability ratio test when accuracy is high (Dragalin, Tartakovsky, & Veeravalli, 1999, 2000). The model's RT predictions are based on the time it takes to reach the criterion (plus time for non-decision processes such as stimulus encoding and response execution), and its accuracy predictions are based on which response alternative is chosen.

Starting with Usher and McClelland (2001), researchers interested in multi-alternative choice have examined various forms of an evidence accumulation model known as the leaky, competing accumulator model. Two key attributes distinguish this model from the generic model just described. First, there can be leakage during the accumulation process, such that accumulated evidence gradually decays over time. Second, there is competition among the accumulators for different response alternatives in the form of lateral inhibition, such that accumulators inhibit each other in proportion to their relative levels of accumulated evidence (i.e., an accumulator with more evidence will exert stronger inhibition over an accumulator with less evidence than vice versa). Leakage and inhibition modulate the amount of evidence gained from noisy, stimulus-driven input, thereby influencing how long it takes the accumulators to reach the criterion. In the special case of choosing among two alternatives with a relative criterion, no leakage and no lateral inhibition, the model becomes

formally equivalent to the standard diffusion model for two-choice behavior, which has a long history of success in cognitive psychology (Ratcliff, Smith, Brown, & McKoon, 2016).

As mentioned earlier, evidence accumulation models require additional accumulators as set size increases, such that there is a separate accumulator for each response alternative. Usher and McClelland (2001; Usher et al., 2002) noted that the addition of accumulators to a model with a fixed, absolute criterion will result in an increased error rate with increased set size (as well as shorter RTs, which is inconsistent with Hick's law) because there is a higher probability that one of the incorrect accumulators will spuriously reach the criterion before the correct accumulator due to noise. They argued that subjects strategically attempt to compensate for this effect by trying to maintain a constant error rate regardless of the set size. This strategy can be implemented in evidence accumulation models by varying the response criterion: As the criterion increases, more evidence has to be accumulated to reach the criterion, which increases accuracy but prolongs the decision-time component of RT. Indeed, varying the response criterion is the primary way in which the general class of evidence accumulation models accounts for speed–accuracy trade-offs in choice behavior (Ratcliff & Smith, 2004).

Usher and McClelland (2001) simulated three types of evidence accumulation models to determine the criterion needed at each set size (ranging from 2 to 10) to maintain an accuracy level of either 95% or 99%. One of the models was the leaky, competing accumulator model with an absolute criterion, whereas the other two models were accumulator models without leakage or inhibition, and either an absolute or a relative criterion. All three models generated simulated mean RTs that varied almost perfectly linearly with the logarithm of set size. Thus, Hick's law was produced by changing the response criterion in such a way that a constant accuracy level was achieved across a range of set sizes, essentially explaining the set-size effect as a speed–accuracy trade-off.

Formal properties of variants of the leaky, competing accumulator model have been explored in multiple studies (Bogacz, Usher, Zhang, & McClelland, 2007; McMillen & Holmes, 2006; Usher et al., 2002). To highlight one example, Usher et al. (2002) formally examined the relationship between response criteria and Hick's law for an accumulator model with an absolute criterion and neither leakage nor inhibition. They found that a constant accuracy level could be attained by having the criterion increase logarithmically with set size, which also resulted in the model yielding an RT pattern that corresponded to Hick's law. Usher and colleagues argued that strategic adaptation of the criterion to reduce errors at greater set sizes might be an intrinsic property of a decision process based on accumulation of noisy evidence.

Complementing the theoretical approach taken by Usher and McClelland (2001) and by others, Leite and Ratcliff (2010) explored which variants of the leaky, competing accumulator model could satisfactorily account for mean RTs, errors and the RT distributions for correct and error responses in empirical data. They assessed models that had (a) decay or no decay, (b) inhibition or no inhibition, (c) variable or identical starting points, (d) unbounded or bounded (i.e., non-negative) evidence accumulation, (e) one or multiple criteria, (f) one or multiple non-decision times and (g) various constraints on inputs. In fits to data from two experiments with set sizes ranging from 2 to 4, they found considerable mimicry among a small set of models, with little need for either decay or inhibition.

The best fitting models involved combinations of differences in three types of parameters. First, the criterion increased approximately logarithmically with set size (but only for the fits to one of the experimental data sets), which was interpreted as evidence that subjects became more cautious when choosing among more alternatives. Second, non-decision time also increased approximately logarithmically with set size, which was interpreted as possibly reflecting longer preparation times for responding at greater set sizes. Third, different input parameters underlying the rates of evidence accumulation were needed for different set sizes, but there was little in the way of a systematic relationship between input and set size, making this aspect of the modeling results harder to interpret. Thus, despite analyzing evidence accumulation models similar to those examined by Usher and colleagues, Leite and Ratcliff found that model fits to empirical data required more than adjusting the response criterion. Changes to non-decision times and inputs to the accumulators were also necessary to capture the RT distributions underlying the overarching pattern associated with Hick's law.

We provide final examples of evidence accumulation models for Hick's law from Brown and colleagues. Brown, Steyvers, and Wagenmakers (2009) investigated how multi-alternative decisions are made when subjects can actually observe evidence being accumulated for different alternatives, thereby reducing memory demands. In their experimental paradigm, several accumulators are displayed onscreen and evidence is represented by "bricks" that fall down from the top of the screen in discrete time steps to create columns on the accumulators. The subject's task is to choose the column with the highest growth rate. Set size is manipulated across trials by varying the number of potential accumulators. Brown et al. found that RTs varied with set size in accordance with Hick's law, although accuracy declined substantially with increasing set size. An important caveat is that the RTs from this task and variations of it (Hawkins et al., 2012a) tend to be relatively long—often 10 s or longer—compared with typical choice RTs of less than 1 s. This difference in time scale, in conjunction with the unique visual representation of evidence

accumulation, makes it unclear whether this experimental paradigm taps into the same mechanisms underlying traditional demonstrations of Hick's law.

Brown et al. (2009) assessed two kinds of Bayesian optimal observer models for their experimental paradigm. Both models followed a multi-hypothesis sequential probability ratio test procedure to determine the probabilities of different hypotheses (about which column has the highest growth rate), given the observed data, and whether evidence accumulation should stop or continue based on a decision criterion. One model assumed that the evidence accumulation probabilities for the target and non-target columns were known exactly (as a consequence of practice), whereas the other model assumed they were known vaguely (due to noisy representations of the growth rates). Brown et al. found that the exact-probability optimal observer model produced Hick's law for RT while maintaining an almost constant accuracy level, and it did so without changes to the criterion as set size increased (cf. Usher & McClelland, 2001; Usher et al., 2002). However, as noted earlier, in their experimental data, they found that accuracy declined with increasing set size. Further modeling efforts revealed that the data were explained better by a sub-optimal heuristic model that makes a decision based on whether the difference in heights between the highest and second highest columns meets a criterion (i.e., a maximum minus next-to-maximum stopping rule). Interestingly, the criterion estimated from the experimental data tended to decrease rather than increase with set size, suggesting that subjects became more impatient over the course of a trial, especially one involving many accumulators. This suggestion received empirical support in a subsequent study by Hawkins, Brown, Steyvers, and Wagenmakers (2012b).

In related work, Hawkins et al. (2012a) closely examined the relationship between RT and accuracy in multi-alternative decision-making, as well as its implications for modeling. In two experiments involving a variation of the visual evidence accumulation paradigm used by Brown et al. (2009), they found signs that contextual factors influence whether accuracy stays constant or declines with increasing set size. A Bayesian optimal observer model was able to accommodate both accuracy patterns—while still producing Hick's law for RTs—by changing the goal to be optimized. In particular, a decline in accuracy with set size could be explained by the model when optimality was redefined in terms of minimizing experiment time for a fixed accuracy level, which can also be considered optimization of the reward rate (see also Hawkins et al., 2012c).

Memory-based model

In contrast with modeling approaches based directly on evidence accumulation, Schneider and Anderson (2011) developed and evaluated a memory-based model of Hick's law and related phenomena (for an application of an exemplar-based memory model to Hick's law, see Jamieson & Mewhort, 2009). Their model was motivated by two observations discussed earlier. First, set-size effects consistent with Hick's law are often obtained with choice RT tasks that require access to S-R associations in memory, whereas negligible or absent set-size effects tend to be found with tasks involving little or no memory demands. Second, given the typical confound between set size and the probability of S-R repetition in experiments addressing Hick's law, a set-size effect can emerge as an artifact of the changing contribution of repetition effects to RT. These observations led to a model that explains Hick's law using a combination of two qualitatively distinct memory mechanisms.

The first mechanism involves memory retrieval of the response associated with the stimulus presented on a given trial. Memory retrieval in the model is based on the structure and functioning of the declarative memory module in the Adaptive Control of Thought–Rational (ACT-R) cognitive architecture (Anderson, 2007). Declarative knowledge in ACT-R is represented in units called chunks, which can be retrieved from memory if they are sufficiently active. The total activation of a chunk is determined by its base-level activation (reflecting its frequency and recency of use) and associative activation (reflecting activation received from associated retrieval cues). As a chunk's activation increases, the time to retrieve it from memory becomes shorter. Critically, associative activation is sensitive to the fan of each retrieval cue, which refers to the number of associations the cue has to chunks in memory. As fan increases, associative activation decreases and retrieval takes longer, resulting in a fan effect (Anderson, 1974; Anderson & Reder, 1999; Schneider & Anderson, 2012).

Schneider and Anderson (2011) proposed that both the stimulus and the set context (defined by the set of S-R alternatives) serve as retrieval cues for accessing chunks in memory representing the alternatives. Associative activation from the stimulus is invariant with set size because each stimulus is typically associated with a single response, but associative activation from the set context decreases as set size increases because the context cue is associated with the entire set of S-R alternatives. Consequently, memory retrieval takes longer as set size increases because set context becomes a less effective retrieval cue. This context-based fan effect on retrieval time is part of how the model produces a set-size effect.

The second mechanism in the model involves checking whether the stimulus presented on a given trial matches the stimulus from the immediately preceding trial. The chunk representing the S-R alternative on the previous trial is assumed to be retained in a memory buffer that can be inspected on the current trial. If there is a mismatch between current and previous stimuli, then the chunk in

the buffer is not appropriate and a new S-R chunk is retrieved from memory, as described earlier. However, if there is a match, then retrieval is skipped and the response associated with the chunk in the buffer is quickly emitted, resulting in shorter RTs for matches than for mismatches.

Schneider and Anderson (2011) allowed for variation in the probability of checking for a stimulus match, generally linking it to the probability of S-R repetition because only a repetition will yield a match that allows retrieval to be skipped. Given that the probability of S-R repetition typically decreases as set size increases, it follows that checking occurs less often with larger set sizes, resulting in fewer matches and less skipping of retrieval for repetitions. Consequently, repetition effects contribute less to RT as set size increases, which is the other part of how the model produces a set-size effect.

In tandem, the context-based memory retrieval process and the repetition-based matching process for occasionally skipping retrieval yield a set-size effect on RT that closely corresponds to Hick's law, as demonstrated by Schneider and Anderson's (2011) fits of the memory-based model to multiple published data sets, including Hick's (1952) data. Besides the basic set-size effect, Schneider and Anderson showed that their model can account for changes in the slope of the set-size effect with practice (Hale, 1968), S-R repetition effects (Kornblum, 1969), and a predicted interaction between set size, S-R transition and stimulus fan (the number of responses associated with a stimulus) that was obtained in two experiments they reported. Although the model was not applied to either errors or RT distributions, they described how both types of data could potentially be accommodated in an extended model involving the addition of noise to chunk activations in memory and the use of an activation threshold in the retrieval process. Moreover, they outlined how the model could be generalized to produce response repetition effects over and above stimulus repetition effects, as well as higher order sequential effects that are often observed in choice behavior.

Future models

This selective review shows that considerable progress has been achieved recently in modeling Hick's law. Future models will undoubtedly build on this success by addressing the limitations of existing models. One limitation that applies in different ways to different models concerns the depth versus the breadth of a model's explanation of Hick's law and related phenomena. On one hand, the evidence accumulation models reviewed here tend to have considerable depth but limited breadth within the domain of Hick's law, in that they can sometimes account for detailed data related to the basic manipulation of set size (e.g., RT distributions; Hawkins et al., 2012a; Leite & Ratcliff, 2010), but they have not been shown to account for modulation of the set-size effect by other experimental factors, such as practice and S-R repetition. On the other hand, the memory-based model of Schneider and Anderson (2011) has considerable breadth but limited depth, in that it accounts for practice and S-R repetition effects in addition to the basic set-size effect, but it has not been shown to account for empirical error patterns and RT distributions.

Schneider and Anderson (2011, p. 217) noted that it might be possible for the two modeling approaches to be integrated in the development of future models. For example, the Retrieval by Accumulating Evidence in an Architecture (RACE/A) model of Van Maanen and colleagues (Van Maanen & Van Rijn, 2007, 2010; Van Maanen, Van Rijn, & Taatgen, 2012) shows how evidence accumulation can be used to account for the dynamics of memory retrieval in ACT-R. Anderson (2007, pp. 131-134) discussed how retrieving a memory in ACT-R can be interpreted as accumulating evidence for a specific chunk in memory, and he showed that the ACT-R equation for retrieval time can be related to the decision time of a leaky accumulator model (see also Anderson & Betz, 2001). Continued efforts toward integration along these lines in the context of multi-alternative choice behavior might eventually produce a more comprehensive model of Hick's law with substantial depth and breadth.

Applications of Hick's law

Hick's (1952) study not only inspired large amounts of basic empirical and theoretical research but it also generated interest in more applied areas. Hick's law (or the Hick-Hyman law) has been the focus of research in intelligence testing for many years (Deary, 2000; Jensen, 2006). It is also foundational knowledge in applied design disciplines. The law is cited in textbooks on human factors and ergonomics (Proctor & Van Zandt, 2008; Wickens, Hollands, Parasuraman, & Banbury, 2013) and in more specialized books and handbooks on the topic (Salas & Maurino, 2010; Salvendy, 2012). The same holds true for HCI textbooks (Jones, 1989; MacKenzie, 2013) and handbooks (Helander, 1991; Jacko, 2012). We briefly cover the role of Hick's law in intelligence testing and HCI in the remainder of this section.

Speed of information processing and intelligence

Recall that Hick's (1951) article dealt with the relevance of information theory to intelligence testing. Given that the slope of the Hick's law function can be taken as a metric of central processing efficiency, it is not too surprising that research focused on it as a correlate of general intelligence. What is surprising, though, is that Hick's (1951) article has not been cited in that literature; on 1 April 2017, PsycINFO showed 2 citations of that article and Web of Science 15 citations, with only one of those by authors studying the

relation of Hick's law to intelligence (Roberts & Stankov, 1999). In a footnote, those authors acknowledged, "It would appear a little known fact that a year earlier Hick (1951) attempted to apply information-theory principles to model the concept of intelligence directly" (p. 99).

The first study examining the relation between intelligence and the slope of the Hick's law function seems to be that of Roth (1964), of which we have access only to summaries in English. According to Jensen and Munro (1979), subjects in Roth's study were required to turn off one of several lights as quickly as possible after its onset by pressing an adjacent button. The slope of the Hick's law function obtained by varying the number of S-R alternatives was found to correlate negatively with intelligence. Roth did not separate RT (time to lift off of a start button) from movement time, which Jensen and Munro did for a task requiring subjects to move from a home button to one of as many as eight buttons arranged about it in a semicircle when a light located adjacent to the button came on. Their results showed a correlation of −0.36 between the slope of the Hick's law function for RT and general intelligence as measured by the Raven's Progressive Matrices test.

Many subsequent studies have been conducted considering the relation of the Hick's law slope to intelligence and other individual differences (Jensen, 2006), including one showing that pigeons produce a shallower slope than humans (Vickrey & Neuringer, 2000). In a review of the literature on intelligence and speed of information processing, Sheppard and Vernon (2008) conclude that studies generally show a negative correlation of RT and measures of intelligence, which tends to increase as the set size increases (i.e., at higher information loads). However, they indicate that ". . . general intelligence correlates with speed of processing in short-term memory to a somewhat greater degree than it does with Hick RTs" (p. 541).

HCI and interface design

As pointed out at the beginning of this article, Hick's law also is a staple in HCI. Card, Moran, and Newell (1983) developed The Model Human Processor as an initial cognitive architecture that would allow predictions to be made relatively easily for situations in which humans interact with a computer workstation. Among its principles is the Information Theory Principle, which is specified as Hick's law. The book *Universal Principles of Design* devotes two pages to an entry on Hick's law (Lidwell, Holden, & Butler, 2010), in which it is stated, "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 (2016) states, "Understanding Hick's law means you can design so that more users will visit and stay on your website."

An early empirical study showing the value of Hick's law to HCI was that of Landauer and Nachbar (1985) on selection from menu trees. They had subjects select a target item (an integer from the ordered integers 1 to 4096, or a word from 4096 alphabetically ordered words) by a series of touch-menu choices between sequentially subdivided ranges. The number of alternatives at each step was 2, 4, 8 or 16, with the number of screens on which a choice had to be made increasing accordingly. Landauer and Nachbar noted that the resulting RT data were fit well by a logarithmic function, as implied by Hick's law and Fitts' law, rather than by a linear function, which would be more in agreement with visual search—for which "there is strong evidence that the visual search mechanism basically proceeds in a serial mode" (Schwarz & Miller, 2016, p. 1656), yielding a linear function. Because a property of the logarithmic relation is that the difference in number of alternatives per screen has no effect on the overall time to select among the total alternatives, the cost associated with number of screens is the only factor affecting time to locate the target. Consequently, Landauer and Nachbar concluded, "Broader, shallower menu trees yield faster search than narrower, deeper ones" (p. 76).

In promoting Hick's law for HCI designers, Ali and Liem (2014, p. 4) pointed out that its range of applications extends beyond menus:

> Within the context of design, this law promotes the use of design methods to simplify decision-making in situations, where the designer is presented with multiple options. In practice, Hick's law has fundamentally proven to be effective in the design of software menus, control display, way finding layout.

Chan, Goswami, and Kim (2012) provided another example application, in this case for alternative referencing styles in a spreadsheet program, Microsoft Excel. The task was to locate the precedent cell from a spreadsheet formula displayed after clicking the dependent cell (e.g., if cell C4 within the spreadsheet has the formula "=3*B1," B1 is the precedent cell referenced by the dependent cell C4). Two different referencing styles for the precedent style available in Excel were used. The A1 style labels columns with letters and rows with numbers (B1 in the formula above), whereas the R1C1 style designates the precedent cell by the number of rows and columns that separate it from the dependent cell (=3 × R[−3]C[−1]). Of importance, two different measures of number of alternative choices can be developed based on these referencing styles, one anchored around A1 and the other around the dependent cell. RT was found to vary in accordance with the uncertainty associated with whichever referencing representation was in effect but not with the other. The authors concluded that their finding has implications not only for spreadsheet designers but also for designers dealing with alternative problem representations.

Although some authors have focused on the logarithmic relation between uncertainty and RT, as in the above

studies, others have emphasized Hick's (1952) progressive classification model. Wang (2012, p. 20) indicates,

> Essentially, Hick's law provides a general guideline for the design and use of hierarchical menu structures. This is consistent with the study (Landauer & Nachbar, 1985) 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.

Progressive classification has also been invoked in the context of data visualization, which is regarded as an important tool when dealing with complex data, as in the area of cybersecurity. In an article on rethinking how risk and security are visualized, Hall, Heath, and Coles-Kemp (2015) wrote,

> The visual information mantra of interactive media-oriented researcher Ben Shneiderman was "overview first, zoom and filter, then details on demand." This position accommodates a technique known as "progressive disclosure" which aims at initial simplification followed by the option of revealing additional content and options. It assumes, after psychologist William Edmund Hick, that the time needed to make a decision increases with the number of variables. (p. 93)

One can also find appeals to Hick's law in guidelines for design of applications (apps) for mobile devices (Nayebi, Desharnais, & Abran, 2013). Interestingly, Hick's law is also referenced in a book on startup coaching (Garcia, 2015), where it is described as a "recurring pattern" of which entrepreneurs should be aware, and as a principle in martial arts training (Patrick, 2016), given that the time for choosing among different maneuvers is limited during combat. Due to its emphasis on the need to minimize alternatives and uncertainty to speed action, applications of Hick's law seem boundless.

Conclusion

It seems fair to say that the widespread impact of Hick's (1952) article goes beyond what most psychological scientists will accomplish in their entire careers. Perhaps the greatest accomplishment that one can hope to achieve is to have not only an immediate influence on closely related research but also a long-lasting impact that extends out beyond the immediate topic. Hick's work also illustrates the value of basic research conducted from an applied perspective, or what Stokes (1997) calls "use-inspired basic research." The basic research pushes others further to examine the empirical relations and theoretical implications of the work, with this effort going in directions that cannot be seen at the time. Moreover, the contribution, though primarily fundamental, ties back into technologies and applications that could not have remotely been envisioned at the time the original research was conducted. Hick's (1952) article has had a clear impact on the field over the past 65 years, and it will likely continue to influence researchers in the future.

Acknowledgements

We thank Guy Hawkins, Sander Los and an anonymous reviewer for their comments on this review article.

Declaration of conflicting interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

References

Albantakis, L., & Deco, G. (2009). The encoding of alternatives in multiple-choice decision making. *Proceedings of the National Academy of Sciences USA*, *106*, 10308–10313.

Ali, A., & Liem, A. (2014, August 27-29). *The use of formal aesthetic principles as a tool for design conceptualisation and detailing*. DS 81: Proceedings of NordDesign 2014, Espoo, Finland.

Alluisi, E. A., Strain, G. S., & Thurmond, J. B. (1964). Stimulus-response compatibility and the rate of gain of information. *Psychonomic Science*, *1*, 111–112.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*, 451–474.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press.

Anderson, J. R., & Betz, J. (2001). A hybrid model of categorization. *Psychonomic Bulletin & Review*, *8*, 629–647.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, *128*, 186–197.

Bates, J. A. V. (1947). Some characteristics of a human operator. *Journal of the Institution of Electrical Engineers-part IIA: Automatic Regulators and Servo Mechanisms*, *94*, 298–304.

Berryhill, M., Kveraga, K., Boucher, L., & Hughes, H. C. (2004). Smooth pursuit under stimulus—response uncertainty. *Cognitive Brain Research*, *19*, 100–102.

Berryhill, M., Kveraga, K., & Hughes, H. C. (2005). Effects of directional uncertainty on visually-guided joystick pointing. *Perceptual and Motor Skills*, *100*, 267–274.

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology*, *13*, 90–102.

Bertelson, P. (1963). S-R relationships and reaction times to new versus repeated signals in a serial task. *Journal of Experimental Psychology*, *65*, 478–484.

Bertelson, P. (1965). Serial choice reaction-time as a function of response versus signal-and-response repetition. *Nature*, *206*, 217–218.

Bogacz, R., Usher, M., Zhang, J., & McClelland, J. L. (2007). Extending a biologically inspired model of choice: Multi-alternatives, nonlinearity and value-based multidimensional choice. *Philosophical Transactions of the Royal Society B*, *362*, 1655–1670.

Brainard, R. W., Irby, T. S., Fitts, P. M., & Alluisi, E. A. (1962). Some variables influencing the rate of gain of information. *Journal of Experimental Psychology*, *63*, 105–110.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*, 153–178.

Brown, S. D., Steyvers, M., & Wagenmakers, E.-J. (2009). Observing evidence accumulation during multi-alternative decisions. *Journal of Mathematical Psychology*, *53*, 453–462.

Brysbaert, M. (2016). Editorial QJEP classics revisited. *Quarterly Journal of Experimental Psychology*, *69*, 1861–1863.

Campbell, K. C., & Proctor, R. W. (1993). Repetition effects with categorizable stimulus and response sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1345–1362.

Cao, J. (2010). *Applying Hick's law to Web design. Free example wireframes*. Retrieved from https://www.uxpin.com/studio/blog/applying-hicks-law-to-web-design-free-example-wireframes/

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Mahwah, NJ: Lawrence Erlbaum.

Chan, H. C., Goswami, S., & Kim, H. W. (2012). An alternative fit through problem representation in cognitive fit theory. *Journal of Database Management*, *23*, 22–43.

Churchland, A. K., Kiani, R., & Shadlen, M. (2008). Decision-making with multiple alternatives. *Nature Neuroscience*, *11*, 693–702.

Conrad, R. (1962). Practice, familiarity and reading rate for words and nonsense syllables. *Quarterly Journal of Experimental Psychology*, *14*, 71–76.

Craik, K. W. (1943). *The nature of explanation*. Cambridge, UK: Cambridge University Press.

Crossman, E. R. F. W. (1953). Entropy and choice time: The effect of frequency unbalance on choice-response. *Quarterly Journal of Experimental Psychology*, *5*, 41–51.

Dassonville, P., Lewis, S. M., Foster, H. E., & Ashe, J. (1999). Choice and stimulus–response compatibility affect duration of response selection. *Cognitive Brain Research*, *7*, 235–240.

Davis, R., Moray, N., & Treisman, A. (1961). Imitative responses and the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *13*, 78–89.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford, UK: Oxford University Press.

Donders, F. C. (1969). *On the speed of mental processes*. In W. G. Koster (Ed.), *Attention and performance II* (pp. 412–431). Amsterdam, The Netherlands: North-Holland. (Original work published 1868)

Dragalin, V. P., Tartakovsky, A. G., & Veeravalli, V. V. (1999). Multihypothesis sequential probability ratio tests—Part I: Asymptotic optimality. *IEEE Transactions on Information Theory*, *45*, 2448–2461.

Dragalin, V. P., Tartakovsky, A. G., & Veeravalli, V. V. (2000). Multihypothesis sequential probability ratio tests—Part II: Accurate asymptotic expansions for the expected sample size. *IEEE Transactions on Information Theory*, *46*, 1366–1383.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*, 381–391.

Fitts, P. M., & Biederman, I. (1965). S-R compatibility and information reduction. *Journal of Experimental Psychology*, *69*, 408–412.

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.

Fitts, P. M., & Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, *46*, 199–210.

Fitts, P. M., & Switzer, G. (1962). Cognitive aspects of information processing: I. The familiarity of S-R sets and subsets. *Journal of Experimental Psychology*, *63*, 321–329.

Garcia, R. (2015). *StartUP SOAR coaching: Coaching the StartUP for successful outcomes*. New York, NY: Buoyant Capital.

Hale, D. J. (1968). The relation of correct and error responses in a serial choice reaction task. *Psychonomic Science*, *13*, 299–300.

Hale, D. J. (1969). Repetition and probability effects in a serial choice reaction task. *Acta Psychologica*, *29*, 163–171.

Hall, P., Heath, C., & Coles-Kemp, L. (2015). Critical visualization: A case for rethinking how we visualize risk and security. *Journal of Cybersecurity*, *1*, 93–108.

Hawkins, G. E., Brown, S. D., Steyvers, M., & Wagenmakers, E.-J. (2012a). Context effects in multi-alternative decision making: Empirical data and a Bayesian model. *Cognitive Science*, *36*, 498–516.

Hawkins, G. E., Brown, S. D., Steyvers, M., & Wagenmakers, E.-J. (2012b). Decision speed induces context effects in choice. *Experimental Psychology*, *59*, 206–215.

Hawkins, G. E., Brown, S. D., Steyvers, M., & Wagenmakers, E.-J. (2012c). An optimal adjustment procedure to minimize experiment time in decisions with multiple alternatives. *Psychonomic Bulletin & Review*, *19*, 339–348.

Helander, M. (Ed.). (1991). *Handbook of human-computer interaction* (2nd ed.). Amsterdam, The Netherlands: North-Holland.

Hick, W. E. (1948). The discontinuous functioning of the human operator in pursuit tasks. *Quarterly Journal of Experimental Psychology*, *1*, 36–51.

Hick, W. E. (1951). Information theory and intelligence tests. *British Journal of Psychology*, *4*, 157–164.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*, 11–26.

Hilgendorf, L. (1966). Information input and response time. *Ergonomics*, *9*, 31–37.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, *53*, 188–196.

Hyman, R., & Umiltà, C. (1969). The information hypothesis and non-repetitions. *Acta Psychologica*, *30*, 37–53.

Jacko, J. A. (2012). *Human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications* (3rd ed.). Boca Raton, FL: CRC Press.

Jamieson, R. K., & Mewhort, D. J. K. (2009). Applying an exemplar model to the serial reaction-time task: Anticipating from experience. *Quarterly Journal of Experimental Psychology*, *62*, 1757–1783.

Jastrow, J. (1890). *The time-relations of mental phenomena* (No. 6). New York, NY: N.D.C. Hodges.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Oxford, UK: Elsevier.

Jensen, A. R., & Munro, E. (1979). Reaction time, movement time, and intelligence. *Intelligence*, *3*, 121–126.

Jones, M. K. (1989). *Human-computer interaction: A design guide*. Englewood Cliffs, NJ: Educational Technology Publications.

Keele, S. W. (1970). Effects of input and output modes on decision time. *Journal of Experimental Psychology*, *85*, 157–164.

Kornblum, S. (1967). Choice reaction time for repetitions and non-repetitions: A re-examination of the information hypothesis. *Acta Psychologica*, *27*, 178–187.

Kornblum, S. (1968). Serial-choice reaction time: Inadequacies of the information hypothesis. *Science*, *159*, 432–434.

Kornblum, S. (1969). Sequential determinants of information processing in serial and discrete choice reaction time. *Psychological Review*, *76*, 113–131.

Kornblum, S. (1973). Sequential effects in choice reaction time: A tutorial review. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 259–288). New York, NY: Academic Press.

Kornblum, S. (1975). An invariance in choice reaction time with varying numbers of alternatives and constant probability. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 366–382). New York, NY: Academic Press.

Kvålseth, T. O. (1989). Longstreth et al.'s reaction time model: Some comments. *Bulletin of the Psychonomic Society*, *27*, 358–360.

Kveraga, K., Boucher, L., & Hughes, H. C. (2002). Saccades operate in violation of Hick's law. *Experimental Brain Research*, *146*, 307–314.

Kveraga, K., & Hughes, H. C. (2005). Effects of stimulus—response uncertainty on saccades to near-threshold targets. *Experimental Brain Research*, *162*, 401–405.

Lacouture, Y., & Marley, A. A. J. (1995). A mapping model of bow effects in absolute identification. *Journal of Mathematical Psychology*, *39*, 383–395.

Laming, D. R. J. (1966). A new interpretation of the relation between choice-reaction time and the number of equiprobable alternatives. *British Journal of Mathematical and Statistical Psychology*, *19*, 139–149.

Laming, D. R. J. (2001). Statistical information, uncertainty, and Bayes' theorem: Some applications in experimental psychology. In S. Benferhat & P. Besnard (Eds.), *Symbolic and quantitative approaches to reasoning with uncertainty* (pp. 635–646). Berlin, Germany: Springer-Verlag.

Landauer, T. K., & Nachbar, D. W. (1985). Selection from alphabetic and numeric menu trees using a touch screen: Breadth, depth, and width. *ACM SIGCHI Bulletin*, *16*(4), 73–78.

Lawrence, B. M. (2010). An anti-Hick's effect for exogenous, but not endogenous, saccadic eye movements. *Experimental Brain Research*, *204*, 115–118.

Lawrence, B. M., & Gardella, A. L. (2009). Saccades and reaches, behaving differently. *Experimental Brain Research*, *195*, 413–418.

Lawrence, B. M., St John, A., Abrams, R. A., & Snyder, L. H. (2008). An anti-Hick's effect in monkey and human saccade reaction times. *Journal of Vision*, *26*, 1–7.

Leite, F. P., & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception, & Psychophysics*, *72*, 246–273.

Leonard, J. A. (1958). Partial advance information in a choice reaction task. *British Journal of Psychology*, *49*, 89–96.

Leonard, J. A. (1959). Tactual choice reactions. Part I. *Quarterly Journal of Experimental Psychology*, *11*, 76–83.

Lidwell, W., Holden, K., & Butler, J. (2010). *Universal principles of design*. Beverly, MA: Rockport.

Longstreth, L. E. (1988). Hick's law: Its limit is 3 bits. *Bulletin of the Psychonomic Society*, *26*, 8–10.

Longstreth, L. E., & Alcorn, M. B. (1987). Hick's law versus a power law: Reply to Welford. *Journal of Experimental Psychology: General*, *116*, 315–316.

Longstreth, L. E., El-Zahhar, N., & Alcorn, M. B. (1985). Exceptions to Hick's law: Explorations with a response duration measure. *Journal of Experimental Psychology: General*, *114*, 417–434.

Luce, R. D. (1986). *Response times*. New York, NY: Oxford University Press.

Luce, R. D. (2003). Whatever happened to information theory in psychology? *Review of General Psychology*, *7*, 183–188.

MacKenzie, I. S. (2013). *Human-computer interaction: An empirical perspective*. Waltham, MA: Morgan Kaufmann.

McMillen, T., & Holmes, P. (2006). The dynamics of choice among multiple alternatives. *Journal of Mathematical Psychology*, *50*, 30–57.

Merkel, J. (1885). Die zeitlichen verhältnesse der willenstätigkeit [The temporal behavior of the will]. *Philosophische Studien*, *2*, 73–127.

Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, *7*, 141–144.

Mordkoff, J. T. (in press). Effects of average uncertainty and trial-type frequency on choice response time: A hierarchical extension of Hick/Hyman law. *Psychonomic Bulletin & Review*.

Morin, R. E., Forrin, B., & Archer, W. (1961). Information processing behavior: The role of irrelevant stimulus information. *Journal of Experimental Psychology*, *61*, 89–96.

Mowbray, G. H., & Rhoades, M. V. (1959). On the reduction of choice reaction times with practice. *Quarterly Journal of Experimental Psychology*, *11*, 16–23.

MRC Applied Psychology Unit. (1946). *Progress report 1946*. Retrieved from http://www.mrc-cbu.cam.ac.uk/timeline/images/1946.pdf

MRC Applied Psychology Unit. (1950). *Progress report 1949-1950*. Retrieved from http://www.mrc-cbu.cam.ac.uk/timeline/images/19491950.pdf

Nayebi, F., Desharnais, J. M., & Abran, A. (2013, October). 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; pp. 147–155), Piscataway, NJ: Institute of Electrical and Electronics Engineers (IEEE).

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliff, NJ: Prentice-Hall.

Pashler, H., & Baylis, G. (1991). Procedural learning: 2. Intertrial repetition effects in speeded-choice tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 33–48.

Patrick, M. (2016). *Hick's law*. Retrieved from http://www.dragonsociety.com/hicks-law

Pierce, J. R., & Karlin, J. E. (1957). Reading rates and the information rate of a human channel. *Bell System Technical Journal*, *36*, 497–516.

Pollack, I. (1963). Speed of classification of words into superordinate categories. *Journal of Verbal Learning and Verbal Behavior*, *2*, 159–165.

Posner, M. I. (1964). Information reduction in the analysis of sequential tasks. *Psychological Review*, *71*, 491–504.

Posner, M. I. (1986) Overview. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. II cognitive processes and performance* (pp. V-3-V-10). New York, NY: John Wiley.

Proctor, R. W., & Van Zandt, T. (2008). *Human factors in simple and complex systems*. Boca Raton, FL: CRC Press.

Proctor, R. W., & Vu, K.-P. L. (2006). *Stimulus-response compatibility principles: Data, theory, and application*. Boca Raton, FL: CRC Press.

Psychreg.org. (2016). *William Edmund Hick*. Retrieved from http://www.psychreg.org/william-edmund-hick/

Rabbitt, P. M. A. (1968). Repetition effects and signal classification strategies in serial choice-response tasks. *Quarterly Journal of Experimental Psychology*, *20*, 232–240.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333–367.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*, 260–281.

Reynolds, L. A. (2004). Wellcome witnesses: The Medical Research Council Applied Psychology Unit. *History of Psychology*, *7*, 85–93.

Reynolds, L. A., & Tansey, E. M. (Eds.). (2001). *The MRC applied psychology unit* (*Vol. 16*, No. 16). London, England: Wellcome Trust Centre for the History of Medicine at UCL.

Roberts, R. D., & Stankov, L. (1999). Individual differences in speed of mental processing and human cognitive abilities: Toward a taxonomic model. *Learning and Individual Differences*, *11*, 1–120.

Roth, E. (1964). Die Geschwindigkeit der Verabeitung yon Information and ihr Zusammenhang mit lntelligenz [The speed of information processing and its relation to intelligence]. *Zeitschrift für Experimentelle und Angewante Psychologie*, *11*, 616–622.

Salas, E., & Maurino, D. (Eds.). (2010). *Human factors in aviation* (2nd ed.). San Diego, CA: Academic Press.

Salvendy, G. (Ed.). (2012). *Handbook of human factors and ergonomics* (4th ed.). Hoboken, NJ: John Wiley.

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America*, *57*, 1030–1033.

Sauro, J. (2013, March). Five HCI laws for user experience design. *MeasuringU*. Retrieved from http://www.measuringu.com/blog/hci-laws.php

Schneider, D. W., & Anderson, J. R. (2011). A memory-based model of Hick's law. *Cognitive Psychology*, *62*, 193–222.

Schneider, D. W., & Anderson, J. R. (2012). Modeling fan effects on the time course of associative recognition. *Cognitive Psychology*, *64*, 127–160.

Schwarz, W., & Miller, J. (2016). GSDT: An integrative model of visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1654–1675.

Schweickert, R. (1993). Information, time, and the structure of mental events: A twenty-five-year review. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV* (pp. 535–566). Cambridge, MA: MIT Press.

Seibel, R. (1962). Discrimination reaction time as a function of the number of stimulus-response pairs and the self-pacing adjustment of the subject. *Psychological Monographs: General and Applied*, *76*(42), Whole No. 561.

Seibel, R. (1963). Discrimination reaction time for a 1,023-alternative task. *Journal of Experimental Psychology*, *66*, 215–226.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences*, *44*, 535–551.

Smith, M. C. (1968). Repetition effect and short-term memory. *Journal of Experimental Psychology*, *77*, 435–439.

Soegaard, M. (2016). Hick's law: Making the choice easier for users. *The Interaction Design Foundation*. Retrieved from https://www.interaction-design.org/literature/article/hick-s-law-making-the-choice-easier-for-users

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.

Teichner, W. H., & Krebs, M. J. (1974). Laws of visual choice reaction time. *Psychological Review*, *81*, 75–98.

Ten Hoopen, G., Akerboom, S., & Raaymakers, E. (1982). Vibrotactual choice reaction time, tactile receptor systems and ideomotor compatibility. *Acta Psychologica*, *50*, 143–157.

Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592.

Usher, M., Olami, Z., & McClelland, J. L. (2002). Hick's law in a stochastic race model with speed-accuracy tradeoff. *Journal of Mathematical Psychology*, *46*, 704–715.

Van Maanen, L., & Van Rijn, H. (2007). An accumulator model of semantic interference. *Cognitive Systems Research*, *8*, 174–181.

Van Maanen, L., & Van Rijn, H. (2010). The locus of the Gratton effect in picture—word interference. *Topics in Cognitive Science*, *2*, 168–180.

Van Maanen, L., Van Rijn, H., & Taatgen, N. (2012). RACE/A: An architectural account of the interactions between learning, task control, and retrieval dynamics. *Cognitive Science*, *36*, 62–101.

Verbruggen, F., & Logan, G. D. (2009). Models of response inhibition in the stop-signal and stop-change paradigms. *Neuroscience & Biobehavioral Reviews*, *33*, 647–661.

Vickrey, C., & Neuringer, A. (2000). Pigeon reaction time, Hick's law, and intelligence. *Psychonomic Bulletin & Review*, *7*, 284–291.

Wang, H. (2012). A guide to assistive technology for teachers in special education. In M. Khosrow-Pour (Editor in chief), *Assistive technologies: Concepts, methodologies, tools, and applications* (Vol. *1*, pp. 12–25). Hershey, PA: IGI Global.

Welford, A. T. (1987). Comment on "Exceptions to Hick's law: Explorations with a response duration measure" (Longstreth, El-Zahhar, & Alcorn, 1985). *Journal of Experimental Psychology: General*, *116*, 312–314.

Wickens, C. D., Hollands, J. G., Parasuraman, R., & Banbury, S. (2013). *Engineering psychology and human performance* (4th ed.). New York, NY: Routledge.

Wiener, N. (1948). *Cybernetics or control and communication in the animal and the machine*. New York, NY: John Wiley.

Wifall, T., Hazeltine, E., & Mordkoff, J. T. (2016). The roles of stimulus and response uncertainty in forced-choice performance: An amendment to Hick/Hyman law. *Psychological Research*, *80*, 555–565.

Woodworth, R. S. (1938). *Experimental psychology*. New York, NY: Henry Holt.

Wright, C. E., Marino, V. F., Chubb, C., & Rose, K. A. (2011). Exploring attention-based explanations for some violations of Hick's law for aimed movements. *Attention, Perception, & Psychophysics*, *73*, 854–871.

Wright, C. E., Marino, V. M., Belovsky, S. A., & Chubb, C. (2007). Visually-guided, aimed movements can be unaffected by stimulus—response uncertainty. *Experimental Brain Research*, *179*, 475–496.