UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Stopping Rule Selection (SRS) Theory Applied to Deferred Decision Making

Permalink

https://escholarship.org/uc/item/9sf1n0sd

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN

1069-7977

Authors

Fific, Mario
Buckmann, Marcus

Publication Date

2013

Peer reviewed

Stopping Rule Selection (SRS) Theory Applied to Deferred Decision Making

Mario Fifić (fificm@gvsu.edu)
Department of Psychology, Grand Valley State University
Allendale, MI 49501 USA

Marcus Buckmann (buckmann@gmx.net)
Max Planck Institute for Human Development,
Center for Adaptive Behavior and Cognition, Berlin Germany

Abstract

The critical step facing every decision maker is when to stop collecting evidence and proceed with the decision act. This is known as the stopping rule. Over the years, several unconnected explanations have been proposed that suggest nonoptimal approaches can account for some of the observable violations of the optimal stopping rule. The current research proposes a unifying explanation for these violations based on a new stopping rule selection (SRS) theory. The main innovation here is the assumption that a decision maker draws from a large set of different kinds of stopping rules and is not limited to using a single one. The SRS theory hypothesizes that there is a storage area for stopping rules—the so-called decision operative space (DOS)— and a retrieval mechanism that is used to select stopping rules from the DOS. The SRS theory has shown good fit to challenging data published in the relevant literature.
.

Keywords: Stopping rule, deferred decision task, optimal, nonoptimal, decision making.

One of the most important steps of decision making is determining when to stop collecting evidence and proceed with the final decision. This is defined as the stopping rule and it is thought to be an irreplaceable component of almost all cognitive models of decision making.

Take, for example, a patient who is facing a risky medical treatment. The treatment can have a good outcome—that is, the patient will benefit from it—or it can have a bad outcome—that is, the patient will suffer serious side effects. To the patient's surprise, doctors don't have a unanimous opinion on whether the treatment is beneficial or harmful. Thus, the patient decides to ask for several doctors' opinions. The patient collects either positive opinions (+1) in favor of the risky treatment or negative opinions (-1) against the risky treatment. The total sum of evidence is defined as the critical difference, d. But how many opinions should he collect to reduce the risk of making the wrong decision? To help the patient with the decision, his best friend, a statistician, tells him that the number of opinions can be calculated based on the most optimal solution.

The Optimal Stopping Rule for Evidence Accumulation and Deviations

The determination of the optimal stopping rule in statistical decision making has been examined in great detail by Wald (1947) and from the Bayesian perspective by Edwards and colleagues (Edwards, 1965). The optimal Bayesian model defines the stopping rule as the minimization of the expected loss, $E(L)$ (De Groot, 1970). The rule prescribed by the optimal model is to continue collecting evidence and to stop only when the expected value of loss is equal to or lower than the expected loss associated with deferring the decision and collecting more evidence.

To calculate the optimal number of doctors the patient should consult, his friend the statistician acquired the conditional distributions of doctors' positive (+) opinions given that the treatment can be either beneficial or harmful, P(+opinion | beneficial treatment), P(+opinion | harmful treatment), and also the prior probabilities of beneficial and harmful treatments, P(beneficial treatment) and P(harmful treatment) (e.g., Edwards, 1965; Schechter, 1988). The statistician used all these probabilities to calculate the so-called posterior odds in favor of the hypothesis that the treatment is beneficial given the evidence acquired from n number of doctors, $\Omega \left(\frac{\text{benefical treatment}}{\text{harmful treatment}} \mid \text{opinion} \right)$. The posterior odds would indicate the best decision for the finite number of collected doctors' opinions, if the costs and payoffs associated with the risky treatment and the expected diagnostic value of a single opinion are considered. Using mathematical software, the statistician got the number 3 as the optimal stopping rule value for that risky decision. This means that the patient should collect positive and negative doctors' opinion (+1s and -1s) as long as their cumulative sum (d) is lower than the value of d=+3 or higher than the value d=-3. The patient should stop evidence collection and make a decision as soon as d=3, in which case the patient should accept the risky treatment, or d=-3, in which case the patient should reject the risky treatment (e.g., Schechter, 1988).

The relevant literature has revealed that humans do not use the optimal stopping rule. (1) In a deferred decision task in which subjects had the option to defer their decision until they had purchased new information, subjects bought either too little evidence (Phillips & Edwards, 1966; Pitz, 1968) or too much evidence (Pitz, 1968) compared to the optimal model's predictions. (2) The critical difference value d can change over the course of sampling evidence in a single trial (e.g., Busemeyer & Rapoport, 1988; Pitz, 1968; Newell, 2005). Subjects tended to make final decisions on smaller critical difference values for larger sets of evidence. To

account for these results, the optimal model should adjust the critical difference value such that it decreases as more evidence is acquired (Pitz, 1968; Viviani, 1979). (3) Subjects frequently terminated evidence collection when the critical difference value was zero (d=0; Pitz, 1968; Pitz, Reinhold, & Geller, 1969). From the optimal Bayesian viewpoint, this means that decision makers made a final decision even though there was no evidence to support any decision. (4) It has also been shown that human decision makers sometimes stop on a nondiagnostic sequence of evidence (Busemeyer & Rapoport, 1988). For example, after a series of three positive pieces of evidence the subjects stopped on a negative piece of evidence, {+, +, +, -}, and made a decision that supported the positive evidence. Note that the last two pieces of evidence were nondiagnostic and stopping on such a pattern of evidence is logically inconsistent with the optimal model.

The optimal approach to decision making has suffered more general criticism. The optimal model can be successfully applied only when a decision maker possesses perfect knowledge of all aspects of a situation. Following Savage (1954) and Binmore (2009), perfect knowledge of an environment is possible if one resides in a so-called small world. Examples of a small world are a controlled laboratory experiment, a lottery, and certain games. In a small world a detailed statistical representation of the environment exists and an optimal model can predict the exact amount of evidence needed to be collected to find the optimal stopping value.

But most decision makers live in a large world. A large world is quite unpredictable and dynamic—it is constantly changing and it is almost impossible to form an exact statistical representation of such an environment. In a large world a decision maker has limited time to make decisions, possesses limited cognitive powers in terms of memory and attention, and usually acts inconsistently (Berg, Biele, & Gigerenzer, 2008; Gigerenzer, 2008; Schooler & Hertwig, 2005; Shanteau, 1992; Tversky & Kahneman, 1974). It is unrealistic to expect that a decision maker living in a large world would be able to employ an optimal model to determine when to stop accumulating evidence. Alternative approaches have been aimed at exploring how to make effective decisions with a limited amount of information and a limited cognitive system.

Bounded Rationality and Nonoptimal Stopping-Rule Models

According to the bounded rationality approach, making decisions involves simple decision strategies and shortcuts that allow for quick and effortless decisions (e.g., Gigerenzer, 2004). Boundedly rational models require neither exact statistical representation of the environment nor optimization. (For a review of different nonoptimal models for evidence collection see Busemeyer & Rapoport, 1988; for examples see Fifić, Little, & Nosofsky, 2010).

Boundedly rational models for determining stopping rules are more suited to real-life decision-making problems and cognitive limitations than is the optimal model. Let us return to our patient example. The patient started to question the optimal value d=3 after he learned that the conditional distributions used to estimate the doctors' diagnostic accuracies do not exist for his country. Instead, his friend the statistician used the data from another, much smaller country across the ocean. Not trusting the optimal solution (d=3), the patient decided to use another rule. He decided to obtain five doctors' opinions and make his decision based on the majority. This is defined as *the fixed-sample-size stopping rule* (s=5 in the example). A decision maker determines a fixed amount of evidence to be collected before the collection starts. Our patient may have used a five-opinion stopping rule before—years ago when he bought a car. Alternatively, the patient could rely on another useful cue—a streak of either positive or negative opinions. The patient could stop looking for more opinions after receiving three successive positive or negative doctor opinions (r=3) and make a decision accordingly. This is defined as the *runs stopping rule* (cf. Audley & Pike, 1965; Estes, 1960). In sports games the runs rule is also known as the hot or cold hand rule (Bar-Eli, Avugos, & Raab, 2006; Gilovich, Vallone, & Tversky, 1985; Wilke & Barrett, 2009). A player who scores a streak of shots in a row is perceived to be "hot" and is a preferred shooter. A player who has a streak of misses is likewise perceived to be "cold."

Although boundedly rational models have been able to explain some observed deviations from the optimal predictions (for details see Busemeyer & Rapoport, 1988), no single such model has been able to account for them all. Take, for example, the fixed-sample-size stopping rule, which can account for the finding that decision makers sometimes stop on a nondiagnostic sequence of evidence. This rule predicts that the probability of termination should be equal for nondiagnostic sequences of identical length. In contrast, it has been observed that subjects prefer some nondiagnostic sequences over others of the same length (Busemeyer & Rapoport, 1988). The runs stopping rule can account for the finding that decision makers stop on d=0, for example {+,+,-,-}. To stop on that evidence, the stopping rule value for the negative evidence has to be set on two pieces of negative evidence (r= -2). The stopping rule for positive evidence has to be set on a value larger than two pieces of positive evidence (say r=+3). However, the runs stopping rule has limited explanatory power (Busemeyer & Rapoport, 1988). For example, it cannot explain stopping when streaks of evidence are missing. In general, more explanatory power is gained by combining several stopping rules (see Pitz et al., 1969) within one framework. We lack a systematic theory to tie together different stopping rules in a single framework for decision making. To remedy this theoretical gap, I propose the stopping rule selection (SRS) theory.

2

The SRS Theory

The SRS theory provides the basis for a general approach to decision-making operations. This theory is consistent with the idea of a boundedly rational decision maker who utilizes simple decision rules in real time. In different environments, a decision maker acts adaptively, constantly looking for the best decision strategies, stopping rules, and critical values.

A formal description of the SRS theory and proposed stopping rules.

The SRS theory aims to provide a unifying framework for the storage and retrieval of multiple stopping rules. It consists of three hypotheses.

Hypothesis 1: Multiple stopping rules. The SRS theory assumes that several different stopping rules can operate concurrently. Decision makers act adaptively to changes in the environment, not only by calibrating different *stopping rule values* (value criterion) but also by switching between different stopping rules if needed. In real life, multiple stopping rules can be combined in a complex fashion (e.g., Pitz et al., 1969). Take, for example, scoring in tennis: The winner of a tennis game is the player whose score is at least two points higher than the opponent's ($d \geq 2$) and if at least four points have been won so far ($s \geq 4$).

Hypothesis 2: Storage for stopping rules—the decision operative space (DOS). A major component of the SRS theory is a storage place for the stopping rules and their values, which is called the decision operative space (DOS). The DOS can be seen as a variant of an "adaptive toolbox," a collection of domain-specific specialized cognitive mechanisms for decision making built through evolution (Gigerenzer & Todd, 1999; Payne, Bettman, & Johnson, 1993; Todd, 1999). Unlike the toolbox concept, the DOS is conceptualized as a structured psychological space. The stopping rules stored in the DOS are sorted on two dimensions: the *cognitive effort* needed for a certain stopping rule, and the *time* needed to make a decision using a certain stopping rule (Figure 1A). Depending on the environment, a decision maker can use these two dimensions to estimate which decision tools are the most appropriate to use.

The time scale, on the *x*-axis, is defined as chronological time. The exact expected duration of each stopping rule can be calculated from an analytic expression (e.g., see Feller, 1957, p. 317; also Busemeyer & Rapoport, 1988; Pitz, 1968; Pitz et al., 1969). Cognitive effort, on the *y*-axis, is defined as the processing complexity of a decision strategy and can be measured by the number of elementary information processes (EIPs, after Payne et al., 1993) engaged in making a decision. As shown in Figure 1A, each point in the DOS represents a stopping rule with a certain stopping value. Stopping values belonging to the same

stopping rule lie on one line: For the runs stopping rule it is *r*, for the critical difference rule, *d*, and for the fixed-sample-size rule, *s*. Overall decision accuracy increases as one chooses as one chooses larger values for the stopping rules. However, the price of improvement is increases in both time and cognitive effort. As depicted in Figure 1, two stopping rules—the critical difference and the fixed-sample-size—are estimated to be of approximately the same complexity. They share the same EIPs, which are counting, differencing, averaging, and memory engagement. They differ on the time needed to complete the operations. The critical difference stopping rule needs more time to finish than the fixed-sample-size rule, for the same critical value. The runs stopping rule uses EIPs that are far simpler than those used by the previous two. To detect runs, a decision maker has only to count evidence, with minimal memory. Although based on simple EIPs, the runs stopping rule requires considerably more waiting time for larger critical values of runs.
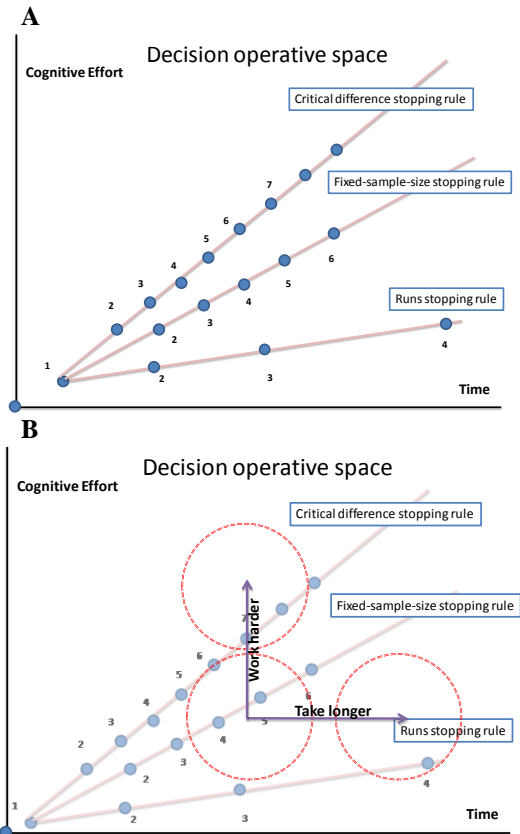

Figure 1: (A) The decision operative space (DOS) for three stopping rules. Each point represents a single stopping rule with a stopping value. A straight line connects the same stopping rule with different stopping values. (B) A cast-net retrieval from the DOS. Dotted circles represent three different cast nets.

3

Hypothesis 3: Retrieval of the stopping rules. A retrieval mechanisms called "cast-net" retrieval is proposed. (cf satisficing approach; Todd, 1997; Todd & Miller, 1999).

Cast-net retrieval. Selection of stopping rules resembles throwing a cast net and catching fish. A decision maker acts much like a fisherman, casting a net into the operative space. Here, on each throw the catch is a subset of possible stopping rules. To behave adaptively in different environments, decision makers adjust the location in the DOS where the net will be cast, and the size of the net. A decision maker who is not familiar with the environment or encounters much uncertainty in evidence collection may cast a larger net. If familiar with the environment, the decision maker may throw a smaller net. The larger the net is, the more different stopping rules are collected to make a single decision. The SRS theory specifies how several stopping rules could be used simultaneously to make a final decision.

The second property of the cast-net retrieval approach is the *double tradeoff*. Depending on where stopping rules are retrieved from the DOS, a decision maker may choose to trade off speed and accuracy (cf. Diederich, 2003; Kocher & Sutter, 2006; Payne et al., 1993) or cognitive effort and accuracy (Payne et al., 1993). Figure 1B shows examples of both tradeoffs. Three cast-net locations are marked by red circles. Moving upward from the lower left circle on the vertical "work harder" path indicates a cognitive effort–accuracy tradeoff, keeping the time value constant. A decision to move vertically in the DOS means choosing to sacrifice frugality of effort to achieve better accuracy. A decision maker works harder to improve overall decision accuracy, as mainly the critical difference stopping rule is sampled. Moving from the lower left circle on the horizontal "take longer" path indicates a speed–accuracy tradeoff, keeping the cognitive effort value constant. A decision to move horizontally means choosing to sacrifice speed to achieve better accuracy. A decision maker takes longer, as mainly the runs stopping rule is sampled. The two tradeoffs can be used to explain adaptive decision making. Under the condition of increased uncertainty, it is expected that a decision maker would increase cognitive effort, and take the "work harder" path. Under time pressure, it is expected that a decision maker would use less time-consuming stopping rules and follow the "take longer" path.

The SRS Theory: A Walkthrough of the Decision Process

In this section I provide a walkthrough of the decision process behind the SRS theory using the cast-net retrieval approach. The SRS model has two stages. The first stage is characterized by the selection and retrieval of stopping rules and their stopping values. The second is characterized by sequential evidence collection and application of stopping-rule criteria. The process is broken into six steps, three in the first stage and three in the second.

Step 1: **Select hypotheses**. Depending on the decision problem, a decision maker chooses the choice hypotheses (e.g., Thomas, Dougherty, Sprenger, & Harbison, 2008). For example, in the patient decision situation described above, the two hypotheses H_1 and H_2 could be about the risky treatment: H_1: The risky treatment is a beneficial procedure, and H_2: The risky treatment is a harmful procedure. This stage is not under the scrutiny of the SRS model.

Step 2: **Cast a net**. The plethora of stopping rules and their values presents a challenge for the selection process. To select a subset of the stopping rules and their values, a decision maker throws a cast net into the DOS. To determine the position of the cast net and its span, a decision maker estimates how much time and cognitive effort can be invested in making the decision (on time and cognitive effort dimensions). These position estimates can be influenced by knowledge the decision maker possesses about this particular environment or similar ones. If no knowledge is available then a random starting point can be chosen. For illustration, assume that the following set of rules determines the cast $\{r=1, r=2, s=2, s=3, d=3, d=2\}$.

Step 3: Select a stopping rule. Once the DOS has been reduced by casting a net, several stopping rules and their values are randomly sampled from the net. All stopping rules and their values contained within the net can be retrieved with the same probability, defined by the probability density function $f(x) = \frac{1}{\text{number of retrieved rules}}$. For example, a decision maker could select the following set of stopping rules and their values from the cast net: $\{r=2, s=2, d=3\}$. Alternatively the probability of retrieving a certain rule from the cast net can be described by the bivariate normal distribution, $\mathbf{x} \sim (N(\boldsymbol{\mu}, \boldsymbol{\Sigma}))$ (where the bold symbols are vectors), allowing rules that are closer to the center of a net to be retrieved with a higher probability than rules that are caught around the edges of the net.

Step 4: Collect evidence. The second stage starts with evidence accumulation. This step is repeated until a decision is made.

Step 5: Check stopping rule. The SRS model tests whether the evidence accumulated so far meets one of the criteria of the stopping rule selected from the net in Step 3. Assume that the model performs a serial test across three selected stopping rules. If none of the criteria have been met the decision maker looks for more evidence and repeats from Step 4. If any of the stopping value criteria are met, the decision maker stops evidence collection and proceeds with making the final decision.

Step 6: Stop and make a decision according to the hypothesis that was supported by the evidence.

Face Validity of the SRS Theory: Preliminary Work and Results of Fitting

To establish face validity, I fit the SRS model to challenging data sets published in two separate studies on

4

determining stopping rules (Busemeyer & Rapoport, 1988; Pitz, 1968). Our preliminary work showed that the SRS computational model can provide an excellent account of reported human data patterns. It is able to account for between 93% and 100% of the variability of Pitz's (1968) data and for about 86% of observed evidence patterns in Busemeyer and Rapoport's (1988) data. The model has 6 parameters describing the "span" of the cast net. Two parameters for each stopping rule define the range of the stopping rule values captured. As a part of the future exploration is the goal to reduce the number of parameters to only three describing the location and the size of the net.

In addition to showing high fitting accuracy, the SRS model was able to account for all four findings that falsified the optimal approach, described above: (1) People bought too much or too little evidence (Pitz, 1968); (2) the value of the critical difference (d) could change over the course of sampling evidence in a single trial (e.g., Busemeyer & Rapoport, 1988; Pitz, 1968); (3) people terminated evidence collection when the critical difference was zero (d=0; Pitz et al., 1969); and (4) people stopped on nondiagnostic patterns. Regarding the accumulation of evidence, the observed data depart from the optimal model predictions (Table 1): For smaller values of d, the subjects collected too much evidence; for larger values of d, the subjects collected too little evidence. The SRS model captures this observed data trend as shown in the SRS model-fitting data. Regarding the value of the critical difference (d), as can be seen in Table 1, less evidence was needed for larger values of d to terminate evidence collection, compared to the optimal model prediction. This trend is accounted for by the SRS model fit. Regarding the termination of evidence collection when the critical difference was zero (d=0), again as seen in Table 1, the SRS model shows that n>0 for d=0. Finally, regarding stopping on nondiagnostic patterns, the SRS model can also predict the nondiagnostic sequence of evidence (see Table 2). The SRS model fitted the observed patterns {1,1,1,0} and {0,0,0,1} (see Table 2; remember that 1 stands for positive and 0 for negative evidence). Note that the last two pieces of evidence in each pattern provide the nondiagnostic information for the optimal model.

Table 1: The average number of pieces of evidence (n, shown in the table's cells) collected as a function of critical difference (d) for three source reliability values (p=.8, .7, and .6). The observed column shows averaged observed human data (from Pitz, 1968). The SRS column shows the best fit values when the stopping rule selection (SRS) model is fitted to the observed data. The optimal column shows the n values predicted by the optimal model. The r^2 values are the proportions of explained variability the SRS model can account for.

d	Source reliability p=.8 r^2=1		
	Observed	SRS	Optimal
0	2.73	2.71	0
1	2.75	2.8	1
2	3	2.92	2.93
3	3.67	3.59	4.71
4	5.04	5	6.41

d	Source reliability p=.7 r^2=0.98		
	Observed	SRS	Optimal
0	3.56	3.92	0
1	3.42	3.65	1
2	4.47	4.21	3.43
3	6.07	6	6.13
4	6.64	6.53	8.86

d	Source reliability p=.6 r^2=0.93		
	Observed	SRS	Optimal
0	3.05	3.89	0
1	4.43	4.51	1
2	5.2	4.75	3.84
3	4.74	5	8.05
4	7.12	6.86	13.37

Table 2: The results of the SRS model fit to Busemeyer and Rapoport (1988) data, from the constant cost condition of their Experiment 2. Table shows the matching patterns correctly recognized by the SRS model, as well as the nonmatching patterns. Evidence refers to the observed patterns of evidence prior decision making, where "1" and "0" stand for positive and negative opinions (recommendations). Response accuracy refers to whether the final decision based on collected evidence was correct. Observed refers to the observed proportion of each pattern. SRS fit refers to the best fitted proportions by the SRS model.

Evidence	Response accuracy	Observed	SRS fit
Observed matched patterns			
{1, 1}	Correct	0.06	0.1
{0, 0}	Correct	0.07	0.1
{1, 1, 1}	Correct	0.19	0.17
{0, 0 ,0}	Correct	0.18	0.16
{1, 0, 1, 1}	Correct	0.05	0.04
{0, 1, 1, 1}	Correct	0.05	0.04
{1, 1, 1, 1}	Correct	0.08	0.07
{1, 1, 1, 0}	Correct	0.001	0.01
{1, 1, 0, 1}	Correct	0.05	0.03
{1, 1, 0, 0}	Incorrect	0.001	0.01
{1, 0, 0, 0}	Correct	0.07	0.04

5

Evidence	Response accuracy	Observed	SRS fit
{0, 0 ,0 ,0}	Correct	0.06	0.07
{0, 1, 0, 0}	Correct	0.06	0.04
{0, 0 , 1 ,0}	Correct	0.05	0.03
{0, 0 , 0, 1}	Correct	0.01	0.01
Observed nonmatched patterns			
{0, 0, 1}	Incorrect	0.002388	0
{0, 1, 1}	Correct	0.009817	0
{1, 0, 0}	Correct	0.002786	0

Acknowledgments

References

Audley, R. J., & Pike, A. R. (1965). Some alternative stochastic models of choice. *British Journal of Mathematical & Statistical Psychology*, *18*, 207–225.

Bar-Eli, M., Avugos, S., & Raab, M. (2006). Twenty years of "hot hand" research: Review and critique. *Psychology of Sport and Exercise*, *7*, 525–553.

Berg, N., Biele, G., & Gigerenzer, G. (2008). Consistency versus accuracy of beliefs: Economists surveyed about PSA. [Meeting Abstract]. *International Journal of Psychology, 43*, 31.

Binmore, K. (2009). *Rational decisions*. Princeton, NJ: Princeton University Press.

Busemeyer, J. R., & Rapoport, A. (1988). Psychological models of deferred decision making. *Journal of Mathematical Psychology*, *32*, 91–134.

DeGroot, M. H. (1970). *Optimal statistical decisions*. New York, NY: McGraw-Hill.

Diederich, A. (2003). MDFT account of decision making under time pressure. *Psychonomic Bulletin & Review*, *10*, 157–166.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology*, *2*, 312–329.

Estes, W. K. (1960). A random walk model for choice behavior. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences* (pp. 265–276). Stanford, CA: Stanford University Press.

Feller, W. (1957). *An introduction to probability theory and its applications* (2nd ed.). New York, NY: Wiley.

Fific, M., Little, D. R., & Nosofsky, R. M. (2010). Logical-rule models of classification response times: A synthesis of mental-architecture, random-walk, and decision-bound approaches. *Psychological Review*, *117*, 309–348.

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 62–88). Oxford, UK: Blackwell.

Gigerenzer G. (2007). *Gut feelings: The intelligence of the unconscious*. New York, NY: Viking.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, *3*, 20–29.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482.

Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (pp. 3–34). New York, NY: Oxford University Press.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*, 295–314.

Kocher, M. G., & Sutter, M. (2006). Time is money—Time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization*, *61*, 375–392.

Newell, B. R. (2005). Re-visions of rationality? *Trends in Cognitive Sciences*, 9, 11–15.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, UK: Cambridge University Press.

Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, *72*, 346–357.

Pitz, G. F. (1968). Information seeking when available information is limited. *Journal of Experimental Psychology*, *76*, 25–34.

Pitz, G. F., Reinhold, H., & Geller, E. S. (1969). Strategies of information seeking in deferred decision making. *Organizational Behavior and Human Performance*, *4*, 1–19.

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.

Schechter, C. B. (1988). Sequential analysis in a Bayesian model of diastolic blood pressure measurement. *Medical Decision Making*, *8*, 191–196.

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, *112*, 610–628.

Shanteau, J. (1992). How much information does an expert use? Is it relevant? *Acta Psychologica*, *81*, 75–86.

Thomas, R. P., Dougherty, M. R., Sprenger, A. M., & Harbison, J. I. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review*, *115*, 155–185.

Todd, P. M. (1997). Searching for the next best mate. In R. Conte, R. Hegselmann, & P. Terna (Eds.), *Simulating social phenomena* (pp. 419–436). Berlin, Germany: Springer.

Todd, P. M. (1999). Simple inference heuristics versus complex decision machines. *Minds and Machines*, *9*, 461–477.

Todd, P. M., & Miller, G. F. (1999). From pride and prejudice to persuasion: Satisficing in mate search. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple Heuristics that make us smart* (pp. 287–308). New York, NY: Oxford University Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.

Viviani, P. (1979). Choice reaction times for temporal numerosity. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 157–167.

Wald, A. (1947). *Sequential analysis*. New York, NY: Wiley.

Wilke, A., & Barrett, H. C. (2009). The hot hand phenomenon as a cognitive adaption to clumped resources. *Evolution and Human Behavior*, *30*, 161–169.

6