

A Signal-Detection Analysis of Fast-and-Frugal Trees

Shenghua Luan

Singapore Management University and Max Planck
Institute for Human Development

Lael J. Schooler and Gerd Gigerenzer

Max Planck Institute for Human Development

Models of decision making are distinguished by those that aim for an optimal solution in a world that is precisely specified by a set of assumptions (a so-called “small world”) and those that aim for a simple but satisfactory solution in an uncertain world where the assumptions of optimization models may not be met (a so-called “large world”). Few connections have been drawn between these 2 families of models. In this study, the authors show how psychological concepts originating in the classic signal-detection theory (SDT), a small-world approach to decision making, can be used to understand the workings of a class of simple models known as fast-and-frugal trees (FFTs). Results indicate that (a) the setting of the subjective decision criterion in SDT corresponds directly to the choice of exit structure in an FFT; (b) the sensitivity of an FFT (measured in d') is reflected by the order of cues searched and the properties of cues in an FFT, including the mean and variance of cues' individual d' s, the intercue correlation, and the number of cues; and (c) compared with the ideal and the optimal sequential sampling models in SDT and a majority model with an information search component, FFTs are extremely frugal (i.e., do not search for much cue information), highly robust, and well adapted to the payoff structure of a task. These findings demonstrate the potential of theory integration in understanding the common underlying psychological structures of apparently disparate theories of cognition.

Keywords: signal-detection theory, fast-and-frugal trees, theory integration

Supplemental materials: <http://dx.doi.org/10.1037/a0022684.supp>

In recollecting September 11, 2001, Louis Cook of the Emergency Medical Services division of the New York City Fire Department noted how the triage system Simple Triage and Rapid Treatment (START) helped his team prioritize victims and identify the ones who needed help the most (Cook, 2001). START classifies the injured into two major categories: those who need medical treatment immediately and those whose treatment can be delayed. When employing START, which is illustrated in Figure 1, a paramedic sequentially checks up to five diagnostic cues to decide which category a person falls into; a decision can be made after each cue is checked. In essence, START is a decision tree with a very simple and transparent structure. Using this type of tree, a person does not need to search for and integrate all the relevant information to reach a decision; thus, a decision can be quickly made with little effort. Such trees are designed to help people make decisions in real settings, potentially achieving a high level of

decision accuracy under the constraints of limited information, time, and resources. In the terminology of Leonard J. Savage (1954) and Ken Binmore (2009), they are explicitly designed to deal with “large world” problems.

Following Savage (1954), the father of modern Bayesian decision theory, we distinguish between the concepts of small and large worlds. Savage's prototypical example of a small world is a lottery, the stock-in-trade of many decision-making studies. In a small world, a person is assumed to have perfect knowledge of all aspects of the situation, including all the states in which he or she could be (e.g., winning or losing), all the acts he or she could perform (e.g., play Gamble A or B), every possible consequence as a result of each act and state (e.g., gain \$6 if Gamble A is played and won), and the probability of each consequence (e.g., 40% of winning \$6 under Gamble A). There are no surprises in a small world, so an optimization calculus such as Bayes's rule can be applied to define the best or “rational” action. A person is further assumed to have abundant time and resources to determine the best action by processing all relevant information.

Louis Cook's paramedic team, in contrast, had to deal with a large world that did not provide full information. In a large world, one may have only limited knowledge about the decision being faced, contingencies and probabilities may change over time, unexpected consequences or surprises may happen, only limited time may be available to contemplate the decision, and there may not exist a single criterion to evaluate decision options so that one of them could be hailed as the best. Savage noted that it is always possible to look before you leap in a small world, but we live in a large world, where it is exceedingly difficult to optimize anything

This article was published Online First March 7, 2011.

Shenghua Luan, School of Social Sciences, Singapore Management University, and Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany; Lael J. Schooler and Gerd Gigerenzer, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development.

We thank Jonathan Evans, Mario Fific, Kevin Gluck, Konstantinos Katsikopoulos, Michael Lee, Björn Meder, Jonathan Nelson, Jan Woike, and members of the ABC Research Group for their helpful comments, and Anita Todd and Rona Unrau for editing the article.

Correspondence concerning this article should be addressed to Shenghua Luan, Level 4, 90 Stamford Road, School of Social Sciences, Singapore Management University, Singapore, 178903. E-mail: shluan@smu.edu.sg

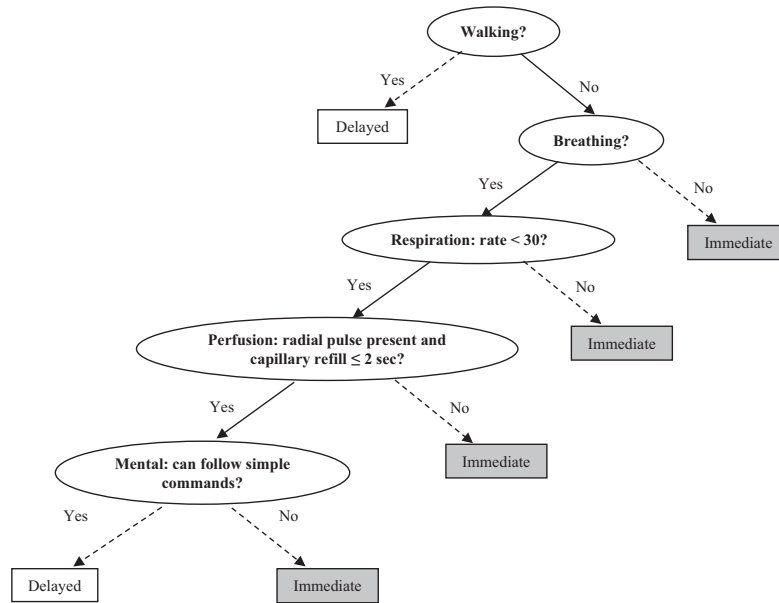


Figure 1. A schematic version of the Simple Triage and Rapid Treatment (START) procedure, which categorizes patients into those who need immediate medical treatment and those whose treatment can be delayed (Super, 1984).

with even ordinary complexity, so we are well advised to live by the adage, “We can cross that bridge when we come to it.”

Savage believed that one could not take optimization models designed for small worlds as models of how people make decisions in large worlds, “because the task implied in making such a decision is not even remotely resembled by human possibility” (Savage, 1954, p. 16). This view serves to highlight the contrast between the one half of his book that is devoted to developing Bayesian decision theory for small-world problems with the other half on heuristics for decision making in large worlds. Today, psychological theorizing is similarly split into the study of how people’s decisions can be described by complex optimization models that operate in small worlds and how people use simple heuristics in large worlds, with scant interaction between the two approaches (but see Hogarth & Karelaia, 2007, and Pleskac, 2007, for exceptions). Does this split imply that we must seek the one true path?

Our answer is no. Instead, we propose dealing with this schism by starting with the following question: Can we better understand heuristic decision making in large worlds with lessons learned from small-world theories? This article is our attempt at an answer. We apply well-studied concepts and methods associated with one small-world theory, *signal-detection theory* (SDT), to understand the nature of a class of large-world models, namely fast-and-frugal trees, of which START is an example. At a general level, this article contributes to what we believe to be a crucial issue in psychological science, the integration of theories. We start by briefly introducing SDT and fast-and-frugal trees.

Signal-Detection Theory

SDT, originating from the statistical theory of Neyman–Pearson hypothesis testing (Gigerenzer, 1991; Gigerenzer & Murray, 1987,

Chapter 2), was first applied in psychology to study perception and sensation (Tanner & Swets, 1954). However, over the years, SDT has been applied in many other areas of psychology, such as categorization (e.g., Ashby & Maddox, 1992), memory (e.g., Wixted & Stretch, 2004), confidence judgment (e.g., Wallsten & Gonzalez-Vallejo, 1994), eyewitness testimony (e.g., Birnbaum, 1983), group decision making (e.g., Sorkin, Luan, & Itzkowitz, 2004), reasoning (Dube, Rotello, & Heit, 2010), and mate choice (e.g., Haselton & Buss, 2000). The most important contribution of SDT to the study of human cognition is distinguishing between sensitivity and decision bias and developing methods to measure the two separately.

The upper part of Figure 2 illustrates the main assumptions and concepts of SDT as it is applied to a binary decision task,¹ and Table 1 lists details of some commonly used SDT terms. The fundamental assumption of SDT is that the two types of events, traditionally called *signal* and *noise*, have overlapping distributions on an observation scale X . After setting a decision criterion on this scale (x_c , as in Figure 2), four decision outcomes are possible: hit, false alarm, miss, and correct rejection. Two measures, d' and c , are popularly adopted in SDT to quantify sensi-

¹ Such binary decision tasks are traditionally referred to as *yes–no tasks* in SDT. In a yes–no task, a stimulus is displayed and the person identifies whether it is a member of one particular category or not. It should be distinguished from a two-alternative forced-choice or *2AFC task*, in which two stimuli, each from a different category, are presented, and the person indicates which of them is a member of a category of interest.

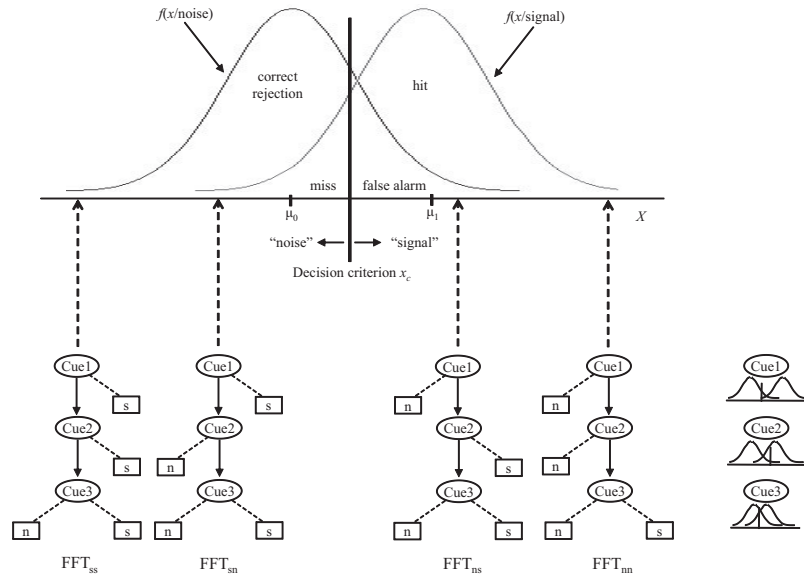


Figure 2. Where is the decision criterion in a fast-and-frugal tree (FFT)? The upper part of the figure illustrates the main assumptions and concepts of signal-detection theory (SDT) in a binary decision task, and the lower part illustrates the four possible FFTs that can be constructed when three cues are searched in a set order. Based on the decisions pointed to by the first two exits, the trees are named from left to right FFT_{ss}, FFT_{sn}, FFT_{ns}, and FFT_{nn}. The arrows connecting the figure parts indicate the rough locations of the four FFTs' decision criteria when they are used to make a binary s/n (for signal and noise, respectively) decision. Among the four, FFT_{ss} has the most liberal decision criterion, and FFT_{nn} has the most conservative one. The decision criteria of FFT_{sn} and FFT_{ns} are less extreme than the other two, with FFT_{sn} being more liberal than FFT_{ns}. The two overlapping normal distributions next to each cue illustrate SDT's assumption about how object values are distributed on a cue and emphasize that each cue comes with its own sensitivity and decision criterion (in *d'* and *c*, respectively; see Table 1).

tivity² (or true decision accuracy) and decision bias, respectively. In an empirical setting, with the knowledge of the frequency of hits and the frequency of false alarms—*P*(Hit) and *P*(FA), respectively—*d'* and *c* can be calculated using the equations in Table 1. More detailed introductions to SDT can be found in Macmillan and Creelman (2005), and Stanislaw and Todorov (1999).

When decision outcomes differ greatly in their consequences, the main concern of a decision maker often turns from simply making an accurate decision to instead making one with high expected value. To maximize expected value, according to Green and Swets (1966), decision makers should base their decisions on the likelihood ratio of an observation— $LR(x) = f(x/signal)/f(x/noise)$ —and adopt a decision criterion based on the likelihood ratio called $\beta_{optimal}$. The equation for calculating $\beta_{optimal}$ (see Table 1) shows that a decision criterion should be set by factors that do not depend on decision sensitivity (*d'*), and depending on the specific values of those factors, it is possible that a biased decision criterion ($\beta_{optimal} \neq 1, c \neq 0$) fits better to a task than an unbiased one ($\beta_{optimal} = 1, c = 0$).

Besides distinguishing between sensitivity and bias and formulating the optimal decision criterion in the processing of information from one source or cue, SDT also contributes to the study of decision making with multiple sources or cues. According to SDT, achieving the best possible sensitivity with *m* cues requires two steps. The first is integration: Observations from all cues ($x_1, x_2 \dots x_m$) need to be converted to a joint likelihood ratio statistic,

$$LR(x_i) = LR(x_1, x_2 \dots x_m) = \frac{f(x_1, x_2 \dots x_m/signal)}{f(x_1, x_2 \dots x_m/noise)}$$

that represents the integrated information from those cues. The second step is decision making: A decision is made by comparing the joint likelihood ratio $LR(x_i)$ with $\beta_{optimal}$. This two-stage process assumes that one possesses full knowledge of the joint distributions of the two decision events, information from all cues is available, no error occurs in calculating $LR(x_i)$, and the value of $\beta_{optimal}$ can be precisely estimated. Because these assumptions are unlikely to be met in realistic settings, the process is often portrayed as ideal, resulting in ideal performance.

One critical assumption of the ideal, or the ideal observer, model is that information from all cues can be obtained, which is often not the case. When information must be sampled in sequence and at a cost, it is desirable to implement a stopping rule to limit information search. Following Wald's generalization of Neyman-Pearson theory to sequential decisions with stopping rules (Wald, 1947), a class of sequential sampling models with Bayesian information updating mechanisms has been developed (e.g., Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Bussemeyer & Rapoport, 1988; Bussemeyer & Townsend, 1993; Edwards, 1965; Vickers, 1979). Figure 3 illustrates a general sequential sampling model

² In some fields, especially medicine, the term *sensitivity* refers to the hit rate and *specificity* refers to the correct rejection rate.

Table 1
List of Commonly Used Signal-Detection Theory Terms, Their Definitions/Measurements, and Some Explanatory Notes

Term	Definition/measurement	Notes ^a
Hit	Respond “signal” given signal	Complementary to miss
False alarm (FA)	Respond “signal” given noise	Complementary to correct rejection (CR)
d'	$z_{Hit} - z_{FA}$	z_{Hit} and z_{FA} are the z scores of the hit rate and the FA rate, respectively; d' reflects the standardized distance between the signal (S) and noise (N) distributions
c	$-0.5 \times (z_{Hit} + z_{FA})$	$c > 0$: <i>conservative</i> bias, making more noise than signal decisions relative to prior probabilities $c < 0$: <i>liberal</i> bias, making more signal than noise decisions relative to prior probabilities $c = 0$: <i>neutral</i> bias, making decisions consistent with prior probabilities
$\beta_{optimal}$	$\frac{V(CR) - V(FA)}{V(Hit) - V(Miss)} \cdot \frac{P(N)}{P(S)}$	$P(N)$ and $P(S)$ are prior probabilities of the two decision events; $V(Hit)$, $V(Miss)$, $V(CR)$, and $V(FA)$ are expected values of the four decision outcomes; with a certain d' : $c_{optimal} = \ln(\beta_{optimal})/d'$.

^a Readers who are not familiar with signal-detection theory may see our online supplementary materials for a more detailed introduction to the theory, including some concrete examples demonstrating the calculations of d' and c .

for a binary decision task. Besides the decision criterion $\beta_{optimal}$, two other criteria exist: β_s for signal and β_n for noise. During the sampling process, whenever the integrated likelihood ratio $LR(x_i)$ exceeds β_s or β_n , search stops and a decision is made. If after searching all cues $LR(x_i)$ still falls between the two criteria, a decision is made by comparing $LR(x_i)$ with $\beta_{optimal}$.

The aim of a normative sequential sampling model is to find criteria (i.e., $\beta_{optimal}$, β_s , and β_n) that maximize the expected value of the decisions at hand. The values of those optimal criteria depend on many factors, including the sensitivity, decision bias, and cost of each cue to be sampled; sampling order of those cues; and all the factors that influence the formation of $\beta_{optimal}$, as shown in Table 1. Moreover, the model assumes that information is integrated in a Bayesian way; that is, the likelihood ratio statistic $LR(x_i)$ is derived based on the distributional characteristics of cues sampled from both the signal and noise distributions and their prior

probabilities. An early application of this modeling approach in SDT can be found in the work of Swets and Green (1961), and a more recent one in Luan, Sorkin, and Itzkowitz (2010).

Compared with the ideal SDT model, an advantage of the optimal sequential sampling (OSS) model is its ability to tackle the problem of information search, which is inherent in almost all large-world tasks. But this does not necessarily validate OSS as a large-world model. On the contrary, the OSS model actually manifests many characteristics of a small-world model because of its demanding knowledge and computational requirements. The OSS model belongs to a broad class of models that optimize under constraints. As pointed out by Gigerenzer and Todd (1999), “the paradoxical approach of optimization under constraints is to model ‘limited search’ by assuming that the mind has essentially unlimited time and knowledge with which to evaluate the costs and benefits of further information search” (p. 11). In contrast, fast-and-frugal trees such as START deal with realistic constraints by sidestepping the process of optimization. In the following, we specify the properties of fast-and-frugal trees and how people use them to solve large-world problems.

Fast-and-Frugal Trees

SDT has been used as a model for both how people make decisions (descriptive) and how they ought to make decisions (prescriptive). Similarly, fast-and-frugal trees (FFT) have also been proposed as descriptive and prescriptive models in large worlds. Before formally defining FFTs, we provide some examples of their applications. START is an FFT that prescribes how paramedics should classify patients when confronting an overwhelming number of casualties. Another FFT, illustrated in Figure 4a, has been used by emergency room doctors to decide whether to send a patient suffering from severe chest pain to a regular nursing bed or the coronary care unit (CCU; Green & Mehr, 1997).

Following this FFT, a doctor should send a patient to the CCU if (a) the patient has ST segment changes (in the ECG) or (b) chest pain is the patient’s chief complaint, plus there is any sign of danger in any one of five other diagnostic cues. Green and Mehr

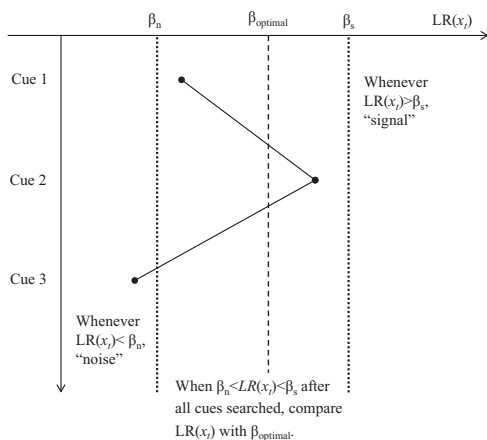


Figure 3. An illustration of a general sequential sampling model in which three cues are searched in a set order. $LR(x_i)$ is the likelihood ratio statistic that results from integrating information from the cues; β_s and β_n are the criteria used to stop search and make a decision in favor of signal or noise, respectively; $\beta_{optimal}$ is the optimal decision criterion used to make a decision if $LR(x_i)$ failed to cross either β_s or β_n after all cues are searched.

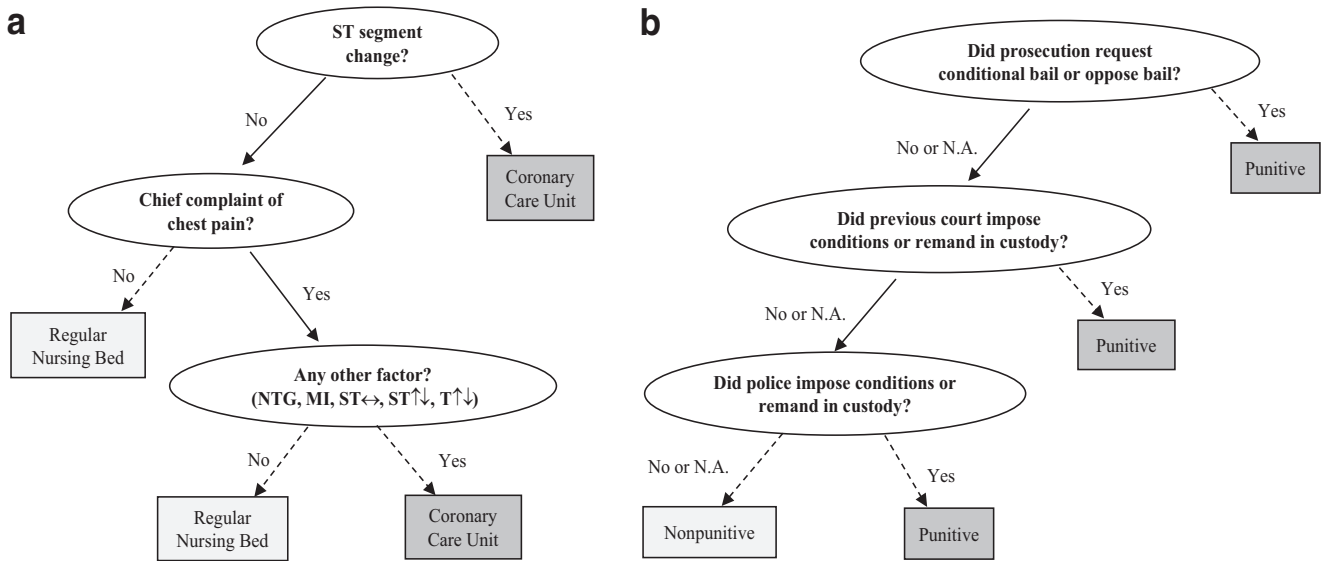


Figure 4. Two examples of fast-and-frugal trees (FFTs) applied to large world problems. The left tree (a) is designed to help emergency room doctors decide whether to send a patient with severe chest pain to the Coronary Care Unit (CCU) or a regular nursing bed (Green & Mehr, 1997). The right tree (b) is a model of how British judges decide whether to make a punitive bail decision (Dhmi, 2003).

(1997) found that, compared with a logistic regression model that uses eight cues simultaneously to make a decision, this FFT had a higher overall predictive accuracy, in addition to its advantages in frugality (i.e., number of cues checked to make a diagnosis) and speed. Moreover, its transparency means that doctors can easily understand how the diagnostic system works, whereas lack of transparency has contributed to doctors' resistance to complex expert systems, such as logistic regression. Medical experts rely on other FFTs as decision aids in clinical practice, such as when performing HIV testing (Gigerenzer, Hoffrage, & Ebert, 1998), performing mammography screening (Welch, 2004, p. 36), and prescribing antibiotics to children (Fischer et al., 2002).

FFTs are not just prescriptive but descriptive models of decision making as well. Figure 4b shows an FFT that judges in London appear to use when deciding whether to make a punitive bail decision (i.e., imprisonment or bail with restrictions) or a nonpunitive one (i.e., unconditional bail). According to Dhmi (2003), there are 25 cues that potentially could be used to make this decision. Based on 342 bail hearings by 116 judges in two London courts, she found that FFTs involving only three cues both fit and predicted the judges' decisions better than a linear model using all 25 cues. Figure 4b shows the tree for one London court (the tree for the other court is identical except for one cue). Moreover, Dhmi and colleagues found that other types of legal decisions made by British judges are similarly well described by FFTs (e.g., Dhmi & Ayton, 2001; Dhmi & Harries, 2001). Additional evidence for FFTs as descriptive models has been reported for decisions based on vignettes about whether to prescribe medication to treat depression (Smith & Gilhooly, 2006) and whether to admit children suffering from asthma to the hospital (Kee et al., 2003), as well as human participants' responses, in terms of both reaction time and accuracy, in classic categorization tasks (Fific, Little, & Nosofsky, 2010).

Tree models of categorization and decision making have been studied in a variety of disciplines, such as medicine, applied statistics, computer science, and psychology (e.g., Breiman, Friedman, Stone, & Olshen, 1984; Busemeyer, Weg, Barkan, Li, & Ma, 2000; Green & Mehr, 1997; Quinlan, 1993). Martignon and colleagues conceptualized FFTs as a class of simple tree models that categorize or make decisions with limited information search (e.g., Martignon, Katsikopoulos, & Woike, 2008; Martignon, Vitouch, Takezawa, & Forster, 2003). Because the lines between categorization and decision making are often murky (Ashby & Berretty, 1997), we frame FFTs in this study as decision models, highlighting the consequences associated with the outcomes of categorization.

In tasks where a binary decision needs to be made (e.g., immediate or delayed treatment as in the triage problem) and there are m cues available for making such a decision, an FFT is defined as follows:

Definition: A fast-and-frugal tree is a decision tree that has $m + 1$ exits, with one exit for each of the first $m - 1$ cues and two exits for the last cue.

An FFT is composed of sequentially ordered cues. To make a decision, an FFT starts by checking an object's value on the first cue. If it meets the exit condition of the cue, which is generally framed as an *if-then* statement (e.g., if a person can walk, then delayed), a decision is made and no other cues need to be checked. Otherwise, an FFT considers the other cues one after another until the exit condition of a cue is met. The last cue of an FFT has two exits, to ensure that a decision will be made in the end. Among all trees that could be constructed from a group of ordered cues, an FFT has the minimal number of exits. In contrast, a full tree has the maximum. For the triage problem to which START is applied (see Figure 1), a full tree

would have 2^5 , or 32, exits, each of which indicates a decision for a unique cue profile (e.g., delayed for patients with a cue profile of [“able to walk,” “able to breathe,” “respiration rate is normal,” “perfusion is normal,” “can follow simple commands”]). In essence, a full tree maps each possible cue profile to a decision but always searches through all available cues.

The structure of a full tree can be simplified. For example, if a person can walk without assistance, the person can surely breathe. Thus, the exits in a full tree that match to cue profiles [“able to walk,” “not able to breathe”] make little practical sense and hence can be removed from consideration or pruned. Because pruned trees often lead to better decision performance than full trees, pruning methods are central to the study of decision trees (e.g., Quinlan, 1993). With respect to structure, an FFT can be viewed as a decision tree whose exits cannot be further pruned, although the construction of an FFT requires no actual pruning, as its number of exits is predefined.

Because of its simple structure, an FFT may possess three advantages over other decision models, including other types of trees. First is its frugality, defined as the number of cues needed to make a decision in a given task (Martignon et al., 2008). Frugality is particularly desirable when information acquisition comes at some cost (e.g., time, effort, money). Because an FFT searches for cues one by one, striving to make a decision after each cue, it can be highly frugal. The second is its simple decision rule. Unlike models that weight and add information, an FFT bases its decisions on the exit condition associated with each cue; thus, implementing an FFT requires knowing only the order in which the cues are checked and their exit conditions. Similar lexicographic decision models (e.g., take the best, elimination by aspects) not only require less mental effort to execute than the weighting-and-adding models (e.g., Mata, Schooler, & Rieskamp, 2007; Payne, Bettman, & Johnson, 1993) but also make the process of decision making more transparent and easier to communicate and understand (e.g., Reilly et al., 2002).

The third advantage of an FFT is its potential for robustness. One important goal of a model is to make accurate predictions based on known data. Complex models with many free parameters can usually fit known data well but may achieve this by fitting spurious relations specific to a sample. As a result, their predictive accuracy in a new sample may suffer (e.g., Myung, Forster, & Browne, 2000). When there is a discrepancy between a model’s fitting and predictive accuracy, the model is said to overfit the data, and the more a model overfits, the less robust it is likely to be. The simplicity of FFTs makes them less susceptible to overfitting and thus more robust. Martignon et al. (2008) offered some supportive evidence by comparing the performances of three models—FFTs, logistic regression, and a complex decision tree named CART—in 30 different tasks. They found that FFTs had the most stable performance across different training sample sizes, and when the training sample sizes were small, FFTs performed nearly as well as the other models in terms of the total number of correct decisions made.

FFTs Through the Lens of SDT

The psychology underlying FFTs appears to have little in common with that of SDT, although both are applied to the same binary decision tasks with multiple cues. According to

SDT, the mind works like an intuitive statistician that sets decision criteria and is characterized by its sensitivity d' . FFTs search through cues sequentially and stop search after the first cue that allows them to do so. From the vantage point of SDT, one may wonder: Where is the decision criterion in an FFT? Where is its sensitivity? In what follows, we apply well-developed methodologies, measures, and concepts of SDT to answer these questions and to gain a better understanding of the psychology underlying FFTs. In Studies 1 and 2, we first ask whether there is a decision criterion in an FFT; if so, how is it reflected in an FFT’s structure, and how can it be adjusted? Second, how is sensitivity implemented in an FFT, and how can it be improved? In Study 3, we compare the decision-making quality of the two theories in terms of sensitivity, frugality, predictive accuracy, and expected value.

Study 1: Where Is the Decision Criterion in an FFT?

The main conceptual contribution of SDT is to disentangle sensitivity (d'), a measure of pure accuracy, from decision criterion (c). It has been widely observed that when making a decision, both sensitivity and decision criterion strongly influence the outcome of a person’s decision (Green & Swets, 1966). Yet there has been no systematic investigation of these two concepts in the research of FFTs. In this study, we show how decision criteria are manifested in the structure of an FFT.

Decision Criterion in the CCU Allocation Tree

To demonstrate the mapping from SDT to FFTs, let us return to the CCU allocation tree illustrated in Figure 4a. On the basis of this tree, physicians assign patients to a regular nursing bed or the special coronary care unit. Maintaining the same cue order as in this tree, four possible FFTs, each with a unique exit structure, can be constructed; they are shown in the lower part of Figure 2. Designating the four FFTs by the decisions associated with their first and second exits, we refer to them from left to right as FFT_{ss}, FFT_{sn}, FFT_{ns}, and FFT_{nn}. If a CCU decision is represented by an “s” exit, then the CCU allocation tree would be designated as an FFT_{sn}. Now the question is: How does the decision criterion of that tree differ from the others?

On the basis of the data used to develop the CCU allocation tree, we calculated the hit and FA rates of the four FFTs, shown in Figure 5. Graphical displays like Figure 5 are often employed in SDT to illustrate the performance of a decision maker by plotting $P(\text{Hit})$ and $P(\text{FA})$, or their transformed z -values $z\text{Hit}$ and $z\text{FA}$, against each other. Such displays illustrate the *receiver operating characteristic* (ROC) of a decision maker, and the $[P(\text{FA}), P(\text{Hit})]$ and $[z\text{FA}, z\text{Hit}]$ spaces are often referred to as the ROC space and the $z\text{ROC}$ space, respectively. Figure 5 shows that there is a systematic shift in decision criteria associated with each of the four FFTs. From FFT_{nn} to FFT_{ss}, it is clear that the decisions made by an FFT become more and more liberal. The bias measures c of the four trees are 0.37, -0.31 , -0.88 , and -1.47 for FFT_{nn}, FFT_{ns},

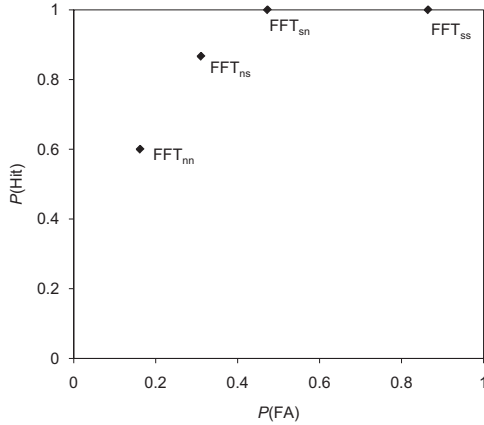


Figure 5. The performance of four fast-and-frugal trees (FFTs) in the receiver operating characteristic (ROC) space. Retaining the same cue order, the four FFTs are composed of the three cues used in the CCU allocation tree (see Figure 4a for an illustration). The CCU allocation tree is of the type FFT_{sn} , in which “s” represents a decision to send a patient to the CCU and “n” a decision to send a patient to a regular nursing bed.

FFT_{sn} , and FFT_{ss} , respectively.³ Thus, it appears that having more “s” exits in an FFT and having these exits earlier in the cue sequence lead to a more liberal bias. However, given the limited data from which the four trees are constructed (sample size = 89), the generality of this result is unclear. To examine whether such a bias shift is systematic, we compared analytically the hit and FA rates of the four FFTs that are composed of three cues ordered in a certain sequence.

Decision Criteria of Three-Cue FFTs

Following the terminology in SDT, an object is assumed to belong to either the signal or the noise category, and there are two corresponding distributions of object values for each cue. If an object is from signal, it has the cue profile $[x_{s1}, x_{s2}, x_{s3}]$, in which x_{si} represents the object’s value on the i th cue; if it is from noise, the cue profile is $[x_{n1}, x_{n2}, x_{n3}]$. Provided that the decision criteria of the three cues are $[x_{c1}, x_{c2}, x_{c3}]$, based on their exit structures, the hit and FA rates of the four FFTs can be expressed as the following:

$$P(\text{Hit})_{\text{FFT}_{\text{ss}}} = P(x_{s1} > x_{c1}) + P[(x_{s1} < x_{c1}) \cap (x_{s2} > x_{c2})] \\ + P[(x_{s1} < x_{c1}) \cap (x_{s2} < x_{c2}) \cap (x_{s3} > x_{c3})]$$

$$P(\text{Hit})_{\text{FFT}_{\text{sn}}} = P(x_{s1} > x_{c1}) \\ + P[(x_{s1} < x_{c1}) \cap (x_{s2} > x_{c2}) \cap (x_{s3} > x_{c3})]$$

$$P(\text{Hit})_{\text{FFT}_{\text{ns}}} = P[(x_{s1} > x_{c1}) \cap (x_{s2} > x_{c2})] \\ + P[(x_{s1} > x_{c1}) \cap (x_{s2} < x_{c2}) \cap (x_{s3} > x_{c3})]$$

$$P(\text{Hit})_{\text{FFT}_{\text{nn}}} = P[(x_{s1} > x_{c1}) \cap (x_{s2} > x_{c2}) \cap (x_{s3} > x_{c3})]$$

$$P(\text{FA})_{\text{FFT}_{\text{ss}}} = P(x_{n1} > x_{c1}) + P[(x_{n1} < x_{c1}) \cap (x_{n2} > x_{c2})] \\ + P[(x_{n1} < x_{c1}) \cap (x_{n2} < x_{c2}) \cap (x_{n3} > x_{c3})]$$

$$P(\text{FA})_{\text{FFT}_{\text{sn}}} = P(x_{n1} > x_{c1}) \\ + P[(x_{n1} < x_{c1}) \cap (x_{n2} > x_{c2}) \cap (x_{n3} > x_{c3})]$$

$$P(\text{FA})_{\text{FFT}_{\text{ns}}} = P[(x_{n1} > x_{c1}) \cap (x_{n2} > x_{c2})] \\ + P[(x_{n1} > x_{c1}) \cap (x_{n2} < x_{c2}) \cap (x_{n3} > x_{c3})]$$

$$P(\text{FA})_{\text{FFT}_{\text{nn}}} = P[(x_{n1} > x_{c1}) \cap (x_{n2} > x_{c2}) \cap (x_{n3} > x_{c3})]$$

Note that the equations for the hit and FA rates under each FFT are the same except for the category from which an object is drawn. The hit (and FA) rate of each FFT is the sum of the hit (and FA) rates contributed by its cues. For an FFT_{sn} , of which the CCU allocation tree is an instance, there are two components in its hit rate: one from the first cue $P(x_{s1} > x_{c1})$, the probability that an “s” decision can be made by this cue, and the other from the third cue $P[(x_{s1} < x_{c1}) \cap (x_{s2} > x_{c2}) \cap (x_{s3} > x_{c3})]$, the probability that this cue makes an “s” decision given that neither of the first two cues can make a decision.

When comparing two decision makers, if both the hit and FA rates of one are higher than those of the other, then the former has a more liberal decision bias than the latter.⁴ Keeping this in mind, let us first compare the hit and FA rates of the two trees FFT_{ss} and FFT_{sn} . Because both FFT_{ss} and FFT_{sn} have an “s” exit on the first cue, the first components of their hit rates are identical. Their second components, however, differ. Particularly, the probability specified in the second component of FFT_{ss} , $P[(x_{s1} < x_{c1}) \cap (x_{s2} > x_{c2})]$, cannot be lower than the probability specified in the second component of FFT_{sn} , $P[(x_{s1} < x_{c1}) \cap (x_{s2} > x_{c2}) \cap (x_{s3} > x_{c3})]$, because the latter is a subset of the former. Given that a third component exists in the hit rate of FFT_{ss} but not in FFT_{sn} , the overall hit rate of FFT_{ss} cannot be lower than that of FFT_{sn} . The same relation holds for the two trees’ FA rates as well.

Now let us examine the hit and FA rates of FFT_{sn} and FFT_{ns} . From the equations of their hit rates, we see that the two components of FFT_{ns} ’s hit rate, $P[(x_{s1} > x_{c1}) \cap (x_{s2} > x_{c2})]$ and $P[(x_{s1} > x_{c1}) \cap (x_{s2} < x_{c2}) \cap (x_{s3} > x_{c3})]$, are actually both subsets of the first component of FFT_{sn} ’s hit rate, $P(x_{s1} > x_{c1})$, and the two are mutually exclusive. Thus, even if the second component of FFT_{ns} ’s hit rate is zero, the overall hit rate of FFT_{sn} cannot be lower than that of FFT_{ns} , nor can its FA rate be lower than that of FFT_{ns} . Following a similar deduction process, it can be shown that the hit and FA rates of FFT_{ns} cannot be lower than those of FFT_{nn} either.

³ When the hit rate is 1, as is the case for both the FFT_{sn} and the FFT_{ss} trees in the CCU allocation example, c and d' cannot be calculated using the standard equation. In keeping with the convention in SDT (e.g., Stanislaw & Todorov, 1999), we used a modified hit rate defined as $(N - 0.5)/N$, where N was the number of signal trials (or patients who really needed to be admitted to the CCU) in the sample. This modified hit rate was then used to calculate the decision criteria (c) and sensitivities (d') of the FFT_{sn} and the FFT_{ss} trees.

⁴ From Table 1, it can be seen that the larger the sum of $z\text{Hit}$ and $z\text{FA}$, the lower the value of c and, hence, the more liberal a decision maker. Because $z(x)$ is a monotonic function of x , if both $P(\text{Hit})$ and $P(\text{FA})$ of a decision maker A are larger than those of another decision maker B, the sum of $z\text{Hit}$ and $z\text{FA}$ of A is also larger than that of B. Thus, A should have a more liberal decision bias than B.

To summarize, both hit and FA rates will decrease successively as we transform the tree from FFT_{ss} to FFT_{sn} , FFT_{ns} , and FFT_{nn} . Hence, its decision criterion will become increasingly conservative. The relative decision criteria of the four FFTs are illustrated in Figure 2. Because FFT_{ss} is the most liberal tree and FFT_{nn} the most conservative, we refer to them as the strictly liberal and strictly conservative trees, respectively. In Figure 2, the trees' decision criteria are projected into two normal distributions of object values as assumed in SDT. Nevertheless, their systematic differences hold regardless of the forms that the signal and noise distributions take.

Decision Criteria of m -Cue Fast-and-Frugal Trees

When there are m cues in an FFT, the following algorithm judges the relative decision bias of a pair of FFTs that are constructed from cues ordered in the same sequence:

Step 1: Check whether the exits on each cue of the two FFTs are the same, starting from Cue 1.

Step 2: Stop at the cue on which the two FFTs have different exits.

Step 3: Infer that the hit and FA rates of the FFT with an "s" exit are no lower than those of the other FFT. Hence, the former has at least as liberal a decision bias as the latter.

The comparison algorithm is implied by a property of FFTs that we call *lexicographic decision bias*, which states that

An FFT of the form $\text{FFT}_{==s**}$ will always be at least as liberal as another FFT of the form $\text{FFT}_{==n**}$, where $==$ represents cues that share the same exits and $**$ represents cues that may or may not have the same exits.

A detailed proof of this property can be found in Appendix A.

Lexicographic decision bias is noncompensatory in nature, meaning that no matter what type of exits the cues have after the first differentiating cue or how many subsequent cues there are, an $\text{FFT}_{==s**}$ will always be at least as liberal as an $\text{FFT}_{==n**}$. For example, suppose that there are two FFTs composed of the same five ranked cues, one being FFT_{snnn} and the other FFT_{nsss} . Although there is only one "s" exit in the first tree but three in the second, the former is still more liberal than the latter, because its sole "s" exit happens to be on the first cue. This counterintuitive result is further confirmed in Study 2, where we compared FFTs with specific cue property values. Knowledge of this lexicographic decision bias property will facilitate constructing an FFT with a particular level of decision bias.

Summary and Discussion

We demonstrated and proved in this study that FFTs composed of the same ranked cues but with different exit structures differ systematically in their decision biases. Such systematic differences enable decision makers to adjust their decision criterion by simply arranging the exit structure of an FFT without changing the order of the cues that make up the tree or their specific properties. This makes FFTs similar to SDT in that a mechanism outside the distributions of cues can be used to adjust one's decision criterion.

The difference is that in SDT, the adjustment can be done continuously because decision makers can set their decision criterion to any value, but in FFTs, the adjustment can only be done discretely, limited by the number of FFTs that can be constructed. With the finding that the psychological concept of decision criterion in SDT can indeed be reflected in the structure of an FFT, we now ask whether there is also a similar correspondence between the two in sensitivity.

Study 2: Where Is the Sensitivity in an FFT?

According to SDT, the sensitivity (d') of a decision maker should depend on the distributional characteristics of the person's observations (see Figure 2). When there are multiple cues involved in the decision-making process, Sorkin and Dai (1994) showed that the d' of an ideal observer in SDT should be sensitive to the mean and variance of cues' individual d' s, their intercue correlations, and the number of cues, but not to cues' decision criteria. Investigating how the sensitivity of FFTs responds to those cue properties is one major goal of this study. Because FFTs search cues sequentially, another goal is to explore how different search orders of cues influence FFTs' sensitivity.

Method

We characterize each cue by its two SDT measures: d' and c . Therefore, a set of m cues can be described by two vectors: $d'_i = [d'_1, d'_2, \dots, d'_m]$ and $c_i = [c_1, c_2, \dots, c_m]$. In addition, the intercue correlation ρ_{ij} for each pair of cues also needs to be specified. Ideally, this would require the specification of a set of ρ_{ij} s. However, to simplify the representation, we assume that all ρ_{ij} s are equal, so that a set of m cues can be described by a single correlation parameter ρ . With m cues, there are $(m!) \times (2^{m-1})$ FFTs that can be constructed, with the first term representing the number of ways cues can be ordered and the second indicating the number of available exit structures under a particular cue order. Although various cue-ordering criteria have been proposed for decision trees in general and FFTs in particular (e.g., Dhimi & Ayton, 2001; Martignon et al., 2008; White & Liu, 1994), here we restrict our analysis to ordering cues with respect to the d' of each cue. Nevertheless, in the General Discussion, we consider the relation between cue orders based on d' and information gain, a measure commonly used to construct decision trees.

By assigning specific values to the three cue properties (d'_i , c_i , and ρ), we created 6,006 three-cue conditions. More specifically, there were 286 combinations of d'_i , generated by the following procedure:

1. Start with $d'_i = [0.5, 0.5, 0.5]$, in which the first value of the vector d'_i represents the highest d' of a cue among the three and d'_3 the lowest;
2. Raise d'_1 incrementally by 0.2 until it reaches 2.5;
3. Raise d'_2 incrementally by 0.2 until it reaches d'_1 ;
4. Raise d'_3 incrementally by 0.2 until it reaches d'_2 .

Thus, the last combination of d'_i was [2.5, 2.5, 2.5]. In addition, there were seven levels of the intercue correlation ρ (0, 0.1, 0.2,

0.3, 0.4, 0.5, and 0.6) and three conditions of c_i : (a) all neutral [0, 0, 0], (b) all conservative [1, 1, 1], and (c) all liberal [-1, -1, -1]. Conditions with unequal c_i among the three cues (e.g., [1, 0, -1]) were also explored. Because the results from those unequal- c_i conditions were fundamentally no different from those of the equal- c_i conditions, we do not report them here but in the supplementary materials.⁵ In addition to the 6,006 (i.e., $286 \times 7 \times 3$) three-cue conditions, one five-cue condition, and two other three-cue conditions were analyzed to compare the effect of the number of cues (m) on sensitivity.

To calculate the d' of an FFT, we assume that cue values of an object are drawn from one of two multivariate normal distributions, signal and noise. For noise, the means of all cue variables were set to 0; and for signal, they were $[d'_1, d'_2 \dots d'_m]$. For both distributions, the variance of each cue variable was 1 and the covariance between any pair of cues was ρ . With these distributional parameters, cues' standardized decision criteria $[c_1, c_2 \dots c_m]$ were further converted to specific cue values $[x_{c1}, x_{c2} \dots x_{cm}]$, so that they could be used to calculate the hit and FA rates of an FFT, and in turn its d' (see Appendix B for the detailed procedure).

Effects of Cue Orders on FFTs' Sensitivity

To examine the effects of cue order on the sensitivity of an FFT, let us start with three-cue FFTs. With three cues, six orders are possible. Given that each cue is named after its d' ranking, with 1 being the highest, the six orders are 1-2-3, 1-3-2, 2-1-3, 2-3-1, 3-1-2, and 3-2-1. Figure 6 shows the sensitivity of four FFTs in one cue-property condition, in which $d'_i = [1.7, 1.5, 1.3]$, $c_i = [0, 0, 0]$, and $\rho = 0.3$, and under each of the six cue orders. Results from the other 6,005 three-cue conditions lead to the same conclusions as in this example.

Three major results can be observed in Figure 6. First, the sensitivity of the strictly liberal and conservative trees, FFT_{ss} and FFT_{nn} , is not affected by cue order at all. Because a consensus among the three cues is needed to make either an "s" decision (FFT_{ss}) or an "n" decision (FFT_{nn}), cue order cannot influence the overall hit and FA rates of those trees and in turn their d' s. Second, for the other two trees, FFT_{ns} and FFT_{sn} , cue order does affect their sensitivity. In general, the more the order deviates from the cues' d' order, the lower the sensitivity of an FFT. In other words, when cues differ in their individual d' , a 1-2-3 cue order generally leads to a higher d' than a 2-1-3 order, which is, in turn, better than a 3-1-2 order.

The third and probably most interesting result is that for FFT_{sn} and FFT_{ns} , the two cue orders 1-2-3 and 1-3-2 result in the same sensitivity, as do the orders 2-1-3 and 2-3-1 and the orders 3-1-2 and 3-2-1. Therefore, once the first cue is set, the sensitivity of an FFT_{sn} or an FFT_{ns} does not depend on which cue is considered second and which third. The explanation for this result is that cases that cannot be decided by the first cue require a consensus of the remaining cues to make either an "s" decision (FFT_{sn}) or an "n" decision (FFT_{ns}); in a sense, the remaining two cues result in strictly liberal and conservative subtrees.

For FFTs composed of more than three cues, the general conclusions from Results 1 and 3 still hold. In fact, for FFTs composed of m cues, the following rule applies:

If the exits of an FFT are the same type from its k th to $(m - 1)$ th cues, where $1 \leq k \leq (m - 1)$, the sensitivity of the FFT is determined

strictly by the order of the first $k - 1$ cues but is unaffected by how cues are ordered between the k th and m th cues.

This rule holds for any number of cues with any combination of properties; we refer to it as the *partial order invariance* rule of FFTs. Assuming that $m = 5$ and $k = 3$, the rule indicates that, after the order of the first two cues is set, the sensitivity of an FFT_{snss} will remain the same no matter how the remaining three cues are ordered. Together with the result in Figure 6 that a cue order further deviated from the d' order will generally lead to a lower sensitivity of an FFT, this rule suggests that a decision maker should be concerned only with the ordering of the first $k - 1$ cues in an FFT and that to achieve a higher d' of the FFT, it would be best to order these cues according to their d' ranking.

Effects of Cue Properties on FFTs' Sensitivity

Figure 7 plots the performance of the four three-cue FFTs in the z ROC space for various cue-property conditions. Their cues are ordered according to their d' values (i.e., order 1-2-3). In such an ROC presentation, sensitivity can be inferred simply by taking the difference between the two coordinate values (i.e., $d' = z\text{Hit} - z\text{FA}$). Moreover, the performance of decision makers with equal sensitivity but employing different decision criteria will result in points in this ROC space that can be fit perfectly by a straight line. Each of the four panels in Figure 7 demonstrates the effects of a single cue property on FFTs' sensitivity. In each panel, three levels of a certain cue property are chosen for comparison purposes. At each level, the four points from left to right in the ROC space correspond to the performance of FFT_{nn} , FFT_{ns} , FFT_{sn} , and FFT_{ss} , respectively, and each point represents the average performance of FFTs that share the same cue-property value. Moreover, for each cue-property condition the mean d' of the four FFTs, d' -FFT, is labeled.

Figure 7A shows the effect of $\mu_{d'}$, the average d' of the three cues used to compose an FFT. For the three levels of $\mu_{d'}$, 0.9, 1.5, and 2.1, there are 147, 378, and 147 conditions at each level, respectively. Figure 7B shows the effect of ρ , the intercue correlation among the cues. There are 1,001 conditions at each of the three levels of ρ : 0, 0.3, and 0.6. Figure 7C shows the effect of $\text{Var}_{d'}$, the variance of the three cues' individual d' s. For the three levels of $\text{Var}_{d'}$, 0, 0.33, and 0.97, there are 231, 252, and 84 conditions at each level, respectively. Finally, Figure 7D shows the effect of c_i , the decision criteria of the cues. There are 2,002 conditions at each level of c_i .

We see from Figure 7A that an FFT's sensitivity is positively related to $\mu_{d'}$, which reflects the overall information quality of the cues, and from Figure 7B that an FFT's sensitivity is negatively related to ρ , which is often used as a measure of information redundancy among cues. Figure 7C shows that the sensitivity of FFT_{sn} and FFT_{ns} increases with $\text{Var}_{d'}$, but the sensitivity of FFT_{ss} and FFT_{nn} does not, as indicated by the points clustered in the leftmost and rightmost parts of the ROC space. The main reason

⁵ We examined 24 unequal- c_i conditions that include all 27 combinations from $c = [-1, 0, 1]$, excluding the three equal- c_i conditions, yielding a total of 48,068 ($256 d'_i \times 7\rho \times 24 c_i$) cue-property conditions. The results from the analysis of those conditions can be found in the online supplementary materials.

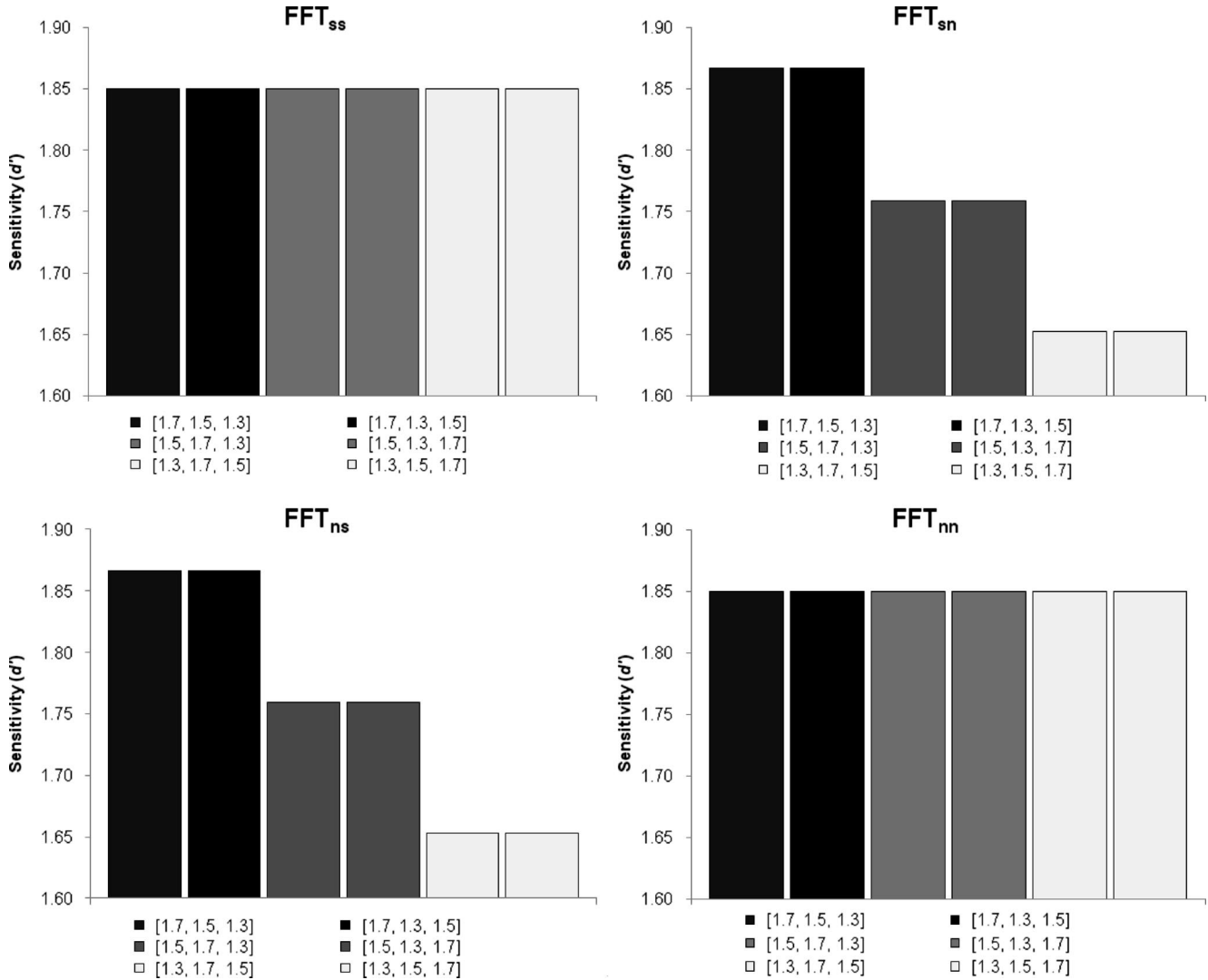


Figure 6. The sensitivity (d') of each three-cue fast-and-frugal tree (FFT) in a cue-property condition where $d'_i = [1.7, 1.5, 1.3]$, $c_i = [0, 0, 0]$, and $\rho = 0.3$. In each panel of the figure, the six bars show the d' s of a certain FFT under six possible cue orders; a cue order is labeled by the d' s of the cues that are searched sequentially. Following the conventions in the main text, the order [1.7, 1.5, 1.3] corresponds to the order 1-2-3, order [1.5, 1.7, 1.3] to the order 2-1-3, and so on.

for this pattern of results is that when $\text{Var}_{d'}$ is high, it is likely that the d' value of the first cue is also high. This helps increase the sensitivity of FFT_{sn} and FFT_{ns}, but not that of FFT_{ss} and FFT_{nn}, which is insensitive to cue orders (see the previous section). Figure 7D reveals two effects of c_i on FFTs' performance: (a) When the decision criteria of the cues progress from being conservative ([1, 1, 1]) to liberal ([-1, -1, -1]), the decision criterion of an FFT also moves from being more conservative to more liberal, and (b) c_i have little effect on the sensitivity of an FFT, as the 12 points in Figure 7D can be fitted well by one straight line.⁶

Because the points plotted in Figure 7 are the result of averaging over many cue conditions, these graphs can obscure the relative impact of different cue properties. To overcome this problem, we ran a series of step-wise regressions using the four cue properties— $\mu_{d'}$, ρ , $\text{Var}_{d'}$, and c_i —as the predictors and the sensitivity of

an FFT or the mean sensitivity of the four FFTs (d' -FFTs) as the predicted variables. The results of these analyses are shown in Table 2. In general, two main results can be seen from the table: (a) The predictors can account for an FFT's sensitivity quite well ($R^2 > .96$), and (b) the most useful predictor is $\mu_{d'}$, which alone leads to an R^2 of around .85. The results of the regression analyses are consistent with those shown in Figure 7. For instance, $\text{Var}_{d'}$ is found to have little effect on the sensitivity of FFT_{ss} and FFT_{nn}, which is also observed in Figure 7C, but it is the second most

⁶ Note that this conclusion is drawn from cue-property conditions where the cues' decision criteria are set to be equal (i.e., $c_1 = c_2 = c_3$). However, even in conditions where they are unequal, the same general conclusion still holds.

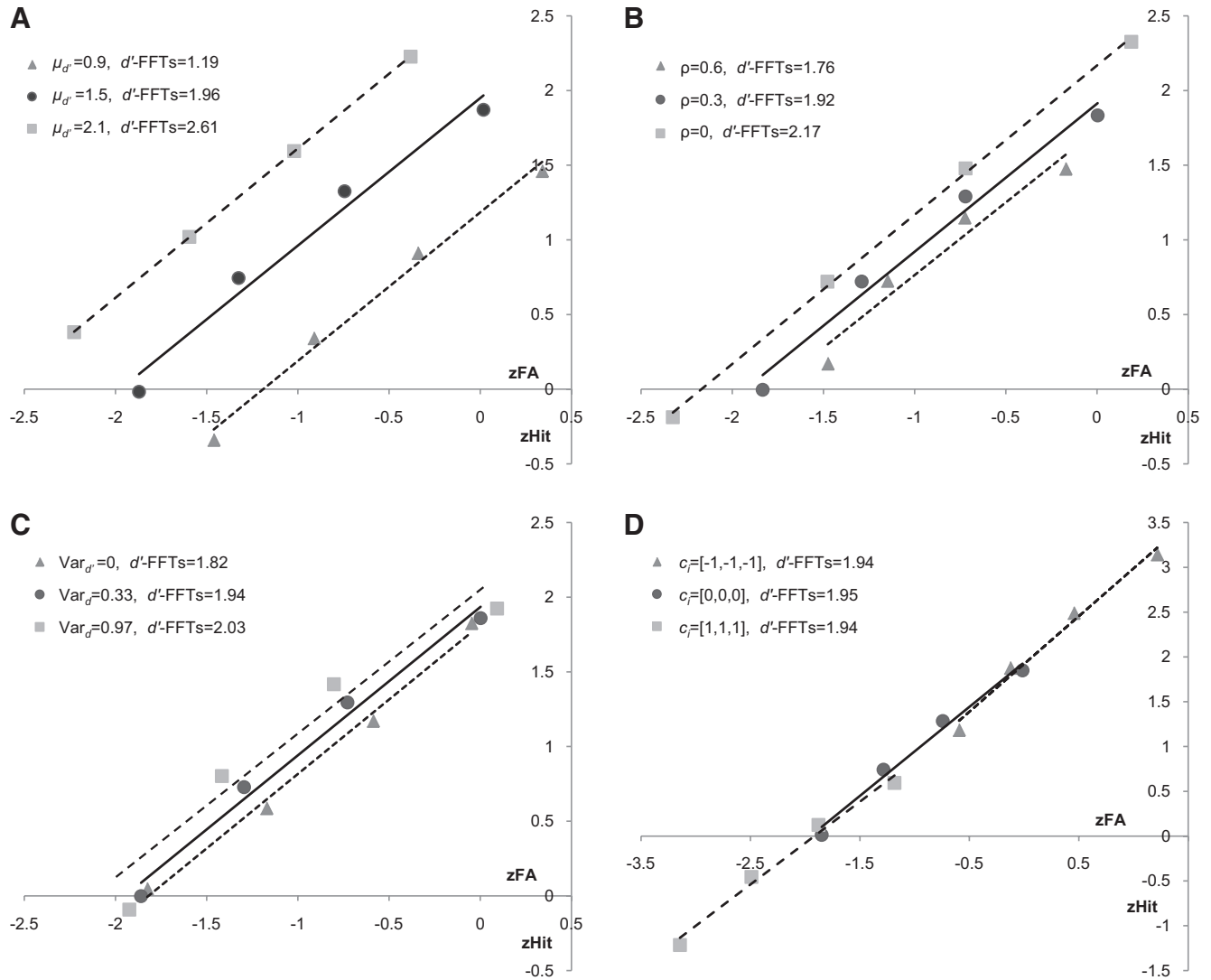


Figure 7. The sensitivity of the four three-cue fast-and-frugal trees (FFTs) averaged over conditions with various cue property values. In each panel, d' -FFTs represents the mean d' of the four FFTs in each parameter condition, and the performances of the four trees— FFT_{nn} , FFT_{ns} , FFT_{sn} , and FFT_{ss} —are reflected as the four points running in turn from bottom-left to upper-right in the zROC space. A. The FFTs' sensitivity at three levels of $\mu_{d'}$, the average d' of the three cues used to construct an FFT. B. Sensitivity at three levels of ρ , the uniform intercue correlation. C. Sensitivity at three levels of $Var_{d'}$, the variance of the three cues' individual d' 's. D. Sensitivity at three levels of c_i , the decision criteria of the cues.

important factor in predicting the sensitivity of FFT_{sn} and FFT_{ns} . According to Sorokin and Dai (1994), the ideal observer of SDT is responsive to the following cue properties: $\mu_{d'}$, ρ , $Var_{d'}$, and m , the number of cues. We find it interesting that the mean d' of the four types of FFTs, d' -FFTs are also sensitive to the first three parameters.

Our analyses show that the sensitivity of an FFT also depends on the number of cues, m . For the sake of brevity, in Figure 8 we illustrate the results of these analyses by comparing two particular three-cue conditions with one five-cue condition. In the five-cue condition, the 16 points running from lower left to upper right represent in turn the performance of FFT_{nnnn} , FFT_{nnns} , FFT_{nnsn} ,

FFT_{nsss} , FFT_{nsnn} , FFT_{nsns} , FFT_{nssn} , FFT_{nsns} , FFT_{snnn} , FFT_{snnns} , FFT_{snnns} , FFT_{snnns} , FFT_{ssnn} , FFT_{ssns} , FFT_{ssns} , FFT_{ssnn} , and FFT_{ssss} . Their relative decision biases provide a concrete demonstration of the lexicographic decision bias property of FFTs (see Study 1). One of the three-cue conditions—the one with $d'_i = [1.43, 1.0, 0.57]$ —is equivalent to the five-cue condition in terms of $\mu_{d'}$, ρ , $Var_{d'}$, and c_i . As can be seen from Figure 8, the sensitivity of the five-cue FFTs appears higher than that of the FFTs from this three-cue condition. However, this does not imply that including more cues in an FFT always has a positive effect on its sensitivity.

Comparing these five-cue FFTs with the FFTs from another three-cue condition with $d'_i = [1.50, 1.25, 1.0]$, in which the three

Table 2
Results of the Step-Wise Regression Analyses on the Sensitivity of Fast-and-Frugal Trees (FFTs)

Predicted variable	Predictors and resulting adjusted R^2 at each step					
	Predictor 1	R^2 (1 ^a)	Predictor 2	R^2 (1, 2 ^a)	Predictor 3	R^2 (1, 2, 3 ^a)
d' -FFTs	$\mu_{d'}$.899	ρ	.974	$\text{Var}_{d'}$.987
d' -FFT _{sn} , d' -FFT _{ns}	$\mu_{d'}$.843	$\text{Var}_{d'}$.912	ρ	.962
d' -FFT _{ss} , d' -FFT _{nm}	$\mu_{d'}$.873	ρ	.968	c_i	.982

^a The number(s) inside the parentheses indicate the adjusted R^2 values using the corresponding predictor(s).

cues are the three best cues of the five-cue condition, we see that FFTs' sensitivities in the two conditions are nearly identical. This result shows that having more cues does not improve an FFT's sensitivity if only cues of lower quality are included. Nevertheless, there is a potential benefit of having more cues in an FFT. As shown in Figure 8, compared with the four FFTs in a three-cue condition, a wider range of decision biases can be covered by the 16 FFTs in the five-cue condition. Therefore, having more cues to work with provides one with more flexibility to construct an FFT with a desired decision bias, although these additional cues may do little to improve its sensitivity.

Finally, from Figures 7 and 8, we see that the points of FFTs in a certain cue-property condition can generally be fitted well by a straight line in the zROC space. This implies that these FFTs are similar in sensitivity, despite their drastic differences in decision criterion. To examine how general a result this is, we calculated the R^2 value of an ROC line that fits the four FFTs in each of the 6,006 three-cue conditions. The average R^2 is .966, indicating that the sensitivity differences among these FFTs are indeed quite small.

Summary and Discussion

In SDT, sensitivity d' is measured by the difference between the means of the signal and noise distributions. Where is d' in an FFT?

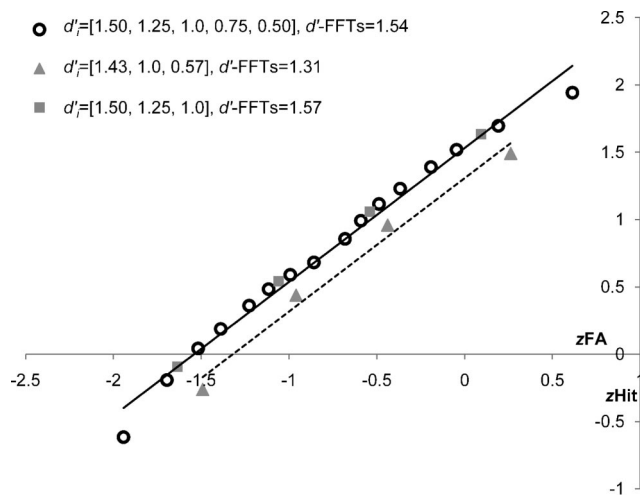


Figure 8. The sensitivity of the 16 fast-and-frugal trees (FFTs) in one five-cue condition, in which $d'_i = [1.50, 1.25, 1.0, 0.75, 0.50]$, and the sensitivity of the eight FFTs in two three-cue conditions, in which $d'_i = [1.50, 1.25, 1.0]$ and $[1.43, 1.0, 0.57]$, respectively. In all three conditions, the intercue correlation ρ is 0.3 and all cues are unbiased ($c_i = 0$).

We showed that sensitivity is found in its cue structure, that is, in the order of the cues and the individual properties of these cues. For three-cue FFTs, which are frequently used in applied settings such as the CCU allocation tree, the first cue determines sensitivity and the order of the remaining cues is irrelevant. In all three-cue conditions investigated in this study, the sensitivity of an FFT decreases when cues with a lower d' are considered first. If there are more than three cues, the partial order invariance rule applies; it implies that an FFT's sensitivity does not depend on the order of consecutive cues that share the same kind of exit. Strictly liberal and conservative FFTs are special cases in which all cues are consecutive cues that share the same exit; therefore, cue order does not affect their sensitivity. Besides cue order, the characteristics of the cues also influence the sensitivity of an FFT.

In Figures 7 and 8, we demonstrated the effects of five cue properties on FFTs' sensitivity. We found that with other cue properties held constant, the d' of an FFT tends to be higher when (a) the mean of cues' individual sensitivity $\mu_{d'}$ is high, (b) the intercue correlation ρ is low, (c) cues' variance $\text{Var}_{d'}$ is high, and (d) there are more and better quality cues available to construct an FFT. We also found that the d' of an FFT is affected only to a small extent by the decision biases of its cues (c_i), which, in contrast, can substantially affect an FFT's overall decision bias. It should be noted that these five cue properties exert similar effects on the sensitivity of the ideal observer of SDT, although the ideal model and FFTs are based on entirely different decision-making mechanisms.

So far, we have only considered sensitivity of FFTs in a small world, where everything is known to a decision maker. But people working in a large world, such as paramedics at the scene of an accident and physicians in an emergency room, often must act quickly on the basis of uncertain information, while being mindful of the serious consequences of their decisions. Therefore, to understand the workings of FFTs better, we consider performance measures that go beyond sensitivity.

Study 3: Performance Beyond Sensitivity

When making decisions, people often care about not only the accuracy of their decisions but also the consequences, which can be roughly captured by the expected value of the decisions. When different decisions yield drastically different consequences, concerns about accuracy often become secondary. To increase the expected value of a decision, it has been argued in SDT that careful attention must be paid to the selection of an appropriate decision criterion. Therefore, from SDT's perspective, a model's capability to adjust its decision criterion in response to the characteristics of a task, mainly the payoffs of the decision outcomes, determines its

expected value performance to a large extent. Besides expected value, speed, which is closely related to the frugality of a model, is another major concern of real-time decision making. Moreover, because a person usually has only limited experience in a task, how effectively a model can help predict the outcomes of unknown cases based on the person's previous learning indicates another important performance measure: predictive accuracy.

In this study, we explore comparative advantages and disadvantages of FFTs by pitting them against three other models. One is the ideal model in SDT, which assumes that information from all cues is integrated into a likelihood ratio and a decision is made by comparing this value to the optimal decision criterion β_{optimal} . The second is the OSS model, which is designed to deal with situations where cue information needs to be sampled sequentially at a cost. The third one is the majority model. Using this model, a person first ranks cues according to their d' values and then checks the binary decisions made by each cue sequentially until a simple majority is reached.

The majority model examined in this study differs from other majority models studied in both the SDT and group decision research (e.g., Hastie & Kameda, 2005; Sorokin et al., 2004) in that it incorporates a search component. For instance, suppose that there are three cues in a task and that their values for an object are x_1 , x_2 , and x_3 . Suppose further that in relation to each cue's decision criterion x_{c_i} , $x_1 > x_{c_1}$, $x_2 > x_{c_2}$, and $x_3 < x_{c_3}$. Thus, the binary decisions of those cues are "s," "s," and "n," respectively. A person using the majority model can stop search after the first two cues, because the majority of cues necessarily favor an "s" decision. As a result of this search algorithm, the majority model does not always search for all available cues to make a decision. The majority model thus shares two common characteristics with FFTs: sequential information search and a simple decision mechanism. However, the two models do differ in one critical aspect: After cues are ordered, one can adjust the decision criterion of an FFT by arranging its exit structure in different ways, whereas the decision criterion of the majority model is fully determined by the decision criteria of the cues. Such inflexibility may hurt the majority model's performance in terms of expected value.

In the following sections, we compare the performance of the FFTs, the ideal SDT model, and the majority model on four different measures: sensitivity, frugality, predictive sensitivity, and expected value. Because the OSS model reduces to the ideal model when search costs are not taken into account, we add OSS into the mix only when considering expected value.⁷

Sensitivity and Frugality

The sensitivity (d') and frugality of each model are derived analytically under the assumptions that (a) cue values are drawn from two multivariate normal distributions, (b) the prior probabilities of the two decision categories $P(S)$ and $P(N)$ are equal, and (c) the precise values of all cue properties are known. Following the processes explained in Appendix B, we derived the d' and frugality of each of the three models—FFT, majority, and ideal—in the 6,006 three-cue conditions specified in Study 2. The average d' s of the three models are 2.37, 1.94, and 1.92 for the ideal, FFT, and the majority model, respectively; their average frugalities are in turn 3.0, 1.75, and 2.36.

When everything is known about the cues, the ideal model optimally uses this information. It achieves its high level of sensitivity at the cost of always considering all three cues. In contrast, the majority model sometimes requires only two cues to make a decision and the FFTs can often get by with just one. Although the FFTs search fewer cues, the FFTs and majority model manage to achieve similar levels of sensitivity. This result shows that sensitivity (or accuracy) and frugality need not trade off; a model can perform well in one aspect without sacrificing the other (see also Gigerenzer, Todd, & the ABC Research Group, 1999).

Predictive Sensitivity

It is important that a model makes accurate predictions about previously unseen cases. In this study, we measure the predictive accuracy of a model in terms of predictive sensitivity. Whereas sensitivity is derived analytically based on the assumption that the exact values of all parameters a model ever needs to know are known, predictive sensitivity depends on parameter values estimated from a sample. Various methods have been applied in psychology to measure models' predictive accuracy (e.g., Myung et al., 2000; Shiffrin, Lee, Kim, & Wagenmakers, 2008). Here, we use accumulative prediction (AP; see Wagenmakers, Grünwald, & Steyvers, 2006), a method similar to cross-validation, to measure models' predictive sensitivity.⁸ We implemented AP as follows:

1. Five objects were randomly drawn from both the noise and signal distributions; the parameter values of a model were then estimated based on these 10 objects.⁹
2. Either the noise or the signal distribution was randomly picked and one new object was randomly drawn from this distribution. Based on the previously estimated parameter values, a model made a prediction about whether this object was a signal or a noise.
3. Model parameters were estimated again based on all objects that had already been seen, including the original 10 plus the new one.
4. Steps 2 and 3 were reiterated for $(N - 10)$ times, with N being the sample size of interest. The accumulated prediction performance, which was measured by d' through tallying hits and FAs from all iterations, is the predictive sensitivity of a model.

⁷ The sensitivity and frugality of the OSS model depend on the specific values of the two stopping criteria β_s and β_n . When they are set at higher values, the sensitivity of the OSS model will be relatively high, but more cues will be searched before making a decision. In some sense, the OSS model can be considered as a constrained version of the ideal model, with β_s and β_n regulating its sensitivity and frugality.

⁸ We thank Michael Lee for suggesting the AP method, which is procedurally similar to leave-one-out cross-validation. However, in contrast to cross-validation, AP can be connected to Bayesian and minimum description length (MDL) model selection methods.

⁹ There is no consensus in the AP literature on the size of the sample initially used to estimate parameters. We settled on 10 because it was problematic to estimate all the ideal model's parameters with smaller sample sizes.

The number of parameters needed by the models varies. The ideal model needs estimates of the means and variance–covariance matrices of two multivariate normal distributions (signal and noise) to calculate the likelihood ratio of a cue value vector $[x_1, x_2, \dots, x_m]$. Both the FFTs and the majority model need to estimate the hit and FA rates of the cues, so that their d' s can be calculated to establish the d' cue order.

Our analysis of predictive sensitivity focuses on conditions where three cues are available (i.e., $m = 3$), as was the case in our previous analysis of sensitivity. However, because of the extensive computation required for deriving the predictive sensitivity of a model, we considered fewer conditions (18) compared to the analysis of sensitivity. Specifically, among the 18 conditions, there are three levels of ρ (0, 0.3, and 0.6), two levels of $\mu_{d'}$ (1 and 1.5), and three levels of $\text{Var}_{d'}$ (0, 0.04, and 0.25). In all conditions, cues' decision criteria (c_i) were set to $[0, 0, 0]$, as our (unreported) preliminary analysis showed that different values of c_i had little effect on the relative performance of the three models. Moreover, five sample sizes ($N = 20, 40, 80, 200,$ and $2,000$) were examined. They match the sample sizes employed in experiments under the SDT framework, in which the number of trials administered to a participant varies from tens to thousands (e.g., Green & Swets, 1966; Macmillan & Creelman, 2005). In each sample size and under every cue-property condition, the results were based on 1,000,000 cases (e.g., for a sample size of 200 objects, 5,000 random samples were run). The average predictive sensitivity of each model over the 18 cue-property conditions is shown in Figure 9.

Three major results can be observed: First, the majority model's predictive sensitivity is unaffected by sample size, an indication that it does not have a tendency to overfit and is extremely robust. This result may appear surprising, because the model orders cues according to their d' values, and any misestimation of the order

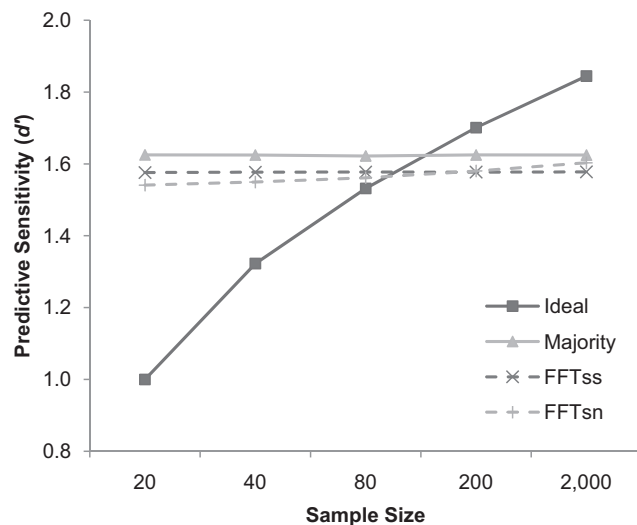


Figure 9. The predictive sensitivity of the three models: ideal, majority, and fast-and-frugal trees (FFTs). For each sample size, the performance of a model is averaged over 18 cue-property conditions (see text). Note that because FFT_{nn} and FFT_{ns} have the same predictive sensitivity as FFT_{ss} and FFT_{sn} , respectively, we only show the results of FFT_{ss} and FFT_{sn} in the figure.

might reduce its predictive sensitivity. However, cue order actually has no effect on the sensitivity of this model, as any search order in a trial will lead to the same decision.¹⁰

Second, the predictive sensitivity of the FFTs is close to that of the majority model and varies little across different sample sizes, indicating that FFTs are very robust as well. For the strictly liberal and conservative FFTs, sample size has no effect on their predictive sensitivity whatsoever; the sensitivity of these FFTs, like that of the majority model, is totally insensitive to cue orders. For the two less biased FFTs, their predictive sensitivity increases gradually with sample size. Our previous analysis has shown that misplacing the first cue in these two FFTs generally reduces their sensitivity (see Study 2). Because such an error is more likely to occur when sample sizes are small, this result was expected. What came as a surprise to us are the rather limited performance gains that accrue as sample size increases from 20 to 2,000. Apparently, it is not too difficult to sort out which cues to search first even when sample size is extremely small.

Finally, the predictive sensitivity of the ideal model responds dramatically to increasing sample size and does not surpass that of the other two models until sample size reaches 200. As the ideal model requires estimates for a large number of parameters, it tends to overfit the data, a characteristic shared by many complex models (e.g., Myung et al., 2000).¹¹

Expected Value

To measure a model's expected value, decision payoffs must be specified. Table 3 shows three payoff conditions, in which the value of a decision outcome is expressed in generic, measurement-free units. Across the three conditions, the positive values or rewards for the two correct decisions, $V(\text{Hit})$ and $V(\text{CR})$, are set to be equal, but the negative values or penalties for the two incorrect decisions, $V(\text{Miss})$ and $V(\text{FA})$, differ. They are set up in such a way that to maximize expected value, the decision criterion should gradually shift from being more liberal when the penalty for a miss is higher than for an FA (Condition I), to unbiased when the penalties for a miss and an FA are equal (II), and finally to more conservative when the penalty for an FA is higher than for a miss (III). The optimal likelihood ratio decision criterion β_{optimal} in each condition, shown at the bottom of Table 3, reflects the shifting requirements.

In addition to decision payoffs, another factor related to expected value is the cost of cue search. We examined four cue-cost conditions: (a) Cost = 0, in which there is no cost for acquiring cue information; (b) Cost = d' , in which the cost of acquiring information from a cue is the same as the cue's d' value; (c) Cost = 1; and (d) Cost = 2, where acquiring each cue costs 1 unit and 2 units, respectively. Because results from the last three conditions share the same general pattern, differing from each other by fixed

¹⁰ Interested readers should consult Appendix B in Karelaia (2006) for a more thorough treatment of why cue order does not affect the majority model's decisions.

¹¹ We replicated all the predictive sensitivity analyses reported in this section using the equal-split cross-validation method. The results, consistent with those obtained using AP, are reported in the supplementary materials.

Table 3
Values of the Four Decision Outcomes and the Optimal Likelihood Ratio Decision Criterion, β_{optimal} in Each of the Three Payoff Conditions

Decision outcome	Payoff condition		
	I (heavy penalty for miss)	II (balanced payoff)	III (heavy penalty for FA)
V(FA)	-10	-10	-50
V(CR)	10	10	10
V(Hit)	10	10	10
V(Miss)	-50	-10	-10
β_{optimal}^a	0.33	1	3

Note. FA = false alarm; CR = correct rejection.

^a β_{optimal} was calculated based on the assumption that the prior odds ratio of noise to signal, $P(N)/P(S)$, equals 1.

constants, only results from the Cost = 0 and Cost = d' conditions are reported.

Expected values of FFTs. Before we compare the expected value results across the models, we first examine the expected value results for the FFTs, which are shown in Figure 10. The expected value of an FFT in each Payoff \times Cue-Cost condition was calculated based on the average *sensitivity* of that FFT over 6,006 three-cue-property conditions. In the Cost = 0 condition, we see that the expected values of an FFT can differ substantially across the payoff conditions. The expected values of the strictly liberal tree FFT_{ss} provide a case in point. In Payoff Condition I, where the penalty for a miss is heavier than that for an FA and a liberal decision criterion is preferred, the expected value of FFT_{ss} is the highest among the four, with another liberal tree FFT_{sn} trailing closely behind. In Payoff Condition II, where a miss and an FA are equally costly and a neutral decision criterion is preferred, FFT_{ss} 's expected value is below that of the two less biased trees, FFT_{sn} and FFT_{ns} , and is the same as that of FFT_{nn} . In Payoff Condition III, where an FA now has a more severe consequence than a miss and a conservative decision criterion is preferred,

FFT_{ss} lags far behind any of the other three FFTs in expected value.

Expected values of the other FFTs all lead to the same general pattern: An FFT's expected value depends largely on the fit between its decision criterion and the optimal decision criterion in a specific task; the better the fit, the higher the expected value. This pattern holds in the Cost = d' condition as well. The only difference between an FFT's expected values in the Cost = 0 and Cost = d' conditions is the cost of cue search.

Because the decisions of FFTs may differ substantially in terms of expected value, it is essential to select the type of FFT most appropriate for a given payoff condition. In contrast to the precision required to set the optimal decision criterion β_{optimal} in SDT, all that is needed to select a proper FFT is the rough relative consequences of the decision outcomes, especially the two decision errors. If one error, say miss, is obviously more consequential than the other, a more liberally biased FFT, such as FFT_{ss} , would be most suitable. If the two errors are roughly equally consequential, then a less biased FFT, such as FFT_{sn} or FFT_{ns} , should be preferred. Experiments have shown that although people's deci-

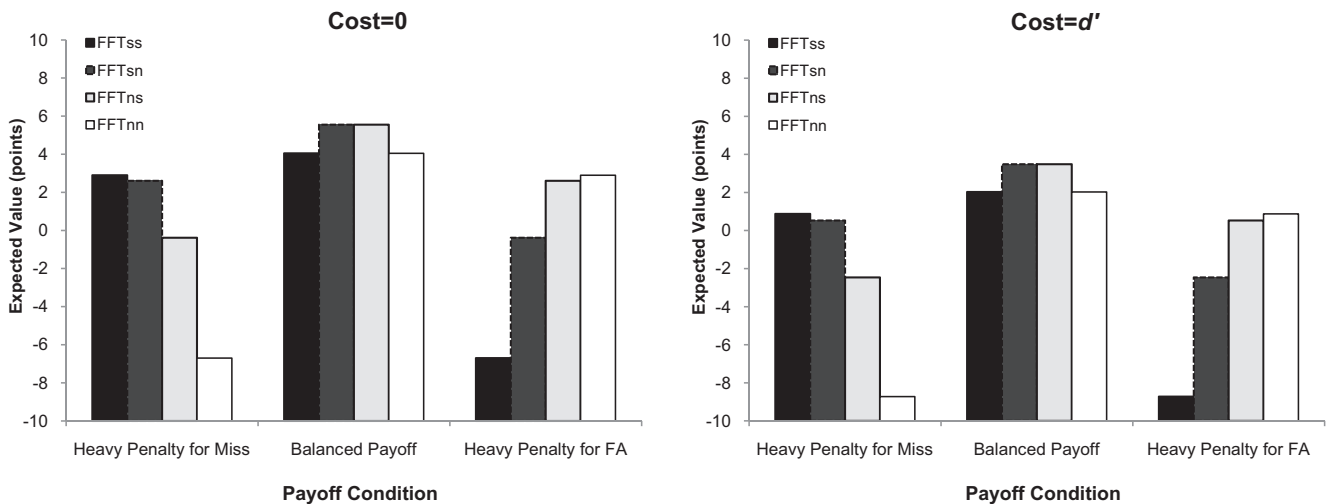


Figure 10. The expected values of the four fast-and-frugal trees (FFT) in three payoff and two cue-cost conditions. In each condition, the expected value of an FFT was derived based on the mean sensitivity of the FFT over 6,006 three-cue conditions. FA = false alarm.

sion criteria often deviate from β_{optimal} , the directions of those criteria are generally consistent with β_{optimal} (e.g., Erev, 1998; Green & Swets, 1966); this directional knowledge is the most critical requirement for selecting an FFT in a task.

Let us now look back at the CCU allocation tree (see Figure 4a), which is an FFT_{sn} with a relatively liberal decision bias. Besides its good sensitivity ($d' = 1.90$), its decision bias also fits well to its task of assigning patients reporting chest pain to either the CCU or a regular nursing bed. In this task, a miss is when a patient at risk for a heart attack is sent to a regular nursing bed, whereas an FA is when a patient at low risk is sent to the CCU. For most people, the cost of a miss will be more severe than the cost of an FA, suggesting that a more liberal criterion is appropriate for this task. Among the four types of FFTs, only an FFT_{ss} is more liberal than an FFT_{sn} . But, the markedly superior sensitivity of an FFT_{sn} compared to that of an FFT_{ss} ($d' = 0.73$) makes FFT_{sn} clearly the overall best choice in this particular task.

Model comparisons in expected value. Unlike the analysis of FFTs' expected values, the comparisons among models' expected values are based on their *predictive sensitivity*. Because the sensitivities of the FFTs and the majority model vary little, if at all, across different sample sizes, the expected values of these two models barely change when sensitivity is replaced by predictive sensitivity as the basis for expected value. However, this is not the case for the ideal model. The expected value of the ideal model is the decision payoff calculated based on the model's predictive sensitivity—which, depending on the size of a sample, can be much lower than its sensitivity—minus the total cue search cost that is fixed in a cue-cost condition, given that the ideal model always searches all cues. Calculating the expected value of the OSS model, however, is more complicated.

Recall that the OSS model limits cue search by setting up two stopping criteria β_s and β_n in addition to β_{optimal} (see Figure 3). Therefore, the OSS model must learn values of β_s and β_n that can maximize expected value. There is no good analytic solution for finding the optimal β_s and β_n , so these values must be found numerically.¹² Moreover, unlike the ideal model, cues must be searched in a certain order in the OSS model. And unlike FFTs and the majority model, the OSS model does not have to search in cues' d' order but in whichever order yields the maximum expected value in a given sample. Thus, in a three-cue situation, all six cue orders must be considered. Finally, the OSS model needs to estimate all the parameters estimated by the ideal model, because it updates information using the integrated likelihood ratio $\text{LR}(x_i)$, whose value depends on knowing the distributional parameters of the cues.

Overall, it is a computationally demanding process to estimate parameters for the OSS model from a sample. Considering that the accumulative prediction procedure requires estimating all parameters afresh for each new case, it was quite time consuming to derive the expected values of the OSS model.¹³ Limited by time and the speed of our computers, we ran just one three-cue condition, in which $d'_i = [2.0, 1.5, 1.0]$, $\rho = 0.3$, and $c_i = [0, 0, 0]$, to compare expected values of the four models. Two decision payoffs were implemented—balanced payoff (Condition II) and heavy penalty for miss (Condition I), together with two cue-cost conditions—Cost = 0 and Cost = d' . In each Payoff \times Cue-Cost condition, four sample sizes were examined: $N = 20, 80, 200$, and 2,000. The number of samples run under each sample size was set

so that a total of 100,000 samples were run for each (e.g., for $N = 80, 1,250$ samples).

The models' expected value performances are shown in Figure 11. In each payoff condition, only the expected values of two FFTs, the two whose decision criteria fit better to the payoff condition, are shown. From the figure, we see that the OSS and the ideal models perform identically in the Cost = 0 conditions. Because the OSS model does not need to set criteria to limit search when there is no cost to sample cues, the expected value of the model depends solely on its predictive sensitivity, as does that of the ideal model. As a result, both models tend to overfit the data when sample size is small and perform less well than the two FFTs until sample size reaches 200. In the Cost = d' conditions, the OSS model does perform better than the ideal model, mainly because of its mechanisms for limiting cue search. However, even more learning is now required for the OSS model to outperform the two better suited FFTs: Its expected value does not surpass that of the two FFTs until sample size is very large ($N = 2,000$), a clear indication that extensive learning is required for the OSS model to stably estimate its many parameters.¹⁴

As was the case for predictive sensitivity, the majority model's expected value is not affected by sample size. Compared with the two FFTs, the majority model performs equally well in the balanced Payoff \times Cost = 0 condition but worse when there is either a heavy penalty for a miss or cue search becomes costly (Cost = d'), and still worse when both factors are present (the heavy penalty for Miss \times Cost = d' condition). Because the majority model on average searches more cues than the FFTs, it follows that its expected value is affected more negatively by cue cost than the FFTs'. The interesting result here is the generally poor performance of the majority model in the condition where a biased decision is desired. With a set of cues, unlike the decision criterion of an FFT, which can be adjusted by rearranging the exit structure, the decision criterion of the majority model is fixed. This inflexibility may not cause a problem when the majority model's decision criterion happens to be consistent with the demands of a payoff condition; however, it is problematic when the payoffs change and a different decision criterion is then required. Thus,

¹² Analytical solutions are available for relatively simple situations in which d' of each cue is the same and cues are independent (e.g., Rapoport & Burkheimer, 1971). The difficulty of deriving an analytical solution drastically increases for most cue-property conditions probed in this study; therefore, we settled on a numerical method. Specifically, two parameters b_s and b_n were used, so that $\ln\beta_s = \ln\beta_{\text{optimal}} + b_s$ and $\ln\beta_n = \ln\beta_{\text{optimal}} - b_n$. We varied b_s and b_n from 0 to 2, with a step size of 0.1. Thus, 21 values of each were tried, which resulted in $21 \times 21 = 441$ pairs of $\ln\beta_s$ and $\ln\beta_n$.

¹³ For instance, it took one computer with a 2.66G-Hz Intel Core 2 Duo CPU and 4GB internal memory about 20 min to finish running a sample of 80 cases in the Cost = d' condition. Therefore, it took us a total of around 420 hr to finish 1,250 samples with that sample size. In contrast, it took less than 2 min to obtain the results of FFTs under the same condition. With 10 computers running simultaneously, it took us more than 3 weeks to get the results of the OSS model reported in Figure 11.

¹⁴ Through experimenting with several sample sizes other than the ones shown in Figure 11, we estimated that in the Cost = d' condition, it would require a sample size of roughly 800 for the OSS model to perform as well as the two better suited FFTs in the balanced payoff condition, and 1,500 in the heavy penalty for miss condition.

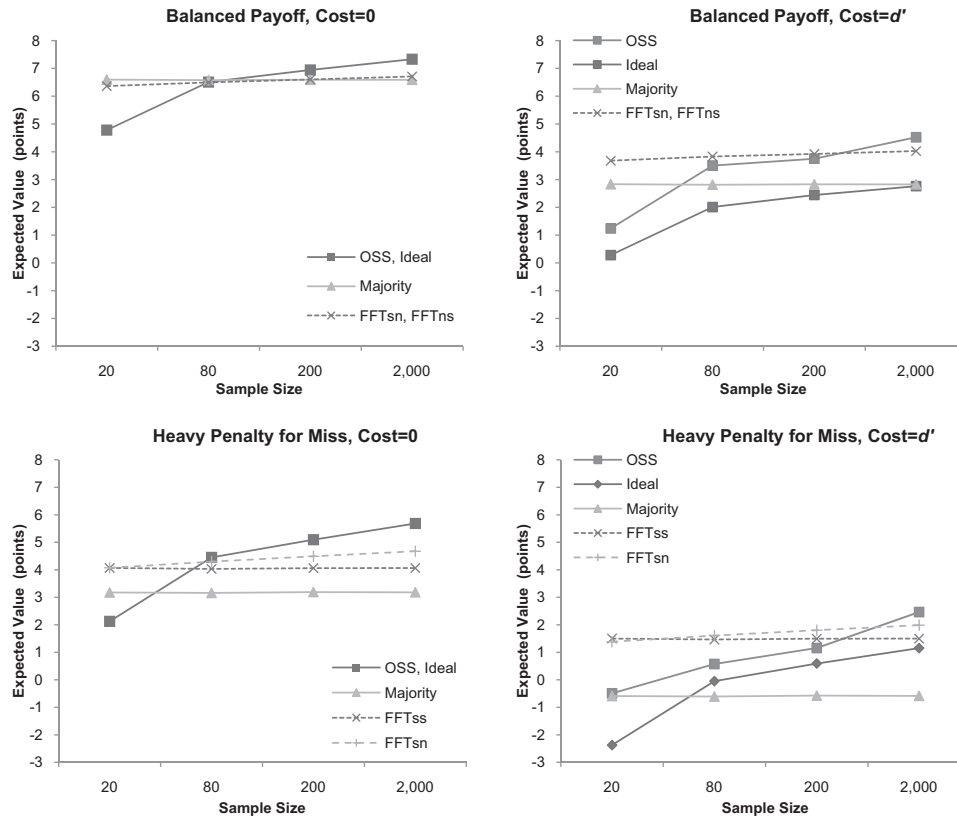


Figure 11. The expected values of four models: optimal sequential sampling (OSS), ideal, majority, and fast-and-frugal trees (FFTs), in two Payoff \times two Cue-Cost \times four Sample Size conditions. The cue property values are $d'_i = [2.0, 1.50, 1.0]$, $\rho = 0.3$, and $c_i = [0, 0, 0]$. Note that the performances of the OSS and ideal models are identical in the Cost = 0 conditions, and the performances of FFT_{sn} and FFT_{ns} are identical in the balanced payoff conditions.

given that one can select an appropriate FFT to use in a payoff condition, FFTs perform better overall than the majority model in expected value.

In general, the expected value of a model depends on the model's sensitivity or predictive sensitivity if learning is required, on its frugality if cue search is costly, and on the ability of the model to adjust its decision criterion in response to the payoff structure of a task. Among the four models discussed here, the majority model lacks a criterion-adjustment mechanism, and the ideal model tends to search too many cues. Therefore, their expected values are generally lower than the other two models'. The strength of the OSS model lies in its ability to strike a good balance between accuracy and frugality with the consideration of all variables that may affect this trade-off, in addition to its ability to select a decision criterion that matches perfectly to a task's payoff structure. Compared with the OSS model, with its arduous pursuit for perfection, the FFTs are lazy models that work only hard enough to find a satisfactory cue order and an exit structure that fits reasonably well to the payoff structure. However, as we see from Figure 11, the FFTs perform nearly as well as the OSS model when sample sizes are fairly large, and are clearly better than the OSS model when sample sizes are relatively small. Although we investigated only one cue property condition, given the characteristics of the models, we are confident that the general results

observed in this condition would be found in other conditions as well.

Model Comparison Summary

In this study, we compared four models—FFTs, majority, ideal, and OSS—in four aspects of performance: sensitivity, frugality, predictive sensitivity, and expected value. FFTs perform similarly to the majority model in sensitivity and predictive sensitivity but are more frugal and can achieve expected values that are equal to or higher than that of the majority model. Hence, FFTs trump the majority model overall. The performances of the ideal model and the FFTs demonstrate a clear trade-off: Whereas the ideal model has higher sensitivity, the FFTs are more frugal. This trade-off is to some extent mitigated when accuracy is measured in predictive sensitivity instead of sensitivity. The ideal model trails behind the FFTs in predictive sensitivity when sample sizes are small (less than 200), because of both its tendency to overfit data and FFTs' robustness to sampling noise. When cues are sampled without any costs, the expected value of the ideal model is identical to that of the OSS model. Whenever there are costs for cue search, however, the expected value of the ideal model falls short of that of the OSS model. As discussed above, FFTs prove again to be highly robust when performance is measured in terms of expected value. They

outperform the OSS model even in conditions that provide abundant learning opportunities (sample size larger than 500) for the OSS model to estimate its parameters.

General Discussion

FFT's have been adopted to deal with decision problems in various domains (e.g., Dhimi, 2003; Fischer et al., 2002; Green & Mehr, 1997; Super, 1984). Through the lens of SDT, our goal is to understand FFT's basic properties, when they do well, and how they can do even better. In Study 1, we found that systematic differences in FFT's exit structures determine the relative differences in their decision criteria, according to the principle of *lexicographic decision bias*. In Study 2, we examined how the sensitivity of FFT's depends on the order of cues used to construct them and the properties of those cues. The effect of cue orders is summarized by the *partial order invariance rule*, which states that switching the order of contiguous cues that share like exits does not affect an FFT's sensitivity. Once cues are ordered by their d' values, an FFT's sensitivity is in general positively related to the mean and variance of cues' individual d' 's ($\mu_{d'}$ and $\text{Var}_{d'}$), negatively related to the intercue correlation (ρ), and scarcely affected by the individual decision biases (c_i) of the cues. In Study 3, we compared FFT's with the majority, ideal, and OSS models. We found that FFT's were frugal and robust, competing well with the other models in terms of predictive sensitivity and expected value, especially with smaller sample sizes.¹⁵ In the following, we discuss several topics related to the findings of our studies and FFT's in general.

Ordering Cues by d' and Information Gain

Constructing a typical decision tree requires ordering cues and pruning unnecessary exits (e.g., Breiman et al., 1984; Dieterich, 2000; Quinlan, 1993). However, constructing an FFT requires no pruning, because the number of exits is set, as per its definition, leaving cue ordering the major concern. In keeping with SDT, we ordered cues according to d' , which, to the best of our knowledge, had not been tried before. Therefore, it is of interest to know how such orders correspond to those based on other cue-ordering indices, among which information gain is probably the most popular (Quinlan, 1993). To connect information gain and the two SDT measures d' and c , we ran a series of analyses,¹⁶ which show that (a) unlike d' , information gain depends on the prior odds of the two decision events, and the closer these odds are to 1, the higher the information gain; (b) a cue's information gain can be predicted well by its d' and c , with the effect of d' being more prominent; (c) in a three-cue set, when cues' decision criteria are randomly assigned, a d' order is totally consistent with an information gain order around 67.5% of the time, whereas the two are completely at odds with each other only about 3.5% of the time; and when the decision criteria of the three cues are set to be equal, the two orders are consistent with each other 100% of the time. Because cues' decision criteria are unlikely to be either totally random or exactly the same in a real task, these latter two extreme conditions suggest that we should observe perfectly consistent cue orders using the two indices between 67.5% and 100% of the

time. Such high consistency in ordering suggests that ordering cues using either d' or information gain should lead to similar performances of an FFT.

Ideal and Real Observers

Achieving the performance of the ideal observer model requires precise knowledge of cue properties, availability of all cue information, and abundant resources dedicated to cue information integration. Failure to meet any of those conditions will result in a less than ideal performance. Participants who have been trained to integrate information optimally in simple tasks involving merely two cues failed to achieve ideal performance even after hundreds of learning trials with timely and accurate feedback (e.g., Ashby & Maddox, 1990, 1992). Yet the demands of the ideal observer model pale in comparison to those required by the OSS model. To achieve the maximum expected value, the OSS model's search for the optimal stopping and decision criteria adds an extra burden of both knowledge and processing. The exceptionally long time it took our computers to find the optimal stopping criteria does not rule out the plausibility of this model—after all, humans are known to achieve feats no computers can match—but it surely does not speak in its favor. Even if humans do make decisions using models that share the same general elements as the OSS model (see Figure 3), as some studies have argued (e.g., Busemeyer & Townsend, 1993; Lee & Cummins, 2004; Lee & Dry, 2006; Vickers, 1979), there is no evidence that the decision and stopping criteria they adopt are optimal.

After years of research experience with the ideal observer model, Wilson Geisler (2003), echoing Savage (1954), remarked that

organisms generally do not perform optimally, and hence one should not think of an ideal observer as a potentially realistic model of the actual performance of the organism. Rather, the value of an ideal observer is to provide a precise measure of the stimulus information available for performing the task, a computational theory of how to perform the task, and an appropriate benchmark against which to compare the performance of the organism. (p. 825)

¹⁵ Part of our results in Studies 2 and 3 was derived on the basis of the assumption that the signal and noise distributions are normal with equal variances. We did explore some conditions with the variance ratio different from 1 in Study 2 and found that such a change did not affect the major results regarding the effects of cue properties on FFT's sensitivity, measured in either d' , which is a proper measure for equal-variance normal distributions, or Az , which has been argued as a more robust sensitivity measure (e.g., Rotello, Masson, & Verde, 2008; Swets, 1986). These results can be found in the supplementary materials. Future studies could examine the robustness of our results more thoroughly by systematically varying the variance ratio of the distributions and/or their forms.

¹⁶ Some results can be found in the online supplemental materials. For the sake of brevity, we do not introduce the formula used to calculate information gain. Oaksford and Chater (1996) and Nelson (2005) provide introductions to information gain and examples of how it has been applied in psychological research.

The same can very well be said for the OSS model.¹⁷ Abandoning the search for optimality, we see FFTs as alternative models designed to deal with realistic constraints, employing search and decision mechanisms that are simple, frugal, flexible, robust, and transparent. FFTs in their various forms have already been adopted by many experts in the field, be they paramedics (e.g., the START triage system), pediatricians (e.g., Fischer et al., 2002), judges (e.g., Dhimi, 2003), and perhaps even professional burglars (e.g., Garcia-Retamero & Dhimi, 2009).

Potential Applications of FFTs to SDT

We have looked at FFTs through the lens of SDT. What might we learn by looking at SDT from the perspective of FFTs? Gigerenzer and Brighton (2009) conjectured that heuristics achieve high predictive accuracy because, in part, bias in their predictions protects them from perturbations in training samples. For FFTs, such bias resides in the coarse nature of their search and decision mechanisms. Specifically, search by FFTs requires only a rank order of cues, and the decision criterion of an FFT is adjusted discretely without taking account of the exact values of the decision payoffs and prior odds. Such coarseness might explain why FFTs outperform the ideal observer and OSS models when parameters cannot be pinned down easily. We wonder whether these models could benefit from forsaking their fastidious pursuit of precision.

By lazily adjusting parameters in a coarse way, these optimization models may achieve higher levels of predictive accuracy, especially when sample sizes are small. Although exactly how to implement a lazy parameter search is beyond the scope of the present study, we offer some thoughts about how such a search might proceed for the OSS model: (a) Instead of testing all possible cue orders, search cues only in the order of their d 's; (b) ignore intercue correlations; (c) select the stopping criteria from a limited pool of values—for instance, one might choose from five criterion values that are “very stringent,” “stringent,” “intermediate,” “lax,” and “very lax” to correspond to task situations where costs are “very high,” “high,” “medium,” “low,” and “very low,” respectively; and (d) instead of estimating β_{optimal} , the decision criterion could be selected from a small set of criteria that correspond to “heavy penalty for miss,” “balanced payoff,” and “heavy penalty for FA” payoff conditions. Simplifications as such would act as strong priors on the OSS model's parameter values, reducing its flexibility but potentially improving its performance for small samples.

Variants of the Majority Model

The majority model has served as an alternative to the ideal type of models in SDT (e.g., Sorokin et al., 2004; Sorokin, West, & Robinson, 1998). We modified the model by adding a search rule; that is, searching cues in d ' order and stopping search when a simple majority is reached. In addition, its decision criterion could be adjusted by choosing the number of cues k_s (for signal) and k_n (for noise) out of m cues needed to make a decision. In its current form, the majority model in a three-cue situation uses the 2s–2n rule—if two out of three cues agree on the same decision, a decision is made, independent of cue order.

Alternatively, one could use a 2s–1n rule, which states that two votes are needed to make an “s” decision while only one is needed for an “n” decision. The change appears minor, but it can affect all aspects of the majority model's performance. First, the order of cue search could affect the frugality and sensitivity of the model, because the first cue searched influences how quickly and accurately a noise decision will be made. Second, assuming that cues are still searched by their d ' order, the sensitivity of the model will likely change, although whether it would be better or worse than the 2s–2n rule is uncertain. Third, because the rule makes it easier to make a noise decision, it leads to a more conservative decision bias compared with the 2s–2n rule. Thus, in payoff conditions where a more conservative bias is desired (e.g., there is a heavier penalty for an FA than a miss), the 2s–1n majority model may outperform the 2s–2n version we tested.

Some majority models with other decision rules turn out to be identical to FFTs. For instance, with three cues, the 3s–1n and 3n–1s versions of the majority model are the same as FFT_{nm} and FFT_{ss}, respectively. Thus, despite differences in how the models are framed and in what domains they are commonly applied, clear connections can be drawn between these two types of decision models. From our perspective, understanding how the variants of the majority model perform will be an interesting and highly relevant extension of the present study.

Theory Integration: Connecting Paradigms

In psychology, the goal of bridging different theories is rarely pursued. As Walter Mischel (2009) put it, many psychologists still tend to treat theories like toothbrushes—no self-respecting person wants to use anyone else's. We see this article as an exercise in theory integration, bridging small and large worlds and corresponding optimizing and satisficing theories. When we began our analysis, we were unsure whether SDT and FFTs, two apparently disparate theories, would have much in common. Even now, we still find it surprising that the key concepts of SDT can be found in the structure of FFTs, albeit represented by different processes, such as the choice of an FFT's exit structure setting its decision criterion. We hope that this step toward conceptual integration will prompt others to search for hidden features shared between seemingly distinctive psychological theories. Whereas borrowing toothbrushes can indeed be dangerous, or at the very least distasteful, sharing theories can be a healthy practice, contributing to more coherent and unified theory in psychology.

¹⁷ It should be noted that we have limited our discussion to the *optimal* sequential sampling model (i.e., OSS). Various nonoptimal sequential sampling models exist and provide good accounts for plenty of behavioral results (e.g., Lee & Corlett, 2003; Lee & Dry, 2006; Vickers, 1979). In general, those models make simplifying assumptions about the distributional properties of the cues and how the stopping criteria are learned and set.

References

- Ashby, F. G., & Berretty, P. M. (1997). Categorization as a special case of decision making or choice. In A. A. J. Marley (Ed.), *Choice, decision, measurement: Essays in honor of R. Duncan Luce* (pp. 367–388). Hillsdale, NJ: Erlbaum.

- Ashby, F. G., & Maddox, W. T. (1990). Integrating information from separable psychological dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 598–612. doi:10.1037/0096-1523.16.3.598
- Ashby, F. G., & Maddox, W. T. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 50–71. doi:10.1037/0096-1523.18.1.50
- Binmore, K. (2009). *Rational decisions*. Princeton, NJ: Princeton University Press.
- Birnbaum, M. H. (1983). Base rates in Bayesian inference: Signal detection analysis of the cab problem. *American Journal of Psychology*, *96*, 85–94. doi:10.2307/1422211
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*, 700–765. doi:10.1037/0033-295X.113.4.700
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Pacific Grove, CA: Wadsworth.
- Bussemeyer, J. R., & Rapoport, A. (1988). Psychological models of deferred decision making. *Journal of Mathematical Psychology*, *32*, 91–134. doi:10.1016/0022-2496(88)90042-9
- Bussemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459. doi:10.1037/0033-295X.100.3.432
- Bussemeyer, J. R., Weg, E., Barkan, R., Li, X., & Ma, Z. (2000). Dynamic and consequential consistency of choices between paths of decision trees. *Journal of Experimental Psychology: General*, *129*, 530–545. doi:10.1037/0096-3445.129.4.530
- Cook, L. (2001). The world trade center attack: The paramedic response: An insider's view. *Critical Care*, *5*, 301–303. doi:10.1186/cc1054
- Dhmi, M. K. (2003). Psychological models of professional decision-making. *Psychological Science*, *14*, 175–180. doi:10.1111/1467-9280.01438
- Dhmi, M. K., & Ayton, P. (2001). Bailing and jailing the fast-and-frugal way. *Journal of Behavioral Decision Making*, *14*, 141–168. doi:10.1002/bdm.371
- Dhmi, M. K., & Harries, C. (2001). Fast and frugal versus regression models of human judgment. *Thinking & Reasoning*, *7*, 5–27. doi:10.1080/13546780042000019
- Dietterich, T. G. (2000). An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine Learning*, *40*, 139–158.
- Dube, C., Rotello, C. M., & Heit, E. (2010). Assessing the belief bias effect with ROCs: It's a response bias effect. *Psychological Review*, *117*, 831–863. doi:10.1037/a0019634
- Edwards, W. (1965). Optimal strategies for seeking information. *Journal of Mathematical Psychology*, *2*, 312–329. doi:10.1016/0022-2496(65)90007-6
- Erev, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychological Review*, *105*, 280–298. doi:10.1037/0033-295X.105.2.280
- 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. doi:10.1037/a0018526
- Fischer, J. E., Steiner, F., Zucol, F., Berger, C., Martignon, L., Bossart, W., . . . Nadal, D. (2002). Using simple heuristics to target macrolide prescription in children with community-acquired pneumonia. *Archives of Pediatric and Adolescent Medicine*, *156*, 1005–1008.
- Garcia-Retamero, R., & Dhmi, M. K. (2009). Take-the-best in expert-novice decision strategies for residential burglary. *Psychonomic Bulletin & Review*, *16*, 163–169. doi:10.3758/PBR.16.1.163
- Geisler, W. S. (2003). Ideal observer analysis. In L. Chalupa & J. Werner (Eds.), *The visual neuroscience* (pp. 825–837). Boston, MA: MIT Press.
- Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, *98*, 254–267. doi:10.1037/0033-295X.98.2.254
- Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, *1*, 107–143. doi:10.1111/j.1756-8765.2008.01006.x
- Gigerenzer, G., Hoffrage, U., & Ebert, A. (1998). AIDS counselling for low-risk clients. *AIDS Care*, *10*, 197–211. doi:10.1080/09540129850124451
- Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.
- Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 3–34). New York, NY: Oxford University Press.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.
- Green, L., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *Journal of Family Practice*, *45*, 219–226.
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, *78*, 81–91. doi:10.1037/0022-3514.78.1.81
- Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review*, *112*, 494–508. doi:10.1037/0033-295X.112.2.494
- Hogarth, R. M., & Karelaia, N. (2007). Heuristics and linear models of judgment: Matching rules and environments. *Psychological Review*, *114*, 733–758. doi:10.1037/0033-295X.114.3.733
- Karelaia, N. (2006). Thirst for confirmation in multi-attribute choice: Does search for consistency impair performance? *Organizational Behaviour and Human Decision Processes*, *100*, 128–143. doi:10.1016/j.obhdp.2005.09.003
- Kee, F., Jenkins, J., McIllwaine, S., Patterson, C., Harper, S., & Shields, M. (2003). Fast and frugal models of clinical judgment in novice and expert physicians. *Medical Decision Making*, *23*, 293–300. doi:10.1177/0272989X03256004
- Lee, M. D., & Corlett, E. Y. (2003). Sequential sampling models of human text classification. *Cognitive Science*, *27*, 159–193. doi:10.1207/s15516709cog2702_2
- Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic Bulletin & Review*, *11*, 343–352.
- Lee, M. D., & Dry, M. J. (2006). Decision-making and confidence given uncertain advice. *Cognitive Science*, *30*, 1081–1095. doi:10.1207/s15516709cog0000_71
- Luan, S., Sorkin, R. D., & Itzkowitz, I. (2010). *A signal detection model of information acquisition and advice taking*. Manuscript submitted for publication.
- Macmillan, N. A., & Creelman, C. (2005). *Detection theory: A user's guide* (2nd ed.). Cambridge, England: Cambridge University Press.
- Martignon, L., Katsikopoulos, K. V., & Woike, J. K. (2008). Categorization with limited resources: A family of simple heuristics. *Journal of Mathematical Psychology*, *52*, 352–361. doi:10.1016/j.jmp.2008.04.003
- Martignon, L., Vitouch, O., Takezawa, M., & Forster, M. (2003). Naïve and yet enlightened: From natural frequencies to fast and frugal decision trees. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment, and decision making* (pp. 189–211). Chichester, England: Wiley.
- Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision

- maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging*, 22, 796–810. doi:10.1037/0882-7974.22.4.796
- Mischel, W. (2009). The toothbrush problem. *Association for Psychological Science Observer*, 21, 11.
- Myung, I. J., Forster, M., & Browne, M. W. (Eds.). (2000). A special issue on model selection. *Journal of Mathematical Psychology*, 44(1).
- Nelson, J. D. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, 112, 979–999. doi:10.1037/0033-295X.112.4.979
- Oaksford, M., & Chater, N. (1996). Rational explanation of the selection task. *Psychological Review*, 103, 381–391. doi:10.1037/0033-295X.103.2.381
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, England: Cambridge University Press.
- Pleskac, T. J. (2007). A signal detection analysis of the recognition heuristic. *Psychonomic Bulletin & Review*, 14, 379–391.
- Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. San Mateo, CA: Morgan Kaufmann.
- Rapoport, M., & Burkheimer, G. J. (1971). Models for deferred decision making. *Journal of Mathematical Psychology*, 8, 508–538. doi:10.1016/0022-2496(71)90005-8
- Reilly, B. M., Evans, A. T., Schaidt, J. J., Das, K., Calvin, J. E., Moran, L. A., . . . Martinez, E. (2002). Impact of a clinical decision rule on hospital triage of patients with suspected acute cardiac ischemia in the emergency department. *Journal of American Medical Association*, 288, 342–350. doi:10.1001/jama.288.3.342
- Rotello, C. M., Masson, M. E. J., & Verde, M. F. (2008). Type I error rates and power analyses for single-point sensitivity measures. *Perception & Psychophysics*, 70, 389–401. doi:10.3758/PP.70.2.389
- Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.
- Shiffrin, R. M., Lee, M. D., Kim, W.-J., & Wagenmakers, E.-J. (2008). A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science*, 32, 1248–1284. doi:10.1080/03640210802414826
- Smith, L., & Gilhooly, K. (2006). Fast and frugal versus regression models of decision-making: The case of prescribing for depression. *Applied Cognitive Psychology*, 20, 265–274. doi:10.1002/acp.1189
- Sorkin, R. D., & Dai, H. (1994). Signal detection analysis of the ideal group. *Organizational Behavior and Human Decision Processes*, 60, 1–13. doi:10.1006/obhd.1994.1072
- Sorkin, R. D., Luan, S., & Itzkowitz, J. (2004). Group decision and deliberation: A distributed detection process. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 464–483). London, England: Blackwell. doi:10.1002/9780470752937.ch23
- Sorkin, R. D., West, R., & Robinson, D. E. (1998). Group performance depends on the majority rule. *Psychological Science*, 9, 456–463. doi:10.1111/1467-9280.00085
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments & Computers*, 31, 137–149.
- Super, G. (1984). *START: A triage training module*. Newport Beach, CA: Hoag Memorial Hospital Presbyterian.
- Swets, J. A. (1986). Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. *Psychological Bulletin*, 99, 181–198. doi:10.1037/0033-2909.99.2.181
- Swets, J. A., & Green, D. M. (1961). Sequential observations by human observers of signals in noise. In C. Cherry (Ed.), *Information theory: Proceedings of the Fourth London Symposium* (pp. 177–195). London, England: Butterworth.
- Tanner, W. P., Jr., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61, 401–409. doi:10.1037/h0058700
- Vickers, D. (1979). *Decision processes in visual perception*. New York, NY: Academic Press.
- Wagenmakers, E.-J., Grünwald, P., & Steyvers, M. (2006). Accumulative prediction error and the selection of time series models. *Journal of Mathematical Psychology*, 50, 149–166. doi:10.1016/j.jmp.2006.01.004
- Wald, A. (1947). *Sequential analysis*. New York, NY: Wiley.
- Wallsten, T. S., & Gonzalez-Vallejo, C. (1994). Statement verification: A stochastic model of judgment and response. *Psychological Review*, 101, 490–504. doi:10.1037/0033-295X.101.3.490
- Welch, H. G. (2004). *Should I be tested for cancer? Maybe not and here's why*. Berkeley, CA: University of California Press.
- White, A. P., & Liu, W. (1994). Bias in information-based measures in decision tree induction. *Machine Learning*, 15, 321–329. doi:10.1007/BF00993349
- Wixted, J. T., & Stretch, V. (2004). In defense of the signal-detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review*, 11, 616–641.

Appendix A

Proof of the Lexicographic Decision Bias Property of FFTs

We prove here that the hit rate of the tree $\text{FFT}_{=s^{**}}$ is no lower than that of the tree $\text{FFT}_{=n^{**}}$. Assume that the first differentiating cue is the k th cue in a sequence of m cues. Before that cue, the two trees have identical exits (represented by $=$); after that cue, the two may or may not have identical exits (represented by **). The hit rate of each of the two trees can be expressed by a three-component equation:

$$\text{FFT}_{=s^{**}}: \sum_1^{k-1} P(\text{Hit}) + P(\text{Hit})_k + \sum_{k+1}^m P(\text{Hit})$$

$$\text{FFT}_{=n^{**}}: \sum_1^{k-1} P(\text{Hit}) + 0 + \sum_{k+1}^m P(\text{Hit}).$$

Because the two trees are identical in their first components and $P(\text{Hit})_k$ of $\text{FFT}_{=n^{**}}$ is 0, if the second component $P(\text{Hit})_k$ of $\text{FFT}_{=s^{**}}$ is no lower than the third component $\sum_{k+1}^m P(\text{Hit})$ of $\text{FFT}_{=n^{**}}$, the overall hit rate of $\text{FFT}_{=s^{**}}$ should be no lower than that of $\text{FFT}_{=n^{**}}$. $P(\text{Hit})_k$ of $\text{FFT}_{=s^{**}}$ is equal to $P[(x_{sk} > x_{ck}) \mid \Omega]$, in which Ω represents all objects that have yet to be classified before the k th cue. $\sum_{k+1}^m P(\text{Hit})$ of $\text{FFT}_{=n^{**}}$ can be expressed by the following equation:

$$\sum_{k+1}^m P(\text{Hit}) = P[(x_{sk} > x_{ck}) \cap \text{exit}_{(k+1)} \mid \Omega] + P[(x_{sk} > x_{ck}) \cap \text{exit}_{(k+1)}^\dagger \cap \text{exit}_{(k+2)} \mid \Omega] + \dots$$

$$+ P[(x_{sk} > x_{ck}) \cap \text{exit}_{(k+1)}^\dagger \cap \text{exit}_{(k+2)}^\dagger \cap \dots \cap \text{exit}_m \mid \Omega].$$

In the equation, each component on the right-hand side represents the hit rate of a cue, from the $(k + 1)$ th cue onward. For all components, if it is an “n” exit on the i th cue, $\text{exit}_i = \emptyset$ and $\text{exit}_i^\dagger = (x_{si} > x_{ci})$; and if it is an “s” exit, $\text{exit}_i = (x_{si} > x_{ci})$ and $\text{exit}_i^\dagger = (x_{si} < x_{ci})$. Because each component is mutually exclusive and all are subsets of $P[(x_{sk} > x_{ck}) \mid \Omega]$, their summation cannot be greater than $P[(x_{sk} > x_{ck}) \mid \Omega]$, which is $P(\text{Hit})_k$ of $\text{FFT}_{=s^{**}}$. Therefore, it is proved that $\sum_{k+1}^m P(\text{Hit})$ of $\text{FFT}_{=n^{**}}$ cannot be greater than $P(\text{Hit})_k$ of $\text{FFT}_{=s^{**}}$; and the overall hit rate of $\text{FFT}_{=s^{**}}$ cannot be lower than that of $\text{FFT}_{=n^{**}}$.

By replacing the term x_s with x_n and following the same steps as above, the FA rate of $\text{FFT}_{=s^{**}}$ can also be proved to be no lower than that of $\text{FFT}_{=n^{**}}$.

Appendix B

Deductions of Models’ Sensitivity and Frugality

The deductions of an FFT’s sensitivity (d') and frugality are demonstrated through the example of FFT_{sn} . First, to calculate the d' of FFT_{sn} , we need to know its hit and FA rates. The $P(\text{Hit})$ and $P(\text{FA})$ of FFT_{sn} are calculated through the following equations:

$$P(\text{Hit})_{\text{FFT}_{\text{sn}}} = P(x_{s1} > x_{c1}) + P[(x_{s1} < x_{c1}) \cap (x_{s2} > x_{c2}) \cap (x_{s3} > x_{c3})]$$

$$P(\text{FA})_{\text{FFT}_{\text{sn}}} = P(x_{n1} > x_{c1}) + P[(x_{n1} < x_{c1}) \cap (x_{n2} > x_{c2}) \cap (x_{n3} > x_{c3})]$$

The value of each probability component on the right-hand side of the equations can be obtained with the knowledge of x_{ci} and the parameter values of the two multivariate normal distributions of the cues. The frugality of FFT_{sn} (F) is calculated through the following equation:

(Appendices continue)

$$F_{\text{FFT}_{\text{sn}}} = P(\text{S}) \cdot \left\{ \frac{1 \cdot P(x_{s1} > x_{c1}) + 2 \cdot P[(x_{s1} < x_{c1}) \cap (x_{s2} < x_{c2})]}{3 \cdot P[(x_{s1} < x_{c1}) \cap (x_{s2} > x_{c2})]} + \right\} +$$

$$P(\text{N}) \cdot \left\{ \frac{1 \cdot P(x_{n1} > x_{c1}) + 2 \cdot P[(x_{n1} < x_{c1}) \cap (x_{n2} < x_{c2})]}{3 \cdot P[(x_{n1} < x_{c1}) \cap (x_{n2} > x_{c2})]} + \right\}$$

$P(\text{S})$ and $P(\text{N})$ are the prior probabilities of signal and noise, respectively. The sensitivity and frugality of other FFTs can be worked out in similar ways.

The deductions of the majority model's sensitivity and frugality are demonstrated through the case where m (total number of cues) is 3 and k (the simple majority number) is 2. As with FFTs, to calculate the d' of the majority model, we need to calculate its hit and FA rates first. These two probabilities are

$$P(\text{Hit})_{\text{Majority}} = P[(x_{s1} > x_{c1}) \cap (x_{s2} > x_{c2})] + P[(x_{s1} > x_{c1}) \cap (x_{s2} < x_{c2}) \cap (x_{s3} > x_{c3})]$$

$$+ P[(x_{s1} < x_{c1}) \cap (x_{s2} > x_{c2}) \cap (x_{s3} > x_{c3})]$$

$$P(\text{FA})_{\text{Majority}} = P[(x_{n1} > x_{c1}) \cap (x_{n2} > x_{c2})] + P[(x_{n1} > x_{c1}) \cap (x_{n2} < x_{c2}) \cap (x_{n3} > x_{c3})]$$

$$+ P[(x_{n1} < x_{c1}) \cap (x_{n2} > x_{c2}) \cap (x_{n3} > x_{c3})]$$

The frugality of this particular majority model is

$$F_{\text{Majority}} = P(\text{S}) \cdot \left\{ \frac{2 \cdot \{P[(x_{s1} > x_{c1}) \cap (x_{s2} > x_{c2})] + P[(x_{s1} < x_{c1}) \cap (x_{s2} < x_{c2})]\}}{3 \cdot \{P[(x_{s1} > x_{c1}) \cap (x_{s2} < x_{c2})] + P[(x_{s1} < x_{c1}) \cap (x_{s2} > x_{c2})]\}} + \right\} +$$

$$P(\text{N}) \cdot \left\{ \frac{2 \cdot \{P[(x_{n1} > x_{c1}) \cap (x_{n2} > x_{c2})] + P[(x_{n1} < x_{c1}) \cap (x_{n2} < x_{c2})]\}}{3 \cdot \{P[(x_{n1} > x_{c1}) \cap (x_{n2} < x_{c2})] + P[(x_{n1} < x_{c1}) \cap (x_{n2} > x_{c2})]\}} + \right\}$$

Similar equations can be developed to calculate other majority models' sensitivity and frugality when m and k take values other than 3 and 2. The Matlab code used in our study to derive the sensitivity and frugality of the FFTs and those of the majority model can be found in the online supplemental materials.

The frugality of the ideal model is always the total number of cues m . An equation was developed by Sorokin and Dai (1994) to calculate the ideal d' of m cues directly:

$$d'_{\text{Ideal}} = \sqrt{\frac{m \cdot \text{Var}_{d'}}{1 - \rho} + \frac{m \cdot \mu_{d'}^2}{1 + \rho \cdot (m - 1)}}$$

In the equation, $\text{Var}_{d'}$ and $\mu_{d'}$ are the variance and mean of the cues' individual d' 's, and ρ is the uniform intercue correlation of the cues.

Received January 1, 2010
Revision received December 6, 2010
Accepted December 6, 2010 ■