ACADEMY OF Management Journal

Ecological Rationality: Fast-and-Frugal Heuristics for Managerial Decision Making under Uncertainty

Journal:	*Academy of Management Journal*
Manuscript ID	AMJ-2018-0172.R2
Manuscript Type:	Special Research Forum: New Ways of Seeing
Keywords:	Lab experiment < Quantitative Orientation < Research Methods, Simulation < Quantitative Orientation < Research Methods, Behavioral decision theory < Theoretical Perspectives, Selection, staffing, and recruiting < Human Resource Management and Industrial Relations < Topic Areas, Decision making (micro perspectives) < Behavior < Organizational Behavior < Topic Areas

Abstract:	Heuristics are often viewed as inferior to “rational” strategies that exhaustively search and process information. Introducing the theoretical perspective of ecological rationality, we challenge this view and argue that under conditions of uncertainty common to managerial decision making, managers can actually make better decisions using fast-and-frugal heuristics. Within the context of personnel selection, we show that a heuristic called Δ-inference can more accurately predict which of two job applicants would perform better in the future than logistic regression, a prototypical rational strategy. Using data from 236 applicants at an airline company, we demonstrate in Study 1 that despite searching less than half of the cues, Δ-inference can lead to more accurate selection decisions than logistic regression. After this existence proof, we examine in Study 2 the ecological conditions under which the heuristic predicts more accurately than logistic regression using 1,728 simulated task environments. Finally, in Study 3, we show in an experiment that participants adapted their strategies to the characteristics of a task, and increasingly so the greater their previous experience in selection decisions. The aim of this article is to propose ecological rationality as an alternative to current views about the nature of heuristics in managerial decisions.

Ecological Rationality: Fast-and-Frugal Heuristics for Managerial Decision Making under Uncertainty

Shenghua Luan
Chinese Academy of Sciences
luansh@psych.ac.cn

Jochen Reb
Singapore Management University
jreb@smu.edu.sg

Gerd Gigerenzer
Max Planck Institute for Human Development
gigerenzer@mpib-berlin.mpg.de

Acknowledgements:
We are deeply grateful to Edgar Kausel for making the data set analyzed in Study 1 available to us. We thank Tian Liu and Jun Fang for programming and running machine learning algorithms on the Study 1 data set, Evaldas Jablonskis for programming the experiment in Study 3, and Filip Lievens and Gokhan Ertug for providing helpful comments on earlier versions of this paper. This research is partially supported by the CAS "Pioneer 100, Category A" grant awarded to Shenghua Luan (Grant #Y8CX153006). The authors contributed equally to this work. Correspondence concerning this article should be addressed to Shenghua Luan and/or Jochen Reb.

ECOLOGICAL RATIONALITY: FAST-AND-FRUGAL HEURISTICS FOR MANAGERIAL DECISION MAKING UNDER UNCERTAINTY

ABSTRACT

Heuristics are often viewed as inferior to "rational" strategies that exhaustively search and process information. Introducing the theoretical perspective of *ecological rationality*, we challenge this view and argue that under conditions of uncertainty common to managerial decision making, managers can actually make better decisions using *fast-and-frugal heuristics*. Within the context of personnel selection, we show that a heuristic called *Δ-inference* can more accurately predict which of two job applicants would perform better in the future than logistic regression, a prototypical rational strategy. Using data from 236 applicants at an airline company, we demonstrate in Study 1 that despite searching less than half of the cues, Δ-inference can lead to more accurate selection decisions than logistic regression. After this existence proof, we examine in Study 2 the ecological conditions under which the heuristic predicts more accurately than logistic regression using 1,728 simulated task environments. Finally, in Study 3, we show in an experiment that participants adapted their strategies to the characteristics of a task, and increasingly so the greater their previous experience in selection decisions. The aim of this article is to propose ecological rationality as an alternative to current views about the nature of heuristics in managerial decisions.

Keywords: ecological rationality, fast-and-frugal heuristics, comparative model testing, Δ-inference, heuristics and biases, personnel selection, selection decisions

It is widely held that managers use heuristics to make decisions, but also that they should not. Heuristics are often considered inferior, or second-best, to strategies that are deemed "rational" because they exhaustively search and process all available information (e.g., Bazerman & Moore, 2008; Dean & Sharfman, 1993, 1996). Strongly influenced by the "heuristics and biases" research program in psychology (e.g., Gilovich, Griffin, & Kahneman, 2002; Tversky & Kahneman, 1974), the underlying assumption is that managers' bounded cognitive capacities lead them to use heuristics, and that doing so leads to dangerous biases (i.e., systematic deviations from logic and probability theories; Hammond, Keeney, & Raiffa, 1998) and ultimately less effective decisions (Dean & Sharfman, 1993). Their use is tolerated by a presumed general effort-accuracy tradeoff, whereby decision makers save on effort but only in exchange for lower accuracy (Beach & Mitchell, 1978; Payne, Bettman, & Johnson, 1993).

Research on *ecological rationality* fundamentally challenges this view of heuristics as second-best and argues that "less can be more," that is, better decisions can be made with less information (Gigerenzer & Gaissmaier, 2011; Todd, Gigerenzer, & the ABC Research Group, 2012). Extending Herbert Simon's theory of bounded rationality (1947; 1955), theorizing on ecological rationality posits that not only fast-and-frugal heuristics search and process less information but many conditions exist under which they can actually lead to better decisions (Gigerenzer, 2016). Thus, even if managers had unbounded cognitive capacities, they could still make more accurate, efficient, and effective decisions using heuristics under many real-world managerial conditions, in contrast to the notion of a general effort-accuracy tradeoff.

In addition, ecological rationality views decision makers as having access to an "adaptive toolbox" of strategies, including both fast-and-frugal heuristics and more complex strategies (Gigerenzer & Selten, 2002). Effective decision makers select and adapt an appropriate strategy from this toolbox according to the structure of the environment. Simple heuristics such as the recognition heuristic, with which decision makers select options they recognize, can work surprisingly well (Gigerenzer & Goldstein, 2011). Research has also shown that experts use heuristics in a variety of contexts, including the medical (Wegwarth, Gaissmaier & Gigerenzer, 2009) and judicial domains (Dhami, 2003).

Furthermore, unlike the laws of logic and probability theories, which are sometimes held as universal standards of good reasoning, an ecological rationality perspective does not naïvely claim that heuristics are always better but instead emphasizes the fit between a decision strategy and task requirements and, more generally, between the organism and its environment (Todd et al., 2012). A strategy is ecologically rational to the degree that it reaches a goal, such as accurate predictions, for a certain type of task. Interestingly, some of the conditions typical of managerial decisions match well with those under which heuristics tend to be particularly effective, including fundamental uncertainty (rather than risk; Knight, 1921) and limited opportunities to learn (Gigerenzer, 2016). Under these conditions, it becomes exceedingly difficult to predict future states or events (rather than fit to past data) such as the performance of a job candidate, the effectiveness of a novel strategy, or the success of a new venture. For example, the future performance of job candidates, a key criterion in personnel selection decisions, is notoriously difficult to predict based on applicant data, with about 70% of the variance unexplained even after using the most valid predictors (Highhouse, 2008; Schmidt & Hunter, 1998). In such tasks, complex strategies tend to extract too much from existing data, mistaking noise for signal; as a result, they overfit. In contrast, by ignoring the less important information, simple heuristics can end up being more robust and better at predicting the outcomes of different options (Gigerenzer, 2016; Newell & Simon, 1972).

The present research introduces ecological rationality to the study of managerial heuristics at both prescriptive and descriptive levels. To advance this novel approach, we present three studies situated within the context of personnel selection. We chose this context for two main reasons. First, recruiting the right people is one of the most influential managerial decisions, with managers considering talent acquisition among the top current priorities (Schwartz, Collins, Stockton, & Wagner, 2017) and organizations devoting tremendous resources to recruiting (e.g., US$124 billion in 2011 alone; Leonard, 2011). Second, whereas a considerable literature in personnel selection has focused on the validities of different cues to predict future performance (Schmidt & Hunter, 1998), comparatively little research has examined how cues are integrated, heuristically or with more complex strategies, to reach decisions (Kausel & Slaughter, 2013).

The goal of Study 1 is to show that a heuristic can result in more frugal (i.e., searching fewer cues) *and* more accurate personnel decisions than a “rational” strategy that considers all available information. This study, however, provides only an existence proof and no ecological analysis of the conditions under which less can be more. Study 2 attempts to fill this gap by systematically investigating such conditions using computer simulations with realistic parameters in personnel decisions. Building on these prescriptive findings, Study 3 is descriptive and asks whether people actually adapt their use of strategies to task characteristics.

Our research makes several theoretical contributions. First, it contributes broadly to the theory of managerial decision making by introducing ecological rationality as a novel perspective on managerial heuristics. We provide the first comprehensive investigation of both descriptive and prescriptive aspects of fast-and-frugal heuristics in managerial decision making. In so doing, we extend initial ventures in this area (Artinger et al., 2015; Luan & Reb, 2017) and challenge views of managerial heuristics as inferior to "rational" managerial decision strategies. Moreover, we argue that the performance of managerial heuristics depends on their fit to the environment. As such, we propose a more nuanced and balanced theory of managerial heuristics and move the conversation towards a contingency theory of managerial decision making, where the rationality of decision strategies, heuristic or otherwise, is primarily ecological rather than economic or logical (in the sense of internal consistency).

Second, our research contributes more specifically to the literature on personnel selection. Much research on personnel selection has focused on the assessment and validities of different cues (e.g., Sackett & Lievens, 2008; Schmidt & Hunter, 1998), and some has examined *what* cues managers actually use, such as general mental ability, conscientiousness, and interview performance (e.g., Dougherty, Ebert, & Callender, 1986; Kausel, Culbertson, & Madrid, 2016). By examining *how* cues are used, our research responds to calls for more work on cue integration and decision strategies in personnel selection (e.g., Kausel & Slaughter, 2013; Ryan & Ployhart, 2014). This process-oriented research not only helps advance understanding of selection decisions but is also relevant to practice: By understanding the process of cue integration, we can find intervention spots and design suitable aids to improve selection decisions.

Finally, our research makes a methodological contribution by introducing methodologies commonly used in the study of ecological rationality. We demonstrate how they can be applied to study managerial heuristics in a robust, falsifiable, and in-depth manner by employing the following principles (Gigerenzer & Gaissmaier, 2011): (1) formal models of heuristics, as opposed to mere verbal labels; (2) comparative testing of heuristics versus other strategies, as opposed to testing a single model; and (3) testing the predictive accuracy of strategies, as in out-of-sample predictions, as opposed to only fitting parameters of a model to known data. We also discuss how these methodologies could be valuable in studying other important questions in organizational and management scholarship.

THEORY

Managerial Heuristics as Products of Bounded Rationality

In 1947, Simon published *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations*, a book with seminal impact on organizational scholarship equaled by few others. In it, Simon argued that administrative behavior (i.e., management) can be viewed as a collection of decision-making activities and that an insightful way to understand organizations is to study the decision-making processes of managers: “Decision-making is the heart of administration, and the vocabulary of administrative theory must be derived from the logic and psychology of human choice” (xiii–xiv). Ever since Simon’s work, decision making, such as personnel selection and strategic decision making, has been considered a (if not *the*) quintessential managerial task.

Simon’s contribution went beyond identifying decision making as an essential managerial activity. In discussing how managers make decisions, Simon fleshed out the ideas of bounded rationality for the first time. In the introduction to the second edition of the book (1957: xxv), he wrote: “While economic man [sic] maximizes—selects the best alternative from among all those available to him; his cousin, whom we shall call administrative man [sic], satisfices—looks for a course of action that is satisfactory or ‘good enough’.” For humans to satisfice, Simon proposed that they rely mostly on heuristics, simple but effective mental tools for problem solving and decision making, because their cognitive capacities are bounded (Simon, 1955, 1990). This view of bounded rationality as resulting from cognitive limitations has been the prevailing explanation of why managers use heuristics.

Yet Simon also argued that in an uncertain world, which characterizes many managerial decisions, no single strategy performs best across all situations. Instead, rationality depends on how well a strategy fits the task environment. He expressed this with a scissors analogy: “Human rational behavior… is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor” (Simon, 1990: 7). This adaptive view of heuristics was the starting point for the systematic study of the ecological rationality—as opposed to the economic or logical (ir)rationality—of heuristics.

Managerial Heuristics as Products of Ecological Rationality

Subsequent to Simon’s original work, the environmental fit aspect of bounded rationality was largely neglected in favor of the limited cognitive capacities aspect as the heuristics and biases program became dominant in research on judgment and decision making (e.g., Gilovich et al., 2002; Tversky & Kahneman, 1974), including managerial decision making (e.g., Highhouse, Dalal, & Salas, 2013). This program focuses on how using heuristics leads to outcomes that depart systematically from those dictated by logical or statistical rules. Similarly, the assumed gold standard for managerial decision making is often economic or logical rationality, and most biases studied in managerial decision making are violations of coherence or consistency (e.g., Bazerman & Moore, 2008). Interestingly, research suggests that there is little evidence that violations of syntactical, content-blind axioms of consistency are costly in terms of less wealth, health, or happiness (Arkes, Gigerenzer, & Hertwig, 2016).

More recently, some research has taken a different approach to managerial heuristics. Eisenhardt, Bingham, and colleagues argued in both qualitative (Bingham & Eisenhardt, 2011) and simulation studies (Davis, Eisenhardt, & Bingham, 2009) that using simple rules to make strategic decisions is not only fast but also highly effective. Their findings suggest that organizations learn portfolios of heuristics for strategic decision making that contribute to their competitive advantage (Bingham & Haleblian, 2012). The authors concluded that "heuristics constitute 'rational' strategy in unpredictable markets" and can be "more effective than information-intensive, cognitively demanding approaches" (Bingham & Eisenhardt, 2011: 1438). Artinger et al. (2015) provided a conceptual review of several fast-and-frugal heuristics together with a discussion of their benefits and potential applications in management. Luan and Reb (2017) meanwhile demonstrated empirically that fast-and-frugal trees, an effective family of heuristics for binary decisions, are valid descriptive models of performance-based managerial decisions and that decision makers respond adaptively to changes in the base rates of a task when using them.

The above studies question the notion that logic and economic rationality are the universal gold standards of managerial decision making. An alternative is ecological rationality, where the rationality of using a heuristic or any other strategy is evaluated by its success in an uncertain world (Todd et al., 2012). This evaluation applies two perspectives: From a prescriptive perspective, researchers study the performance of a heuristic in different environmental conditions, which has implications for whether decision makers should use the heuristic and under what conditions; from a descriptive perspective, researchers examine whether decision makers actually use the heuristic and if so, whether they use it adaptively, based on the requirements of the task environment.

The Bias-Variance Dilemma

That leads to the question, under what conditions will heuristics perform better than more complex strategies? To answer this, it is useful to begin with the distinction between risk and uncertainty (e.g., Knight, 1921; Savage, 1954; Simon, 1990). In a situation of risk, the exhaustive and mutually exclusive set of future states are known and their consequences and probability distribution can be foreseen with certainty. In such situations of perfect knowledge, exemplified by lotteries, it is true that heuristics are generally second-best. Situations of uncertainty, in contrast, are defined by the absence of perfect foresight, where the full set of states, their consequences, and/or the probabilities are not known or knowable. Optimization is by definition impossible here, and heuristics can outperform complex strategies that try to fine-tune on past data (Gigerenzer & Brighton, 2009).

Specifically, *optimization* means to select a strategy that can lead to the best outcome in the future. Under uncertainty, even large amounts of historical data do not guarantee that a strategy that was optimal in the past will also be the best in the future. For instance, Google researchers analyzed some 50 million search terms to build Google Flu Trends, an algorithm for predicting influenza-related doctor visits. Other researchers, however, showed that using a single variable, the number of influenza-related doctor visits two weeks ago, predicted better than Google's big data algorithm (Lazer, Kennedy, King, & Vespignani, 2014).

Many strategic, investment, entrepreneurial, personnel, and other types of managerial decisions have to be made under uncertainty rather than risk (Artinger et al., 2015). Decisions are based on models that need to predict the future (e.g., the future performance of a job candidate), and where there is uncertainty, there will be prediction errors. According to the bias-variance decomposition of prediction error (e.g., Geman, Bienenstock, & Doursat, 1992), the prediction error (the sum of squared error) of a model is the sum of three separate components:

$$\text{Prediction error} = \text{bias}^2 + \text{variance} + \text{random error} \qquad (1).$$

Bias is the average difference between a model's predictions and the true status of an event and reflects how accurately a decision maker's model represents reality. *Variance* is a model's sensitivity to sampling error when a decision maker needs to estimate values of the model's free parameters in one sample and apply them for prediction in another sample (e.g., a manager develops a regression model of several cues and future job performance based on a sample of hired applicants and then uses the model to evaluate future applicants). Finally, *random error* is the irreducible and unavoidable error, independent of which model is used (for a more detailed exposition of this decomposition of prediction error, see Brighton and Gigerenzer [2012]).

The key insight from the bias-variance analysis of prediction error is that under situations of uncertainty, it is difficult for a model to have both a small bias and a small variance. Variance tends to be larger for more complex models that have a greater number of free parameters or of parameters whose precise values are difficult to estimate; the bias of such a model, by contrast, tends to be smaller. Less complex models, including heuristics, have the opposite tendencies. For instance, the 1/*N* heuristic, with which one allocates resources equally among *N* options, may be highly biased but has zero error due to variance because it has no free parameters and does not need to estimate anything from the past; thus, it often predicts better than highly complex allocation models in finance (e.g., Gigerenzer & Gaissmaier, 2011). This trade-off between bias and variance is known as the "bias-variance dilemma" (Geman et al., 1992).

Depending on how this fundamental trade-off plays out in a specific context, heuristics can perform better than more complex, seemingly rational strategies, especially under situations of uncertainty. However, because various methods of fitting, rather than predicting, have been predominantly used in studies of managerial decision making, complex strategies have (unintentionally) been shown to be superior to heuristics. This result is unfortunate, given that in the real world of managerial decision making, uncertainty is arguably very common and predicting the future is more important than fitting to the past.

The Δ-Inference Heuristic in Selection Decisions

To make the discussion more concrete, we now situate it within the context of selection decisions. Given the crucial role of human capital for organizational success, personnel decisions such as whom to hire, fire, or promote are among the most influential managerial decisions (Guion, 2011). Personnel selection features among the classic areas in industrial psychology, dating back to over a century ago (Münsterberg, 1912). Much of the research is applied, with the aim of helping organizations make better selection decisions, and a large amount of research has examined predictors of criteria such as job performance, adverse impact, and fairness perceptions, as well as methods to assess these (e.g., Ryan & Ployhart, 2014; Sackett & Lievens, 2008). Based on this research, meta-analytic studies have estimated cue validities for various predictors of job performance. A key finding is that the upper bound of predictability is at around 30% of the variance (Schmidt & Hunter, 1998). As such, uncertainty is rampant in this context, although many practitioners fail to consider it appropriately (Highhouse, 2008).

In contrast to research on cue validities (i.e., *what* cues to use), research on cue integration (i.e., *how* to use cues) has received relatively little attention (Kausel et al., 2016). When there are multiple cues that managers could use concurrently to make a personnel decision, existing research has largely applied regression analysis and thus implicitly assumed that managers decide on the basis of a compensatory weighting-and-adding strategy. At the same time, however, research also suggests that actual recruitment decisions are not made in this way (Highhouse, 2008). All in all, personnel selection provides an ideal context for our study because it is among the most crucial managerial decisions, involves substantial uncertainty, and allows us to address the important yet poorly understood issue of the process of cue integration.

Consistent with others, we study selection decisions between two final candidates (e.g., Kausel et al., 2016). Such decisions are also referred to as "paired comparisons," in which one chooses between two options on the basis of multiple relevant cues. The cues we focus on are three of the most commonly used and most valid predictors of job performance suggested by meta-analytic research (e.g., Farr & Tippins, 2010; Schmidt & Hunter, 1998): general mental ability (GMA), conscientiousness (CON), and structured interview performance (SIP).[1] The standard strategy to predict a binary dependent variable is logistic regression, a method that considers all available cues and estimates a beta weight for each.

A general heuristic for paired-comparison tasks is called *Δ-inference* (Luan, Schooler, & Gigerenzer, 2014). The heuristic can be described by three rules:

1. *Search*: Examine cues in the order of their importance or validities.
2. *Stopping*: If the difference between a pair of options on a cue exceeds a threshold value Δ, then stop search.
3. *Decision*: Choose the option with the higher (lower) cue value if higher (lower) cue values are more desirable. If no difference exceeds Δ for all cues, then restart the search from the first cue and make a decision as soon as any difference is found between the options (i.e., setting Δ to zero).

Unlike logistic regression, the Δ-inference heuristic is lexicographic. This means that the process is *sequential*, searching cues one after another instead of considering all cues at once. It is also *noncompensatory*, meaning that a decision is made based on the cue that stops search and subsequent cues in the search hierarchy have no effect on the decision, that is, their values cannot compensate for the values of the decisive cue. Finally, the heuristic is *frugal*, meaning that on

[1] We acknowledge that there are other valid cues, such as work samples, biodata, and integrity tests, and that cues used in practice often depend on the stage in the selection process. However, to reduce the complexity of the investigation carried out in the three studies reported in this article, we decided to limit our attention to these three cues. Also, because our interest is in cue integration, we do not discuss the methods used to generate cue values and present cue values as given in the simulations (Study 2) and to our research participants (Study 3).

average, it looks up fewer cues than are available.

Lexicographic heuristics have been studied in several areas of decision making, including consumer behavior and risky choices (e.g., Bettman, Johnson, & Payne, 1990; Kohli & Jedidi, 2007; Tversky, 1969), and the evidence generally shows that people often make decisions in such a sequential, noncompensatory manner. Luan and colleagues (2014) found that Δ-inference led to the same level of predictive accuracy as did regression and other complex models in 39 real-world tasks, such as predicting which professor earns a higher salary or which car has a better fuel efficiency. However, no studies have examined Δ-inference in managerial decisions.

An Illustration

Logistic regression and Δ-inference are prototypical examples of "rational" and heuristic strategies for paired-comparison decisions. Let us illustrate how they may be used with a specific example. Imagine that a manager must make a hiring decision. After several rounds of screening, the top two candidates are left. In an effort to practice evidence-based management, our manager considers a set of valid cues that predict future job performance (FJP) of these two candidates: their GMA, CON, and SIP scores, each of which correlates positively with FJP. Figure 1 (taken from Study 3) shows their scores on these cues: Clearly, no candidate dominates the other. How should the manager integrate these cues to arrive at a decision?

--- Figure 1 around here ---

From the perspective of "more information is always better," the prevailing view of managerial rationality would suggest a compensatory weighting-and-adding strategy such as logistic regression because it considers all information, allows for trade-offs among cue values, and maximizes. In this process, our manager would try to derive the weight of each cue, multiply the weight by the value of the cue, add up the products across all cues, and select the candidate with the higher score. If, in contrast, our manager uses Δ-inference, cue use would be sequential and noncompensatory. Assuming that the manager ranks cues based on their validities derived from a meta-analysis, GMA would be considered first. If the difference in GMA score between the two candidates is deemed sufficiently large (i.e., surpasses the difference threshold), the manager would choose the candidate with the higher score, without even considering the other two cues. Only when the difference is smaller than the threshold would the manager move on to consider the next-ranked cue, and so on.

Logistic regression and Δ-inference are at the center of our investigations carried out in three studies. We now describe these studies, including the goal, predictions, and results of each.

STUDY 1

In this study, we compared the prediction performances of Δ-inference and logistic regression in a real-world data set, a common approach in research investigating ecological rationality (e.g., Czerlinski, Goldstein, & Gigerenzer, 1999; Luan et al., 2014; Marewski & Schooler, 2011). The goal was to examine whether managers using the fast-and-frugal Δ-heuristic would make selection decisions that are as good as—or even better than—those of managers using logistic regression. In addition, Study 1 provided an initial exploration of the role of task environment with respect to learning opportunities. Our expectation was that Δ-inference would predict better when learning opportunities are limited, resulting in greater uncertainty in the task.

Methods

Data set. The data set was taken from Study 1 in Kausel et al. (2016) and includes data from 236 actual applicants at an airline company. Each applicant did GMA and CON assessments and received an unstructured interview performance (USIP) score by a line manager. All applicants in this data set were eventually hired and were assessed by their supervisors on their overall job performance approximately three months later. Table 1 shows the key statistical properties of the four relevant variables, FJP, GMA, CON, and USIP. Among the

236 individuals in the data set, there were 25 unique values in FJP. By exhausting all pairs of individuals with different FJP scores—so that the correctness of a paired-comparison decision could be unambiguously established—we ended up with a total of 50,334 pairs. These pairs served as the database from which random samples were drawn in our subsequent analyses.

--- Table 1 around here ---

Model testing and strategy performance. To measure a strategy's performance, we used cross-validation to assess its accuracy in predicting which of two job candidates will have a better FJP score and thus should be hired. Cross-validation is one of the most commonly applied model-testing methods in statistics, machine learning, and cognitive sciences (e.g., Czerlinski et al., 1999; Geisser, 1993; Stone, 1974). Operationally, in a sample consisting of *n* cases (e.g., paired comparisons), a certain proportion are used to "train" a model, estimating the model's free parameters (e.g., the beta weights in logistic regression), and the remaining cases are used to "test" the model's prediction accuracy (e.g., how often it chooses the better job candidate), with parameter values learned from the training cases.

Δ-inference and logistic regression needed to learn, or estimate, very different sets of parameters. For Δ-inference, the parameters were cue search order, which was estimated by calculating the bivariate correlations between the three cues and the decisions and then ordering the correlations by their absolute magnitudes (Luan et al., 2014), and the three threshold values, one for each cue; thus, adding up to four parameters in total. For logistic regression, we assumed no interactions among the three cues, meaning that there were also four parameters to be estimated: the beta weights for the three cues and an intercept term.

Learning opportunities. We varied learning opportunities in two ways. First, there are many ways to conduct cross-validation, depending on how training and testing cases are split in a sample. We applied three splits in this study: 50-50, 60-40, and 80-20, in which 50%, 60%, and

80% of a sample were used respectively for training. Second, we tested the strategies with three sample sizes: 30, 100, and 1,000, which represent situations where learning opportunities are generally few, moderate, and abundant, respectively. In each sample, the three splits of cross-validation were applied, resulting in a 3 × 3 factorial design with learning opportunities ranging from 15 cases (50% of 30) to 800 (80% of 1,000). To obtain reliable results on performance of the two strategies, 10,000 random samples were drawn from the paired-comparison database in the n = 30 and n = 100 conditions, whereas 1,000 were drawn for the n = 1,000 condition.

In general, our analysis can be situated in the context in which a manager first tries to learn the parameters of a model by observing or making a number of decisions with feedback and then proceeds to apply the model to make more decisions without feedback. We essentially tested and compared the predictive accuracy of two types of managers, one using logistic regression and the other using Δ-inference. In nine learning conditions, we examined which manager would predict more accurately in the real-world data set investigated in this study.

Results and Discussion

Figure 2 shows the prediction accuracy of Δ-inference and logistic regression in each (sample size) × (training proportion) condition. Two general patterns can be observed: (1) each strategy became more accurate when provided with more learning opportunities in terms of both a larger sample size and a higher proportion of training cases, and (2) Δ-inference achieved higher prediction accuracy than did logistic regression in all conditions. The difference between the two strategies was especially pronounced when there were generally few opportunities for learning and decreased as learning opportunities became more abundant.

--- Figure 2 around here ---

Despite the all-around superior predictive accuracy of Δ-inference, one should note that neither strategy predicted well: Even with many opportunities for learning, the highest prediction accuracy remained below 63%. This certainly has something to do with the relatively low predictive validities of the three cues in the data set (Table 1) and is consistent with meta-analytic evidence suggesting that FJP is very difficult to predict (e.g., Schmidt & Hunter, 1998).

In addition to being more accurate in prediction, Δ-inference on average searched fewer than 1.5 cues to make a decision, compared to all three cues used by logistic regression, and the fewer the learning opportunities, the fewer cues it searched. This result is highly relevant to practice because assessing job applicants' GMA, CON, and especially USIP is costly and time-consuming. In the data set studied here, USIP had the lowest validity and was rarely searched by Δ-inference. Thus, using Δ-inference would not only lead to higher predictive accuracy but also save managers cost and time, making it a better strategy across the board[2].

In sum, in an ecologically valid, real-world data set, Study 1 provides an existence proof that Δ-inference can lead to better decisions than logistic regression while searching less information (i.e., "less is more"). The performance advantage of the heuristic was particularly large under conditions of high uncertainty due to limited learning opportunities (i.e., smaller sample sizes and fewer training trials), conditions that are common to many personnel selection decisions and real-world managerial decisions in general.

STUDY 2

The goal of Study 2 was to examine in more detail the ecological conditions under which Δ-inference is likely to outperform logistic regression, and vice versa. Based on the bias-variance analysis of prediction error, we made the following two predictions on the relative performance

[2] The detailed frugality, cue search, and additional model-testing results can be found in the Supplementary Materials. Because the validity of USIP was so low in this data set, we tested Δ-inference and logistic regression with only the GMA and CON cues. The two-cue models did have higher predictive accuracy when sample size was small; however, the improvements were generally limited and Δ-inference stood to benefit even more than logistic regression. Lastly, we ran three popular machine learning algorithms, LASSO regression, random forest, and support vector machine (SVM), in our data set. None of the three algorithms predicted more accurately than Δ-inference in any of the learning conditions. These results show that for tasks of high uncertainty, even top-of-the-line machine learning algorithms may not outperform simple heuristics. We thank two anonymous reviewers for suggesting these additional analyses.

of the two strategies:

1. The smaller the sample size, the larger the relative advantage of Δ-inference. Sample size affects the variance component of prediction error in that the smaller the sample size, the larger the error due to variance. However, this is typically less of a problem for lexicographic heuristics than for models that try to integrate all available information (e.g., Brighton & Gigerenzer, 2012).
2. The more skewed the distribution of cue validities, the larger the relative advantage of Δ-inference. Lexicographic heuristics, including Δ-inference, often rely on only the first cue to decide. If the first cue is substantially more useful than others, then with regard to the bias component of prediction error, Δ-inference and logistic regression will be similarly biased (e.g., Gigerenzer, 2016; Martignon & Hoffrage, 1999). This will increase the overall advantage of Δ-inference, which generally has less variance than logistic regression.

Besides sample size and distribution of cue validities, we also investigated the effects of two other environmental properties, described further below. These ecological investigations were carried out in 1,728 simulated task environments, and the statistical parameters in those environments (i.e., cue validities and intercue correlations) were chosen according to the results of a meta-analytic study of personnel selection (Bobko, Roth, & Potosky, 1999).

Methods

Task environments. The task in each simulated environment was to choose which of two job candidates would have better FJP based on the candidates' scores on GMA, CON, and standardized interview performance (SIP). Unlike Study 1, we used SIP here because of its higher validity for predicting job performance (Schmidt & Hunter, 1998).

Table 2 shows a correlation matrix that contains the values of six parameters critical to a simulated environment: the validity of each cue (a, b, and c) and the intercue correlations (d, e, and f). These values were taken from Table 1 reported in a meta-analysis study by Bobko and colleagues (1999). The last column in Table 2 lists the parameter values we used to construct the simulated environments. There were four levels for each cue validity parameter: its meta-analytic value, the plus and minus .10 of this value, and a fixed value of .05 that renders the cue close to being useless (similar to USIP in the Study 1 data set). For each intercue correlation parameter, three levels were included: its meta-analytic value and the plus and minus .10 of this value. A total of 1,728 combinations can be formed with these parameter levels, and each combination provided values on the basis of which parameters of a simulated environment were set.

--- Table 2 around here ---

Each simulated environment was specified by a multivariate Normal distribution with four variables, a criterion (FJP) and three cues (GMA, CON, and SIP). The variance of each variable was set to 1, and the six pairwise correlations among the four variables were given by one of the 1,728 combinations of parameter values. To create a sample of n paired comparisons, we first randomly drew $2n$ cases (i.e., job candidates) from the multivariate Normal distribution and then paired the i^{th} case ($i = 1$ to n) with the $(i + n)^{th}$ one. Whichever of the pair had a higher criterion value was the correct choice. Environments simulated with this procedure are linear, meaning that the best model for predicting the criterion should be a linear combination of the cues (Hogarth & Karelaia, 2007). Therefore, a linear strategy such as logistic regression should have an inherent advantage, however small, over nonlinear ones with respect to the bias component of prediction error (i.e., be less biased).

Environmental properties. We varied four environmental properties. The first is *sample size*, which limits the amount of learning available to a decision maker to estimate a strategy's free parameters and directly affects the variance component of prediction error. The second is the *distribution of cue validities*. Two types were distinguished: (a) J-shaped (coded as "1"), where the highest cue validity ρ_1 is higher than the other two to the extent that $\rho_1 > (\rho_2 + \rho_3)$, and (b) not J-shaped (coded as "0"), where validities are distributed otherwise. As described above, we predicted that, relative to logistic regression, Δ-inference should perform better when sample size is small and in environments where the distribution of cue validities is J-shaped.

The third property is *linear predictability*, which is defined as the R^2 of the best linear regression in an environment and is a critical environmental property to a lens model analysis (e.g., Cooksey, 1996). We measured linear predictability by first simulating 1,000,000 cases in an environment and then getting the R^2 of the linear regression that used the three cues as predictors of the criterion variable. Because each environment was linear in this study, linear predictability represents how predictable an environment was when the theoretically best model was used. Hogarth and Karelaia (2007) showed that the performance of both linear and nonlinear strategies increases in environments with higher linear predictabilities but that the direction of their relative performance is ambiguous.

Finally, we varied the *best cue's relative predictiveness*, which is the ratio between the R^2 of a linear regression using only the best cue and the linear predictability in an environment. It represents the amount of information contained in the best cue relative to others and should be higher in J-shaped environments and in environments where intercue correlations are higher. When the best cue's relative predictiveness is higher, Δ-inference does not need to search much further beyond the best cue to reach a decision (Luan et al., 2014); however, as with higher linear predictability, it is unclear how different types of strategies would perform relative to each other in such environments.

Of the four properties, sample size does not depend on the statistical characteristics of a simulated environment, whereas the other three do and their values vary across environments.

Among these three properties, distribution of cue validities is highly correlated with relative predictiveness of the best cue (r = .73) and negatively correlated with linear predictability (r = −.26), which is also negatively correlated with the best cue's relative predictiveness (r = −.32).

Model testing. As in Study 1, we tested the predictive accuracy of logistic regression and Δ-inference using cross-validation in three sample-size conditions: 30, 100, and 1,000. However, instead of testing different proportions of training cases in a sample, we applied a fixed 60-40 split in this study. In each environment, the two strategies' performances were based on 10,000 random samples in the n = 30 and n = 100 conditions and 1,000 in the n = 1,000 condition.

Because we tested only two strategies and were concerned exclusively with their relative performance in this study, the main measure we assessed is the relative frequency of logistic regression predicting better than Δ-inference across all samples in a sample-size condition. For example, there were 10,000 samples in the n = 100 condition. In a specific environment, suppose that the prediction accuracy of logistic regression was higher than that of Δ-inference in 5,000 samples, lower in 4,000, and tied with it in the other 1,000 samples. The relative frequency was calculated by adding half of the frequency when the two were tied to the frequency that logistic regression was truly better. In the above example, it is then .50 + .50 × .10 = .55.

Results and Discussion

Table 3 shows the mean relative frequency of logistic regression predicting better than Δ-inference across all simulated environments in each sample-size condition. When sample sizes were smaller (i.e., 30 and 100), the mean frequencies were below .50 (.44 and .49, respectively), meaning that logistic regression on average predicted less accurately than Δ-inference when learning opportunities were limited, consistent with our finding in Study 1. When sample size was very large (i.e., 1,000), the performance of each strategy approximated its maximum level; there, logistic regression finally became the generally more predictive model (mean relative frequency = .65)[3]. Overall, the results support our prediction that Δ-inference should perform relatively better when sample size is small and learning opportunities are limited.

--- Table 3 around here ---

Table 3 also shows how the other three environmental properties affected the strategies' relative performance. First, on average, the relative frequency of logistic regression predicting better than Δ-inference was lower in J-shaped environments, although the difference became smaller as sample size declined. This result generally supports our prediction regarding the effect of cue validity distribution.

Second, because both linear predictability and relative predictiveness of the best cue are continuous variables, we calculated the bivariate correlation between each and the relative frequency of logistic regression predicting better[4]. In each sample size condition, the correlation was negative for the best cue's relative predictiveness, suggesting that when useful information concentrated more in the best cue, logistic regression tended to perform relatively worse than Δ-inference. In contrast, the correlation was positive for linear predictability, indicating that when FJP was more predictable by a linear combination of the three cues, logistic regression tended to perform relatively better. Figure 3 displays the scatter plots of the relative frequency of logistic regression predicting better against these two properties in the n = 100 condition.

--- Figure 3 around here ---

The last column of Table 3 reports the mean frugality of Δ-inference (i.e., average number

[3] This result was expected because all the environments are linear and linear models should predict best, given enough learning. We also simulated some environments in which decisions were outcomes of a lexicographic process of the three cues. There, Δ-inference outperformed logistic regression regardless of the sample size, and its relative advantage was influenced by the threshold value set in each cue to stop searching, the amount of random noise on the thresholds, and the pairwise correlations among the cues. A description of these environments and a summary of the results of our ecological analysis can be found in the Supplementary Materials.

[4] We also calculated the correlation of each property when controlling for other properties. Values of these partial correlations differed only slightly from those of the bivariate correlations, and the pattern of results remains the same as the one shown in Table 3.

of cues searched) across all environments in each sample-size condition. It shows that Δ-inference not only searched on average less than half of the available cues but also searched fewer cues when sample size was smaller, consistent with our findings in Study 1. We think that this is a "smart" way for Δ-inference to deal with the high level of noise in small sample situations. Specifically, sparse learning makes it difficult for Δ-inference to estimate exact cue validity values and then the correct cue search orders. Paradoxically, this does not hinder—and sometimes even facilitates—its ability to identify the best cue (e.g., Katsikopoulos, Schooler, & Hertwig, 2010; Şimşek & Buckmann, 2015). To reduce overall error, Δ-inference can thus set a small threshold on the best cue, relying more on it to make decisions and searching less.

In summary, we conducted an analysis of the ecological rationality of Δ-inference and logistic regression in 1,728 simulated task environments, whose parameters cover what is likely to be encountered in real-world personnel selection tasks. Consistent with our predictions, we found that, relative to logistic regression, Δ-inference predicts better when sample size is smaller and the distribution of cue validities is skewed (J-shaped). Explorations of two other environmental properties show that Δ-inference is more likely to outperform logistic regression when the best cue is particularly useful, whereas the opposite tends to occur when the linear predictability of a task is higher. Furthermore, despite the linearity of all the environments, which imposes a handicap on Δ-inference, Δ-inference on average predicted more accurately than logistic regression except when sample size was very large and did so by searching less than half of the available cues. This provides another demonstration of the "less-is-more effect" and further evidence that Δ-inference is a useful and effective heuristic for personnel selection and potentially other managerial decisions.

STUDY 3

Results of Studies 1 and 2 suggest that under many circumstances, managers should use the

Δ-inference heuristic rather than logistic regression to predict which of two job applicants will show better future job performance. Extending these findings, we now turn our focus from prescription to description, examining actual decision processes in an experimental setting. In the experiment, participants made a series of selection decisions similar to those in Studies 1 and 2. We asked them to decide on candidates for two job positions using a within-subjects design, which allowed us to test how frequently their strategies were consistent with Δ-inference or logistic regression in each condition, whether they adjusted their strategies to the features of a task, and how they might switch strategies between conditions.

Unlike in the previous two studies, participants' decisions made in this study could not be judged as right or wrong; thus, it was impossible to judge whether their strategies were ecologically rational or not. Even so, we measured participants' previous experience in selection decisions, expecting that the more experienced ones would behave more similarly to an ecologically rational or adaptive decision maker. Some previous studies have shown that more experienced decision makers tend to adopt heuristics more frequently (e.g., Luan & Reb, 2017; Pachur & Marinello, 2013; Wegwarth et al., 2009) and apply strategies more selectively across different task conditions (e.g., Rieskamp & Otto, 2006). Whether this would hold in the present study was an important question in our investigation and could provide clues as to what adaptive behaviors would look like when managers use Δ-inference or logistic regression.

Methods

Participants. In order to have sufficient variation in selection decision experience, we recruited three groups of students at a management university in Southeast Asia: (1) first- or second-year undergraduates who were taking introductory management courses (N = 101), (2) third- or fourth-year undergraduates majoring in organizational behavior and human resources and taking a course on personnel selection (N = 37), and (3) part-time master's students in a master of human capital leadership program with typically five or more years of working experience in an HR function (*N* = 28). Undergraduate students participated in exchange for partial course credit, whereas master students received no credit. Given that experience with selection decisions varied substantially within each participant group and the uneven group sizes, we collapsed data across the three groups and used self-reported experience rather than data source as the grouping variable.

Out of the 166 participants, 23 were excluded from data analysis for one or more of the following reasons: (1) answering "*no, because I was distracted and did not pay full attention*" to the question "*in your honest opinion, should we include your responses in our study*?" that was asked at the end of the experiment (*N* = 11); (2) taking less than 15 minutes to complete the experiment, which we consider abnormally short (*N* = 4); and (3) not selecting the job candidate who scored better than the other on all three cues more than once (*N* = 10). Our final sample thus consisted of 143 participants (85 female, 59.4%) with a mean age of 24.2 (*SD* = 6.5).

Design and procedure. Participants were informed that the purpose of this study was to understand how people make recruiting decisions based on their judgments of candidates' qualifications. They were instructed to assume the role of an HR manager in a multinational corporation and were provided with information on job candidates' GMA, CON, and SIP scores. They were asked to make decisions on the basis of these cues with the assumption that the two candidates scored similarly on all other relevant qualifications and characteristics.

A within-subjects design was applied: Each participant was asked to recruit for two different positions: data analyst (a more complex job) and receptionist (a less complex job). In each job condition, participants were instructed to read a description of the required responsibilities for the position before engaging in 105 paired-comparison decisions one by one. The description was a minimally edited version of a real job description for either a data analyst or a receptionist position posted in a popular job search website and can be found in Appendix A. After making their decisions, participants were asked to judge the importance of each cue for hiring the best person for the position, doing so by distributing 99 points among the three cues. After that, they moved on to make decisions for the second position and then again judged cue importance. The orders of the two job conditions were randomized for each participant.

Materials. In each experimental trial, participants viewed two job candidates' scores on GMA, CON, and SIP side-by-side and were asked to select the one they preferred to hire (see a sample display in Figure 1). A brief definition of each cue was provided and could be seen on screen in each trial. Presentation orders of the cues were first determined randomly and then remained fixed throughout the whole experimental session for each participant; the cue values were generated by a computer program. In the receptionist condition, the program drew values from three Normal distributions, whose means and standard deviations were 100 and 15 for GMA, 50 and 10 for CON, and 3.65 and 0.52 for SIP; the intercue correlations were set to 0 between GMA and CON, 0.24 between GMA and SIP, and 0.12 between CON and SIP. These parameter values were adopted from results reported in the literature (Bobko et al., 1999; McCrae, Martin, & Costa Jr., 2005; Roth, Switzer, Van Iddekinge, & Oh, 2011). In the analyst condition, the only difference was that the mean was 120 instead of 100 on GMA (Schmidt & Hunter, 2004). Information on the mean, the lowest, and the highest scores of each cue was also displayed on screen in each trial.

In each condition, we created 105 pairs of candidates using the computer program. Among them, 100 were results of pairing 100 program-generated candidates with another 100, and five were created so that one candidate dominated the other, that is, had better values on all three cues. Participants' decisions in these five pairs provided one way for us to check whether they had paid attention in the experiment. After creating the 105 pairs, five and 100 were selected respectively to make up the practice and experimental trials; the display orders of the pairs in each block were randomized for each participant.

Measures. Participants' choices and reaction times in each trial were recorded. At the end of each experimental condition, we asked participants for their ratings on the importance of each cue for a position, and at the end of the experimental session, besides requesting demographic information, we asked participants whether they had ever been in a position to formally recruit others as part of their job and in how many selection decisions they had previously been involved (1: 0-3; 2: 4-12; 3: 13-24; 4: 25-36; 5: >36). Finally, participants were asked whether they thought their responses in the experiment should be included in our analyses.

Model testing. To test which strategy—logistic regression or Δ-inference—a participant was more likely to adopt in an experimental condition, we applied the same method as used in Luan and Reb (2017). As in Studies 1 and 2, cross-validation was core to this method. However, in addition to investigating the accuracy of each model's predictions of a participant's choices, we also considered the model's predictions of a participant's reaction times (RT). In essence, the method is a modified version of the multiple-measure maximum likelihood method by Glöckner (2009) and tested how well each model could predict a participant's choice and RT, by estimating the conditional likelihood of the data given the model.

In the 100 decisions made by a participant in an experimental condition, we estimated parameters of logistic regression and Δ-inference in the first 60 trials (i.e., the training cases) and examined the models' predictions in the next 40 trials (i.e., the testing cases). The models were compared in terms of their maximum likelihoods. For logistic regression, seven parameters were estimated: four linear terms (i.e., three beta weights and the intercept term), one error rate in applying the model, and two parameters for RTs (i.e., the mean and standard deviation). For Δ-inference, eight were estimated: cue order, threshold values on the three cues, error rate, and three RT parameters (i.e., mean, standard deviation, and a scaling parameter). Rationales for why parameters for error rate and RTs were needed can be found in Glöckner (2009).

Model testing and comparison were always conducted at the individual level. For each participant and in each experimental condition, we identified the model that had the larger maximum likelihood in prediction as the one more likely adopted by the participant.

Results and Discussion

Reaction times. We started our analysis by inspecting participants' RTs and found that it sometimes took participants an exceptionally long or brief time to complete a trial. To reduce the effects of these trials on our analyses, we calculated the mean and standard deviation (*SD*) of each participant's RTs across all trials in an experimental condition and replaced RTs longer and shorter than 2.5*SD*s with mean plus and minus 2.5*SD*, respectively. The mean and *SD* of a participant's RTs were re-calculated after this treatment. Table 4 shows the percentage of abnormal RT trials, the mean RT, and the *SD* of RT, all averaged over all participants, in each experimental condition[5].

--- Table 4 around here ---

Cue importance. How important was each cue to the participants when they made hiring decisions? Table 4 shows the average ranks of the three cues based on participants' subjective ratings of cue importance for each job position. By this aggregate measure, the orders were CON > SIP > GMA for the receptionist position and GMA > CON > SIP for the analyst position. They are consistent with results from model testing (see Supplementary Materials) and show that the importance of GMA depended heavily on the job position: It was the most important cue for

[5] To test how robust our model-testing results are against abnormal RTs, we analyzed our data without the RT treatment. The results that match Figures 4 and 5 reported below can be found in the Supplementary Materials. In general, leaving abnormal RTs untreated affected some aspects of the results, albeit without changing the main conclusions of our study. We also performed robustness checks on our main results by adding back the excluded participants' data. These results can also be found in the Supplementary Materials.

the more complex analyst position but the least important cue for the less complex receptionist position. Meanwhile, CON was deemed as an overall important cue for both positions.

Did participants use Δ-inference and under what conditions? The last rows of Table 4 show the proportions of participants who were classified as using either Δ-inference or logistic regression. In the receptionist condition, almost half (49%) used Δ-inference, but that proportion dropped to 38% for the more complex analyst condition. Thus, many participants did adopt Δ-inference to make decisions, but their preference for the strategy depended on the job position.

To further understand their strategy selections, we divided participants into two categories using the median of previous experience: those who had been involved in four or more selection decisions ($N = 64$) and those who had been involved in fewer ($N = 79$). Moreover, within each job condition, we distinguished two types of participants by the distribution of their subjective ratings of cue importance: those with skewed ratings such that the highest rating was more than the sum of the other two, that is, $r_1 > (r_2 + r_3)$, and those with more equal ratings, that is, $r_1 \leq (r_2 + r_3)$. A skewed distribution here is the subjective version of a J-shaped environment in Study 2; participants with such a distribution were in the minority in both the receptionist and the analyst conditions: $N = 30$ and 33, respectively. Figure 4 shows the proportion of participants classified as using Δ-inference in each (experience) × (cue rating distribution) category for each condition.

--- Figure 4 around here ---

Study 2 indicated that a J-shaped distribution of cue validities is an environment condition under which Δ-inference has a relative performance advantage over logistic regression. Figure 4 shows that a higher proportion of participants were classified as using Δ-inference when the distributions of their cue importance ratings were skewed, a result that held for both the less and the more experienced participants—and particularly so for the latter—in both job conditions. This suggests that our participants were more likely to use Δ-inference when the heuristic was a prescriptively better strategy than logistic regression.

Figure 4 also shows that a higher proportion of the more experienced participants were classified as using Δ-inference, and this was the case in both job conditions and regardless of whether participants' cue importance distributions were skewed or not. This result is consistent with previous findings that the use of heuristics is often positively related to experience in a domain and experts are more likely than novices to use heuristics (e.g., Garcia-Retamero & Dhami, 2009; Wegwarth et al., 2009). Furthermore, of all the participants, those with more experience and a skewed cue importance distribution adopted Δ-inference most frequently, whereas those with less experience and a more equal distribution adopted it least frequently. This suggests that the effects of experience and cue importance distribution on strategy selection could be additive.

Finally, we ran a logistic regression with a participant's classified strategy as the predicted variable and the participant's experience and type of subjective cue importance distribution as the predictors in each job condition. The results show that in both conditions, the beta weight of experience was statistically significant or close to significant (p = .046 and p = .082 for the receptionist and the analyst conditions, respectively). However, the beta weight of subjective cue importance distribution was not statistically significant in the receptionist condition (p = .261) but close to significant in the analyst condition (p = .057). The relatively small number of "skewed" participants in each condition might contribute to the nonsignificant results.

Experience and strategy switching. The within-subjects design of this study allowed us to examine whether and how participants switched strategies between the two job conditions. Figure 5 shows the rates of strategy switching from the receptionist condition to the analyst condition for the less and the more experienced participants. For both, more switched from Δ-inference to logistic regression than vice versa (recall that logistic regression was the more common strategy for the more complex analyst position). However, only a minority (33%) of the less experienced switched, whereas the majority (53%) of the more experienced did so. It is difficult to know exactly why the more experienced switched strategies more frequently. It is possible that through learning, they better understood the requirements of each job position and became more discerning as to how information in the cues should be integrated, leading to a more selective adoption of strategies. This pattern has also been observed in studies where the accuracy of decisions can be firmly established (e.g., Rieskamp & Otto, 2006).

--- Figure 5 around here ---

In sum, the results of Study 3 show that participants adopted both heuristic and weighting-and-adding strategies when making paired-comparison decisions in personnel selection and many of them adopted qualitatively different strategies in different task conditions. The results also suggest that participants were sensitive to a crucial condition for the ecological rationality of Δ-inference: a skewed (J-shaped) distribution of cue importance or validities. Specifically, participants, especially the more experienced ones, were more likely to use Δ-inference when they judged one cue to be much more important for the hiring decision than other cues. Moreover, compared to the less experienced participants, those with more experience were generally more likely to adopt the heuristic and switched strategies more frequently between job conditions. If the behavior of the more experienced is indeed closer to that of an ecologically rational, adaptive manager, then that manager would be more selective with regard to which strategy to use under which condition and more inclined to use the Δ-inference heuristic, particularly when deeming one cue as much more important or informative than others.

GENERAL DISCUSSION

A widely-held view in management research and teaching is that heuristics are inferior to “rational” strategies. Under the influence of the heuristics and biases program (Gilovich et al., 2002; Tversky & Kahneman, 1974), much management research has focused on how heuristics can lead to pernicious biases in areas such as performance appraisal, negotiation, personnel selection, portfolio investment, and strategy (e.g., Bazerman & Moore, 2008; Highhouse et al., 2013). It has been assumed that managers use heuristics because of their cognitive limitations or because of heuristics' advantage in saving search and processing costs, not because they can lead to more accurate decisions. Viewing heuristics from the perspective of ecological rationality, the present research challenges these assumptions and argues that under conditions of uncertainty, heuristics can lead to more accurate managerial decisions than rational strategies do while using less information; that is, less can actually be more.

Situating our studies within the context of personnel selection decisions, we compared Δ-inference, a fast-and-frugal heuristic, to logistic regression, a compensatory strategy that weights and adds all available information. In Study 1, we analyzed data from 236 applicants at an airline company and showed that Δ-inference was better than logistic regression at predicting which of two applicants would have superior future job performance. This effect held in all conditions (Figure 2) and was particularly strong when sample sizes were small. Study 1 thus provides an existence proof that a heuristic strategy can lead to more accurate predictions and decisions in a real-world personnel selection task.

Ecological rationality implies that the performance of a strategy depends on its fit to the task environment. In Study 2, we examined the effects of four environmental properties on the relative performance of logistic regression and Δ-inference in 1,728 simulated environments. Despite all environments being linear, we found that logistic regression performed worse than Δ-inference under a substantial set of conditions. In general, Δ-inference was more likely to predict better than logistic regression when (1) learning opportunities were limited, (2) one cue was substantially more informative than other cues, and (3) the criterion variable (i.e., future job performance) was less predictable by a linear model of cues.

Finally, we conducted an experiment in Study 3 to examine whether participants use Δ-inference and whether and how they adapt their strategies to the characteristics of the task. The analyses showed that many participants did adopt the heuristic and that participants tended to do so more often when they judged one cue as being much more important than the other cues, a condition identified in Study 2 as ecologically beneficial for Δ-inference. Both tendencies were particularly strong for participants with more experience in personnel selection.

Overall, findings from our studies are not consistent with the notion that heuristics are generally inferior, or second-best, to more complex strategies in managerial decision making. Instead, under conditions common in managerial decisions, heuristics can perform well and better than complex strategies, and decision makers seem to be sensitive to some of these conditions, using heuristics adaptively.

Our research makes several theoretical contributions. Most importantly, we introduce ecological rationality as a vision of managerial rationality. Economic rationality considers logic, probability theory, and maximization as the universal standards for good managerial decision making. Managers who violate these standards by using heuristics are often viewed as biased decision makers. Taking ecological rationality as a basis, we propose a more positive view on managerial heuristics: In addition to saving search and processing costs, using fast-and-frugal heuristics can also result in more effective and higher-quality decisions. We also propose a more balanced view on strategies traditionally thought of as “rational,” in that taking more information into consideration does not guarantee better decisions in situations of uncertainty. Our view is consistent with work on managerial intuition that rejects economic rationality as the universal standard (Dane & Pratt, 2007; Hodgkinson et al., 2009), albeit from a very different theoretical and methodological perspective.

Ecological rationality posits that the effectiveness of any strategy, heuristic or otherwise, depends on its fit to the task environment. We found that Δ-inference worked particularly well under conditions of uncertainty that are common in many managerial decision environments. This confirms recent research findings that managers use simple rules to make effective strategic decisions in tasks of great uncertainty (Bingham & Eisenhardt, 2011; Davis et al., 2009). Moreover, by emphasizing ecological fit, we move the conversation towards a contingency theory of managerial decision making. Similar to contingency theories of leadership, which argue that the effectiveness of a leadership style depends on the situation (Fiedler, 1964; Vroom & Jago, 2007), such a theory posits that there is no single best decision strategy and managers should use multiple strategies adaptively rather than relying on just one. The imperative for future research is to uncover these contingencies in managerial decision making.

Research on ecological rationality has already shed light on some of the general contingencies. In particular, as opposed to the assumption of a general accuracy-effort trade-off (Payne et al, 1993), ecological rationality emphasizes the distinction between uncertainty and risk and the resulting bias-variance trade-off (Brighton & Gigerenzer, 2012; Geman et al., 1992). *Variance* represents the sensitivity of a strategy to samples (or idiosyncratic learning experiences), and *bias* reflects the extent to which the strategy departs from reality. Due to uncertainty, simpler strategies tend to have a larger bias but a smaller variance than those of more complex strategies. Thus, the challenge in strategy selection is to strike a good trade-off between bias (complexity) and variance (simplicity). Lexicographic heuristics such as Δ-inference can reduce variance because of their simplicity, and if the distribution of cue validities is highly skewed, then they and a linear rule are similarly biased (Martignon & Hoffrage, 2002). Therefore, a manager working on tasks with this property can make more frugal *and* more accurate decisions by using lexicographic heuristics than by compensatorily weighting and adding all available information.

Our research also contributes to the literature on process models of managerial decision making (Luan & Reb, 2017). *As-if* models (Friedman, 1953), such as expected utility theory, prospect theory, or inequity aversion theory, are popular models of managerial decision making; however, they are meant to model the outcomes, not the process. Furthermore, many studies of managerial heuristics rely on qualitative labels provided by researchers (e.g., “availability”) or by managers themselves through qualitative interviews (Bingham & Eisenhardt, 2011; Manimala, 1992). Notwithstanding the valuable contributions of these approaches, a potential danger is that they allow researchers and practitioners to apply the labels flexibly (and sometimes incorrectly) to different processes, concluding that certain heuristics are used more commonly than they actually are. As an alternative, here we study heuristics as process models of decision making, specifying rules of how to search, when to stop search, and how to make a decision. With these specifications, heuristics can be implemented more easily in computer simulations and be tested in empirical settings more rigorously.

Moreover, our research adds to the literature on personnel selection by examining the processes through which cues are integrated in selection decisions. Past research in this area has been mainly interested in understanding the validities of different cues (Schmidt & Hunter, 1998) and has seldom examined how decision makers should integrate these cues (e.g., De Corte, 1999; De Corte, Lievens, & Sackett, 2007) or how they actually do so (e.g., Dougherty et al., 1986; Kausel et al., 2016; Lievens, Highhouse, & De Corte, 2005). In addition, the limited existing research has relied largely on regression models and optimization procedures, explicitly or implicitly assuming that managers integrate cues in a “rational” manner. Our research suggests that these assumptions may not be warranted and that personnel selection research should take heuristic process models into consideration to better understand and subsequently improve selection decision processes.

Furthermore, an ecological rationality approach helps clarify a confusion in the personnel selection literature that tends to equate heuristic processing with intuition. For example, selection researchers have pointed out practitioners' stubborn reliance on intuition and subjectivity in selection decisions (e.g., Highhouse, 2008). Heuristics, including Δ-inference, are not necessarily intuitive or subjective. Instead, by explicitly specifying the search, stopping, and decision rules, the objectivity and transparency of heuristics can be higher than those of complex strategies, whose processes and inputs (e.g., utilities) are often a black box and subject to interpretations. Thus, it is important to differentiate between how information on cues is sought (e.g., subjectively through unstructured interviews or mechanically through personality tests) and how the cues are being processed (e.g., subjectively through intuition or mechanically through well-specified heuristics or other rules; Gatewood, Feild, & Barrick, 2015).

The Methodology of Ecological Rationality

Novel research programs and paradigms often require different methods (e.g., research on organizational networks; Borgatti & Foster, 2003). Drawing on cognitive sciences, modeling, and statistics, research in ecological rationality has developed a set of methodologies to examine questions related to the performance and use of decision strategies. These methods are not currently typical in management research and may thus present a potential entry barrier for researchers interested in studying ecological rationality in organizations. Ultimately, however, we believe that the relative sophistication of these methods usefully complements existing methods, allows for the rigorous study of managerial decision making, and offers opportunities for researchers in other areas of organizational scholarship.

For example, building on recent research in model testing (e.g., Czerlinski et al., 1999; Glöckner, 2009), we applied a comparative model testing method in Study 3. Comparative model testing is widely applied in cognitive sciences owing to several advantages, in particular increased precision and reduced ambiguity. The difference between comparative model testing and testing only a single model is analogous to that of alternative hypothesis testing and the widely criticized practice of null hypothesis testing (Cohen, 1994). Moreover, we examined the performance of models in prediction rather than in fitting. Prediction is of more practical use than fitting—consider the value of foresight over hindsight. Prediction is also better at capturing a model's prescriptive and descriptive performance in an uncertain world, in which observations are limited, random noise is abundant, and true model parameter values may change dynamically, much like the decision environments managers commonly face when attempting to predict future states of their organization or their business environment (Gigerenzer & Brighton, 2009). Principles and methods for predictive and comparative model testing can be easily applied beyond managerial decision making to examine the performances of different entrepreneurial, collaborative, or investment strategies.

Practical Implications

Ecological rationality provides not only a novel theoretical perspective on managerial decision making but also novel practical implications. Often, advice is based on the notion that heuristics lead to biases. Managers have thus been warned of heuristics and their biases ("forewarned is forearmed"; Hammond et al., 1998: 58), on the perhaps naïve assumption that once decision makers know the dangers of heuristics and biases, they will change their thinking. Along the same line, decision makers have been urged to move from the unconscious and heuristic "System 1" to the conscious and analytical "System 2" to process information (Milkman, Chugh, & Bazerman, 2009). Finally, in the event that attempts to make decision makers think more "rationally" fail, policy makers and organizations have been encouraged to use "nudges" to protect people from their own decision-making incompetence (Thaler &

Sunstein, 2008).

An ecological rationality perspective, albeit not oblivious to the limitations of heuristics, rejects the view that analytical thinking is generally superior (Gigerenzer, 2008; Kruglanski & Gigerenzer, 2011). This research consistently found that a simple heuristic made more accurate personnel selection decisions when compared with a prototypical rational strategy. Importantly, this advantage became larger as the decision environment became arguably more typical of many managerial decisions, with more uncertainty and fewer learning opportunities. At the same time, it should be acknowledged that the accuracy advantages of the heuristic were sometimes small. That said, when the stakes are high (e.g., hiring the right executives) or when the same types of decisions are repeated many times, small increases in the probability of making the right decisions can mean large differences in the long run for an organization.

Even if the accuracies of both types of strategies are similar, heuristics tend to have substantial advantages in terms of frugality. Our studies showed that the Δ-inference heuristic needed to search on average less than half of the cues to make a decision. This means lower cue assessment and search costs and allows for quicker decisions, a desirable objective in managerial decision making (Baum & Wally, 2003). Additionally, as Simon pointed out, information processing consumes attention: "A wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it" (1971: 41-42). In this age of information explosion and attention overload, the value of "fast and frugal" decision making is becoming increasingly salient for managers.

Ecological rationality highlights the value of being adaptive: Managerial competence lies in applying the appropriate strategy given the task environment and the decision maker's objectives, such as accuracy, speed, frugality, or efficiency. Therefore, training programs should focus on helping managers develop their repertoire of heuristic and analytical decision strategies and apply them in an adaptive manner, informed by the decision context and purpose. Programs along these lines could include the explicit teaching of heuristics and their specific search, stopping, and decision rules, such as those in Δ-inference and fast-and-frugal trees (Luan & Reb, 2017), with the help of visualization programs (Phillips, Neth, Woike, & Gaissmaier, 2017). An advantage of learning heuristics over relying on intuition is that the rules of fast-and-frugal heuristics can be formulated and are transparent, whereas intuitive processes by definition are unconscious and thus lack transparency (Hogarth, 2001). In addition to training, selection and promotion could also be used to identify managers who flexibly and effectively use different decision strategies. This would likely require a shift from selection systems that prioritize analytical competence to systems that value adaptive decision making.

In the context of personnel selection, advice for practitioners has emphasized cue validities and the effects of cues on criteria such as performance and adverse impact. The rationale is that once researchers discover which cues managers should use, they can disseminate this information and managers will behave accordingly. However, the continued reliance on cues with questionable validities, despite decades of accumulated knowledge, casts doubt on the effectiveness of this approach (Highhouse, 2008). In its place, we suggest seeking a deeper understanding not only of the cues but of the decision strategies managers use as well as of their decision environments. On this basis, researchers and practitioners can co-develop interventions and decision aids, such as decision trees, that align with both managers' natural tendencies and their task environments. As a result, these strategies may be easier to adopt, more transparent, and more effective. Decision aids of this sort have been successfully developed in other fields, including medicine (e.g., Green & Mehr, 1997; Jenny et al., 2015) and the military (e.g., Keller & Katsikopoulos, 2016). Finally, the less-is-more principle can also be applied in job interview processes, given that under many realistic conditions, one good interviewer may be better than two or more because adding less capable interviewers is likely to detract from the performance of the best one (Fific & Gigerenzer, 2014).

Strengths, Limitations, and Future Research

This research has both strengths and limitations that point towards directions for future research. A strength of our studies lies in comparative model testing, which is often more insightful than examining only a single strategy. However, we recognize that our studies are limited by examining primarily two strategies, Δ-inference and logistic regression (but see the results of some other strategies we tested in Study 1 in the Supplementary Materials). As such, care needs to be exercised in extrapolating the current findings to other decision strategies. While we chose these two strategies because of their suitability for the current research setting, future research could test additional strategies, such as take-the-best (Gigerenzer & Goldstein, 1996), the recognition heuristic (Gigerenzer & Goldstein, 2011), the 1/*N* heuristic (Hertwig, Davis, & Sulloway, 2002), and more sophisticated machine learning algorithms.

Also, when testing strategies' prescriptive performance in Studies 1 and 2, we assumed that managers would learn parameters of the strategies efficiently without calculation errors, a typical assumption made in most studies involving simulations. In practice, however, this assumption is unlikely to hold, and managers are likely to encounter varying degrees of difficulty while learning different strategies. A strategy with parameters that are easier to learn and more robust against learning errors would be advantageous over others, even though it may not perform best in simulations. Research in ecological rationality that has studied strategy learning at both prescriptive and descriptive levels is rare (e.g., Rieskamp & Otto, 2006). To better understand the practical performance of Δ-inference and logistic regression, issues related to learning hence need to be addressed and studied in future research.

In our studies, we focused on paired-comparison decisions between two options: the final two candidates for a job position. Although this is consistent with previous research and organizational practices (e.g., Kausel et al., 2016), we need to be cautious when generalizing the present findings to other selection contexts, such as decisions about a larger set of candidates or a single candidate. Indeed, as discussed above, it is one of the foundations of ecological rationality that a heuristic's performance depends on its match with the task environment. We therefore neither suggest nor expect that Δ-inference will always perform well or be used for different decision tasks. In the course of selection, for example, tallying (Einhorn & Hogarth, 1975) may be used for initial screening of applicants, elimination-by-aspects (Tversky, 1972) for deciding among several candidates, Δ-inference for deciding between two final candidates, and fast-and-frugal trees (Luan & Reb, 2017) for deciding whether a single candidate is sufficiently qualified. More research will be needed to understand the most common and effective heuristics inside managers' "adaptive toolbox" of decision strategies for personnel selection.

An interesting extension of the present research on the sequential Δ-inference heuristic would be to multi-stage selection systems, which are also sequential in nature. Such systems have been studied largely from a prescriptive perspective, examining their effects on selection quality and adverse impact (e.g., De Corte, Lievens, & Sackett, 2006; Finch, Edwards, & Wallace, 2009; Roth, Bobko, Switzer, & Dean, 2001). This research also points to another limitation of the present studies: We limited our investigation to a single decision criterion (i.e., future job performance). In selecting employees, however, organizations may try to achieve multiple goals, including predicting contextual performance and counterproductive behaviors, reducing adverse impact on applicant groups (e.g., minorities), and increasing applicant fairness perceptions (e.g., Sackett & Lievens, 2008). Future research could examine simple heuristics that are suitable for multi-criteria decision making.

Some other features of our studies also suggest that caution is called for when generalizing the present results, even after we took steps to reduce such concerns. For example, although we investigated a real-world data set in Study 1 to increase ecological relevance, extending the analysis to other real-world data sets would be highly desirable to strengthen the generalizability of our findings. In Study 3, because of the requirement of a large number of decisions for model testing and cross-validation (e.g., Lewandowsky & Farrell, 2011), we chose a scenario-based experimental design, and many of our participants had limited experience in personnel selection decisions. We tried to address these concerns by using only slightly adapted job descriptions, basing our tasks on existing research on cue validities and cue intercorrelations (e.g., Roth et al., 2011), and recruiting participants with varied degrees of experience in personnel selection.

Finally, our studies were conducted in a personnel selection context only. Future research should examine heuristics in other types of managerial decisions, such as strategy, finance, and marketing, allowing for broader conclusions to be made about the ecological rationality and the effectiveness of heuristics in managerial decision making. The present research is best viewed as a stepping stone in the pursuit of this large endeavor.

Conclusion

In conclusion, in this research, we (a) propose ecological rationality as an alternative theoretical framework through which managerial heuristics should be viewed and studied, (b) challenge the common view in management research that heuristics are second-best through two prescriptive studies, (c) investigate in a descriptive study how decision makers integrate different cues in making decisions pertaining to personnel selection, and (d) introduce a set of novel methods to study the performance and processes of managerial decision making.

REFERENCES

Arkes, H. R., Gigerenzer, G., & Hertwig, R. 2016. How bad is incoherence?. ***Decision***, 3(1): 20–39.

Artinger, F., Petersen, M., Gigerenzer, G., & Weibler, J. 2015. Heuristics as adaptive decision strategies in management. ***Journal of Organizational Behavior***, 36(S1): S33–S52.

Baum, R. J., & Wally, S. (2003). Strategic decision speed and firm performance. *Strategic Management Journal*, *24*(11), 1107-1129.

Bazerman, M. H., & Moore, D. A. 2008. ***Judgment in managerial decision making***. New York: Wiley.

Beach, L. R., & Mitchell, T. R. 1978. A contingency model for the selection of decision strategies. ***Academy of Management Review***, 3(3): 439–449.

Bettman, J. R., Johnson, E. J., & Payne, J. W. 1990. A componential analysis of cognitive effort in choice. ***Organizational Behavior and Human Decision Processes***, 45(1): 111–139.

Bingham, C. B., & Eisenhardt, K. M. 2011. Rational heuristics: the 'simple rules' that strategists learn from process experience. ***Strategic Management Journal***, 32(13): 1437–1464.

Bingham C., Haleblian J. 2012. How firms learn heuristics: Uncovering missing components of organizational learning. ***Strategic Entrepreneurship Journal***, 6(2): 152–177.

Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. ***Personnel Psychology***, 52(3): 561–589.

Borgatti, S.P. and Foster, P.C. 2003. The network paradigm in organizational research: A review and typology. ***Journal of Management***, 29(6): 991–1013.

Brighton, H. and Gigerenzer, G. 2012. Homo heuristicus: less-is-more effects in adaptive cognition. ***The Malaysian Journal of Medical Sciences***, 19(4): 6–16.

Cohen, J. 1994. The earth is round (p < .05). ***American Psychologist***, 49(12): 997–1003.

Cooksey, R. W. 1996. ***Judgment analysis: Theory, methods, and applications***. Academic Press.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. 1999. How good are simple heuristics? In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), ***Simple heuristics that make us smart***: 97–118. New York, NY: Oxford University Press.

Dane, E. and Pratt, M.G. 2007. Exploring intuition and its role in managerial decision making. ***Academy of Management Review***, 32(1): 33–54.

Davis J. P, Eisenhardt K. M., Bingham C. B. 2009. Optimal structure, market dynamism, and the strategy of simple rules. ***Administrative Science Quarterly***, 54(3): 413–452.

Dean Jr, J. W., & Sharfman, M. P. 1996. Does decision process matter? A study of strategic decision-making effectiveness. ***Academy of Management Journal***, 39(2): 368–392.

Dean Jr, J. W., & Sharfman, M. P. 1993. Procedural rationality in the strategic decision-making process. ***Journal of Management Studies***, 30(4): 587–610.

De Corte, W., Lievens, F., & Sackett, P. R. 2006. Predicting adverse impact and mean criterion performance in multistage selection. ***Journal of Applied Psychology***, 91(3): 523–537.

De Corte, W., Lievens, F., & Sackett, P. R. 2007. Combining predictors to achieve optimal trade-offs between selection quality and adverse impact. ***Journal of Applied Psychology***, 92(5): 1380–1393.

De Corte, W. 1999. Weighing job performance predictors to both maximize the quality of the selected workforce and control the level of adverse impact. ***Journal of Applied Psychology***, 84(5): 695–702.

Dhami M. K. 2003. Psychological models of professional decision making. ***Psychological Science***, 14:175–80.

Dougherty, T. W., Ebert, R. J., & Callender, J. C. 1986. Policy capturing in the employment

interview. ***Journal of Applied Psychology***, 71(1): 9–15.
Einhorn, H. J., & Hogarth, R. M. 1975. Unit weighting schemes for decision making. ***Organizational Behavior and Human Performance***, 13(2): 171–192.
Farr, J. L., & Tippins, N. T. 2010. ***Handbook of employee selection***. New York: Routledge.
Fiedler, F.E. 1964. A contingency model of leadership effectiveness. ***Advances in Experimental Social Psychology***, 1: 149–190.
Fific, M., & Gigerenzer, G. 2014. Are two interviewers better than one? ***Journal of Business Research***, 67: 1771–1779.
Finch, D. M., Edwards, B. D., & Wallace, J. C. 2009. Multistage selection strategies: Simulating the effects on adverse impact and expected performance for various predictor combinations. ***Journal of Applied Psychology***, 94(2): 318–340.
Friedman, M. 1953. ***Essays in positive economics***. University of Chicago Press.
Garcia-Retamero, R. and Dhami, M.K., 2009. Take-the-best in expert-novice decision strategies for residential burglary. ***Psychonomic Bulletin & Review***, 16(1): 163–169.
Gatewood, R., Feild, H.S., & Barrick, M. 2015. ***Human resource selection***. Nelson Education.
Geisser, S. 1993. ***Predictive inference: An introduction***. Boston, MA: Springer US.
Geman, S., Bienenstock, E. and Doursat, R. 1992. Neural networks and the bias/variance dilemma. ***Neural Computation***, 4(1): 1–58.
Gigerenzer, G. 2016. Towards a rational theory of heuristics. In R. Frantz & L. Marsh (Eds.), ***Minds, models, and milieux: Commemorating the centennial of the birth of Herbert Simon***: 34–59. New York: Palgrave Macmillan.
Gigerenzer, G. 2008. Why heuristics work. ***Perspectives on Psychological Science***, 3(1): 20–29.
Gigerenzer, G., & Brighton, H. 2009. Homo heuristicus: Why biased minds make better inferences. ***Topics in Cognitive Science***, 1(1): 107–143.
Gigerenzer, G., & Gaissmaier, W. 2011. Heuristic decision making. ***Annual Review of Psychology***, 62: 451–482.
Gigerenzer, G., & Goldstein, D. G. 1996. Reasoning the fast and frugal way: Models of bounded rationality. ***Psychological Review***, 103(4): 650–669.
Gigerenzer, G., & Goldstein, D. G. 2011. The recognition heuristic: A decade of research. ***Judgment and Decision Making***, 6: 100–121.
Gigerenzer, G., & Selten, R. (Eds.). 2002. ***Bounded rationality: The adaptive toolbox***. MIT press.
Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). 2002. ***Heuristics and biases: The psychology of intuitive judgment***. Cambridge: Cambridge University Press.
Glöckner, A. 2009. Investigating intuitive and deliberate processes statistically: The multiple-measure maximum likelihood strategy classification method. ***Judgment and Decision Making***, 4: 186–199.
Green, L., & Mehr, D. R. 1997. What alters physicians' decisions to admit to the coronary care unit? ***Journal of Family Practice***, 45: 219–226.
Guion, R. M. 2011. ***Assessment, measurement, and prediction for personnel decisions***. New York, NY: Routledge.
Hammond, J. S., Keeney, R. L., & Raiffa, H. 1998. The hidden traps in decision making. ***Harvard Business Review***, 76(5): 47–58.
Hertwig, R., Davis, J. N., & Sulloway, F. J. 2002. Parental investment: How an equity motive can produce inequality. ***Psychological Bulletin***, 128(5): 728–745.
Highhouse, S. 2008. Stubborn reliance on intuition and subjectivity in employee selection. ***Industrial and Organizational Psychology***, 1(3): 333–342.
Highhouse, S., Dalal, R. S., & Salas, E. 2013. ***Judgment and decision making at work***. New

York, NY: Routledge.
Hodgkinson, G.P., Sadler-Smith, E., Burke, L.A., Claxton, G. and Sparrow, P.R. 2009. Intuition in organizations: Implications for strategic management. ***Long Range Planning***, 42(3): 277–297.
Hogarth, R.M. 2001. ***Educating intuition***. University of Chicago Press.
Hogarth, R. M., & Karelaia, N. 2007. Heuristic and linear models of judgment: Matching rules and environments. ***Psychological Review***, 114(3): 733–758.
Jenny, M. A., Hertwig, R., Ackermann, S., Messmer, A. S., Karakoumis, J., et al. 2015. Are mortality and acute morbidity in patients presenting with nonspecific complaints predictable using routine variables? ***Academic Emergency Medicine***, 22(10): 1155–1163.
Katsikopoulos, K.V., Schooler, L.J. and Hertwig, R. 2010. The robust beauty of ordinary information. ***Psychological Review***, 117(4): 1259–1266.
Kausel, E. E., Culbertson, S. S., & Madrid, H. P. 2016. Overconfidence in personnel selection: When and why unstructured interview information can hurt hiring decisions. ***Organizational Behavior and Human Decision Processes***, 137: 27–44.
Kausel, E. E., & Slaughter, J. E. 2013. Employee selection decisions. In Highhouse, Dalal, & Salas (Eds.) ***Judgment and decision making at work***: 77–99. Routledge.
Keller, N., & Katsikopoulos, K. V. 2016. On the role of psychological heuristics in operational research; and a demonstration in military stability operations. ***European Journal of Operational Research***, 249(3): 1063–1073.
Knight, F. H. 1921. ***Risk, uncertainty and profit***. Boston, MA: Hart, Schaffner & Marx; Houghton Mifflin Co.
Kohli, R., & Jedidi, K. 2007. Representation and inference of lexicographic preference models and their variants. ***Marketing Science***, 26(3): 380–399.
Kruglanski, A., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review, 118,* 97–109. doi:10.1037/a0020762
Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). Big data. The parable of Google Flu: Traps in big data analysis, *Science*, *343*(6176), 1203–1205. (Appendix). doi:10.1126/science.1248506
Leonard, K. 2011. ***Talent acquisition factbook 2011: Benchmarks and trends of spending, staffing and key talent metrics*** (Report). Oakland, CA: Bersin & Associates.
Lewandowsky, S., & Farrell, S. 2011. ***Computational Modeling in Cognition: Principles and Practice***. Thousand Oaks, CA: Sage.
Lievens, F., Highhouse, S., & De Corte, W. 2005. The importance of traits and abilities in supervisors' hirability decisions as a function of method of assessment. ***Journal of Occupational and Organizational Psychology***, 78(3): 453–470.
Luan, S., & Reb, J. 2017. Fast-and-frugal trees as noncompensatory models of performance-based personnel decisions. ***Organizational Behavior and Human Decision Processes***, 141: 29–42.
Luan, S., Schooler, L. J., & Gigerenzer, G. 2014. From perception to preference and on to inference: An approach–avoidance analysis of thresholds. ***Psychological Review***, 121(3): 501–525.
Manimala, M. J. 1992. Entrepreneurial heuristics: A comparison between high PL (pioneering-innovative) and low PI ventures. ***Journal of Business Venturing***, 7(6): 477–504.
Marewski, J. N., & Schooler, L. J. 2011. Cognitive niches: An ecological model of strategy selection. ***Psychological Review***, 118(3): 393–437.
Martignon, L., & Hoffrage, U. 2002. Fast, frugal, and fit: Simple heuristics for paired

comparison. ***Theory and Decision***, 52(1): 29–71.
Martignon, L., & Hoffrage, U. 1999. Why does one-reason decision making work. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), ***Simple heuristics that make us smart***: 119–140. New York, NY: Oxford University Press.
McCrae, R. R., Martin, T. A., & Costa, P. T. 2005. Age trends and age norms for the NEO Personality Inventory-3 in adolescents and adults. ***Assessment***, 12(4): 363–373.
Milkman, K.L., Chugh, D., & Bazerman, M.H., 2009. How can decision making be improved?. ***Perspectives on Psychological Science***, 4(4): 379–383.
Münsterberg, H. 1912. ***Psychologie und Wirtschaftsleben: ein Beitrag zur angewandten Experimental-Psychologie***. Barth.
Newell, A., & Simon, H.A., 1972. ***Human problem solving***. Englewood Cliffs, NJ: Prentice-Hall.
Pachur, T., & Marinello, G. 2013. Expert intuitions: How to model the decision strategies of airport customs officers?. ***Acta Psychologica***, 144(1): 97–103.
Payne, J.W., Bettman, J.R., & Johnson, E.J., 1993. ***The adaptive decision maker***. Cambridge University Press.
Phillips, N. D., Neth, H., Woike, J. K., & Gaissmaier, W. 2017. FFTrees: A toolbox to create, visualize, and evaluate fast-and-frugal decision trees. ***Judgment and Decision Making***, 12(4): 344–368.
Rieskamp, J., & Otto, P.E. 2006. SSL: a theory of how people learn to select strategies. ***Journal of Experimental Psychology: General***, 135(2): 207–236.
Roth, P. L., Bobko, P., Switzer, F. S., & Dean M.A. 2001. Prior selection causes biased estimates of standardized ethnic group differences: simulation and analysis. ***Personnel Psychology,*** 54: 591–617.
Roth, P. L., Switzer III, F. S., Van Iddekinge, C. H., & Oh, I. S. 2011. Toward better meta-analytic matrices: How input values can affect research conclusions in human resource management simulations. ***Personnel Psychology***, 64(4): 899–935.
Ryan, A. M., & Ployhart, R. E. 2014. A century of selection. ***Annual Review of Psychology***, 65(1): 693–717.
Sackett, P. R., & Lievens, F. 2008. Personnel selection. ***Annual Review of Psychology***, 59(1): 419–450.
Savage, L. J. 1954. ***The foundations of statistics***. John Wiley and Sons
Schmidt, F. L., & Hunter, J. 2004. General mental ability in the world of work: Occupational attainment and job performance. ***Journal of Personality and Social Psychology***, 86(1): 162–173.
Schmidt, F. L., & Hunter, J. E. 1998. The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. ***Psychological Bulletin***, 124(2): 262–274.
Schwartz, J., Collins, L., Stockton, H., Wagner, D., & Walsh, B. 2017. ***Rewriting the rules for the digital age***. Deloitte Global Human Capital Trends.
Simon, H. A. 1947. ***Administrative behavior: A study of decision-making processes in administrative organizations***. New York, NY: Macmillan.
Simon, H. A. 1955. A behavioral model of rational choice. ***The Quarterly Journal of Economics***, 69(1): 99.
Simon, H. A. 1957. ***Administrative behavior: A study of decision-making processes in administrative organizations*** (2nd ed.). New York, NY: Macmillan.
Simon, H. A. 1971. Designing organizations for an information-rich world. In M. Greenberger (Ed.), ***Computers, communications and the public interest***: 37–72. Baltimore, MD: The

Johns Hopkins Press.
Simon, H. A. 1990. Invariants of human behavior. ***Annual Review of Psychology***, 41(1): 1–20.
Şimşek, Ö., & Buckmann, M. (2015). Learning from small samples: An analysis of simple decision heuristics. ***Advances in Neural Information Processing Systems***, 3159–3167.
Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. ***Journal of the Royal Statistical Society***, Series B, 36: 111–147.
Thaler, R. H., & Sunstein, C. R. 2008. ***Nudge***. New Haven: Yale University Press
Todd, P. M., Gigerenzer, G., & the ABC Research Group. (Eds.). 2012. ***Ecological rationality: Intelligence in the world***. New York, NY: Oxford University Press.
Tversky, A. 1969. Intransitivity of preferences. ***Psychological Review***, 76(1): 31–48.
Tversky, A. 1972. Elimination by aspects: A theory of choice. ***Psychological Review***, 79(4): 281–299.
Tversky, A., & Kahneman, D. 1974. Judgment under Uncertainty: Heuristics and Biases. ***Science***, 185: 1124–1131.
Vroom, V. H., & Jago, A. G. 2007. The role of the situation in leadership. ***American Psychologist***, 62: 17–24.
Wegwarth, O., Gaissmaier, W., & Gigerenzer, G. 2009. Smart strategies for doctors and doctors-in-training: Heuristics in medicine. ***Medical Education***, 43: 721–728.

Table 1. Statistical properties of the criterion variable (FJP) and the three cues from 236 job applicants at an airline company, Study 1.

	Range	Mean	SD	Correlation matrix			
				FJP	GMA	CON	USIP
Future job performance (FJP)	[1.75, 4.50]	3.16	0.44	1	---	---	---
General mental ability (GMA)	[0.42, 0.96]	0.70	0.12	0.30	1	---	---
Conscientiousness (CON)	[2.27, 5.00]	3.95	0.45	0.22	0.10	1	---
Unstructured interview performance (USIP)	[2, 5]	3.20	1.02	0.06	0.11	0.02	1

Note: The scale range for each variable is as follows: FJP: 1–5, GMA: 0–1, CON: 1–5, and USIP: 1–5.

Table 2. Parameter values (i.e., cue validities and intercue correlations) used to construct the simulated task environments of Study 2.

	Meta-analytic correlation matrix				Parameter values					
	FJP	GMA	CON	SIP	a	b	c	d	e	f
Future job performance (FJP)	1	---	---	---	0.05	0.05	0.05	−0.10	0.14	0.02
General mental ability (GMA)	0.30 (a)	1	---	---	0.20	0.08	0.20	0	0.24	0.12
Conscientiousness (CON)	0.18 (b)	0 (d)	1	---	0.30	0.18	0.30	0.10	0.34	0.22
Structured interview performance (SIP)	0.30 (c)	0.24 (e)	0.12 (f)	1	0.40	0.28	0.40			

Table 3. Results pertaining to the ecological rationality of logistic regression and Δ-inference in Study 2.

Sample size	Mean relative frequency of logistic regression predicting better than Δ-inference			Bivariate correlation with relative frequency of logistic regression predicting better than Δ-inference		Mean frugality of Δ-inference (cues searched)
	Overall	J-shaped	Not J-shaped	Best cue's relative predictiveness	Linear predictability	
$n = 30$	0.44	0.44	0.45	−0.38	0.31	1.26
$n = 100$	0.49	0.46	0.50	−0.61	0.78	1.36
$n = 1,000$	0.65	0.59	0.71	−0.69	0.73	1.45

Table 4. Some key measures of Study 3 by experimental condition.

Measure		Receptionist condition	Analyst condition
Reaction time (RT)	% of abnormal RTs	2.03	1.89
	Mean (in seconds)	5.28	5.21
	SD (in seconds)	2.79	2.65
Average rank of a cue by participants' subjective importance ratings	GMA	2.43	1.60
	CON	1.60	1.98
	SIP	1.97	2.42
Proportion of participants classified as using a certain strategy	Logistic regression	0.51	0.62
	Δ-inference	0.49	0.38

Notes. GMA = general mental ability; CON = conscientiousness; SIP = structured interview performance.

Figure 1. An example of a paired-comparison decision in which two job candidates' scores on three cues, general mental ability (GMA), conscientiousness (CON), and structured interview performance (SIP), are provided.

	Candidate A	Candidate B
General mental ability	116	102
Conscientiousness	47	55
Structured interview performance	3.6	3.9

Figure 2. The prediction accuracy of Δ-inference and logistic regression in a paired-comparison selection task based on data of 236 actual job applicants at an airline company, Study 1. *n* = sample size.

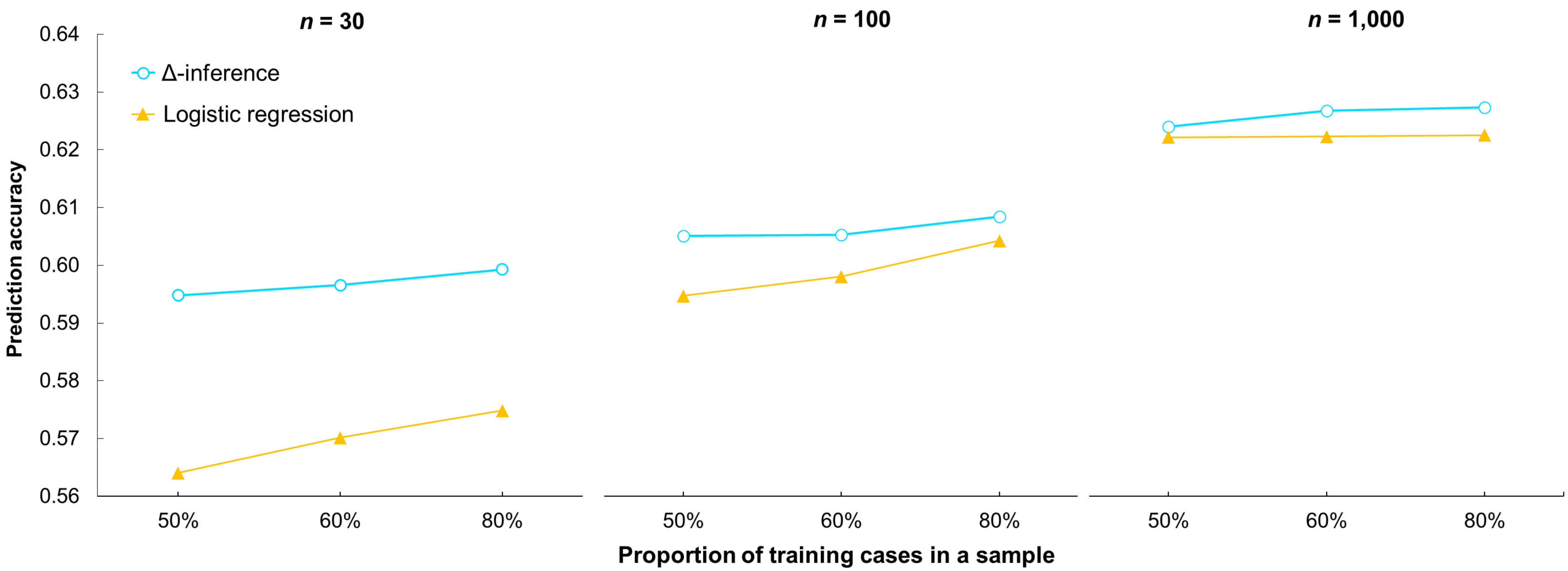


Figure 3. Scatter plots of the relative frequency of logistic regression predicting better than Δ-inference against the best cue's relative predictiveness (left) and the linear predictability of an environment (right), when sample size was 100 in Study 2.

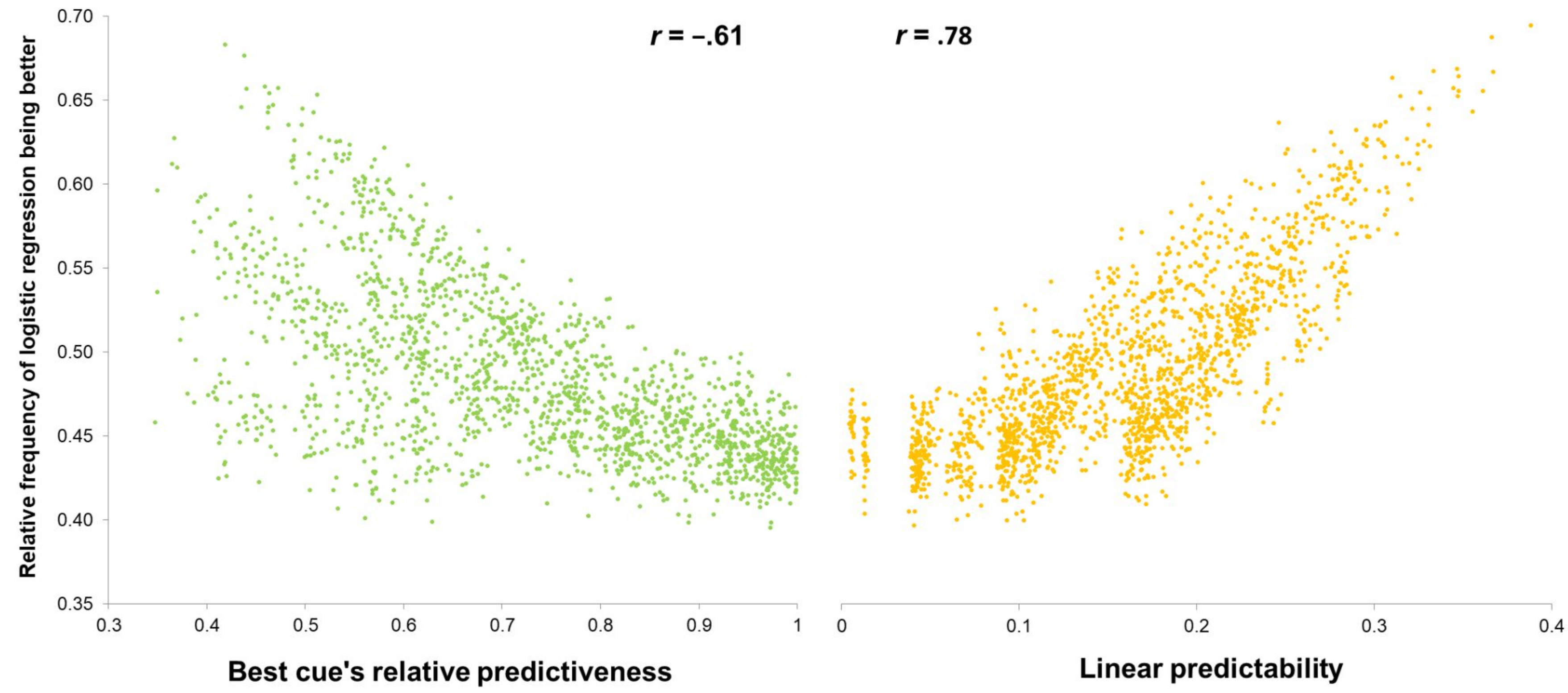


Figure 4. Proportion of participants in Study 3 classified as using Δ-inference, depending on participants' previous experience in selection decision, whether the distribution of their cue importance ratings was skewed, and the job condition in which decisions were made. The dotted line in each panel indicates the overall proportion of participants classified as using Δ-inference in each job condition. Error bars indicate standard errors.

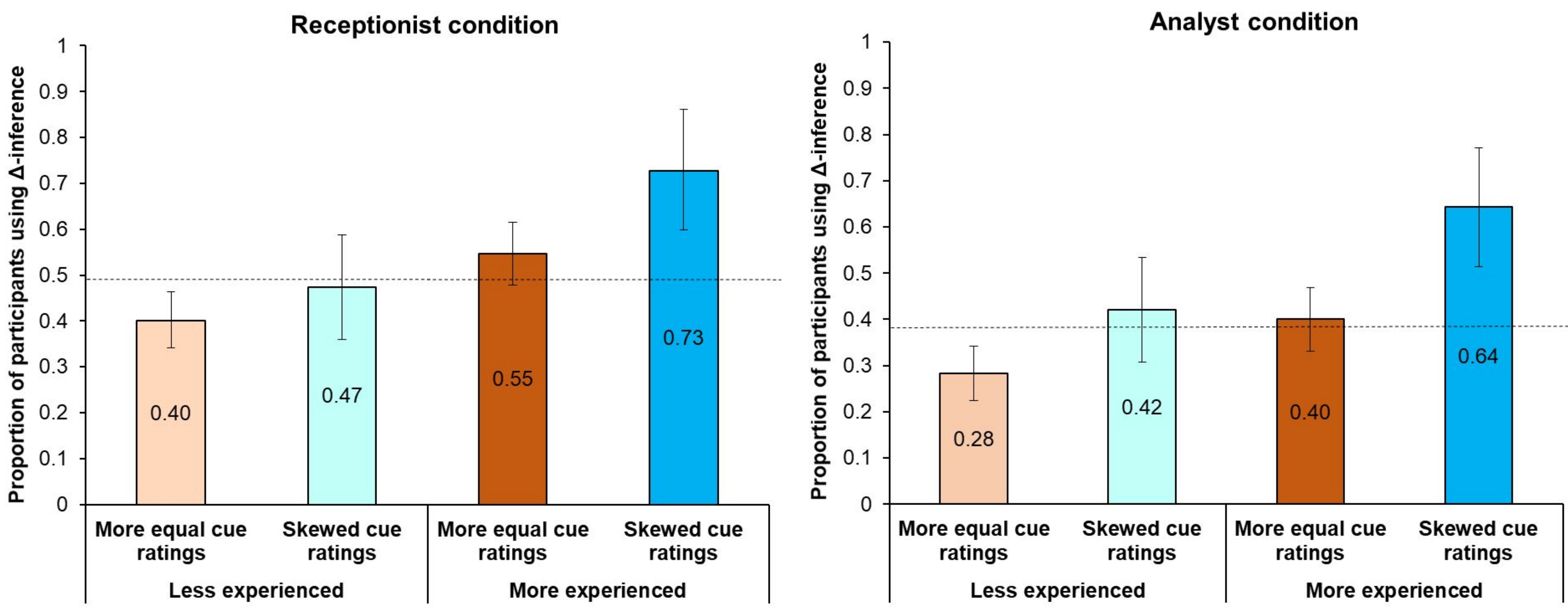
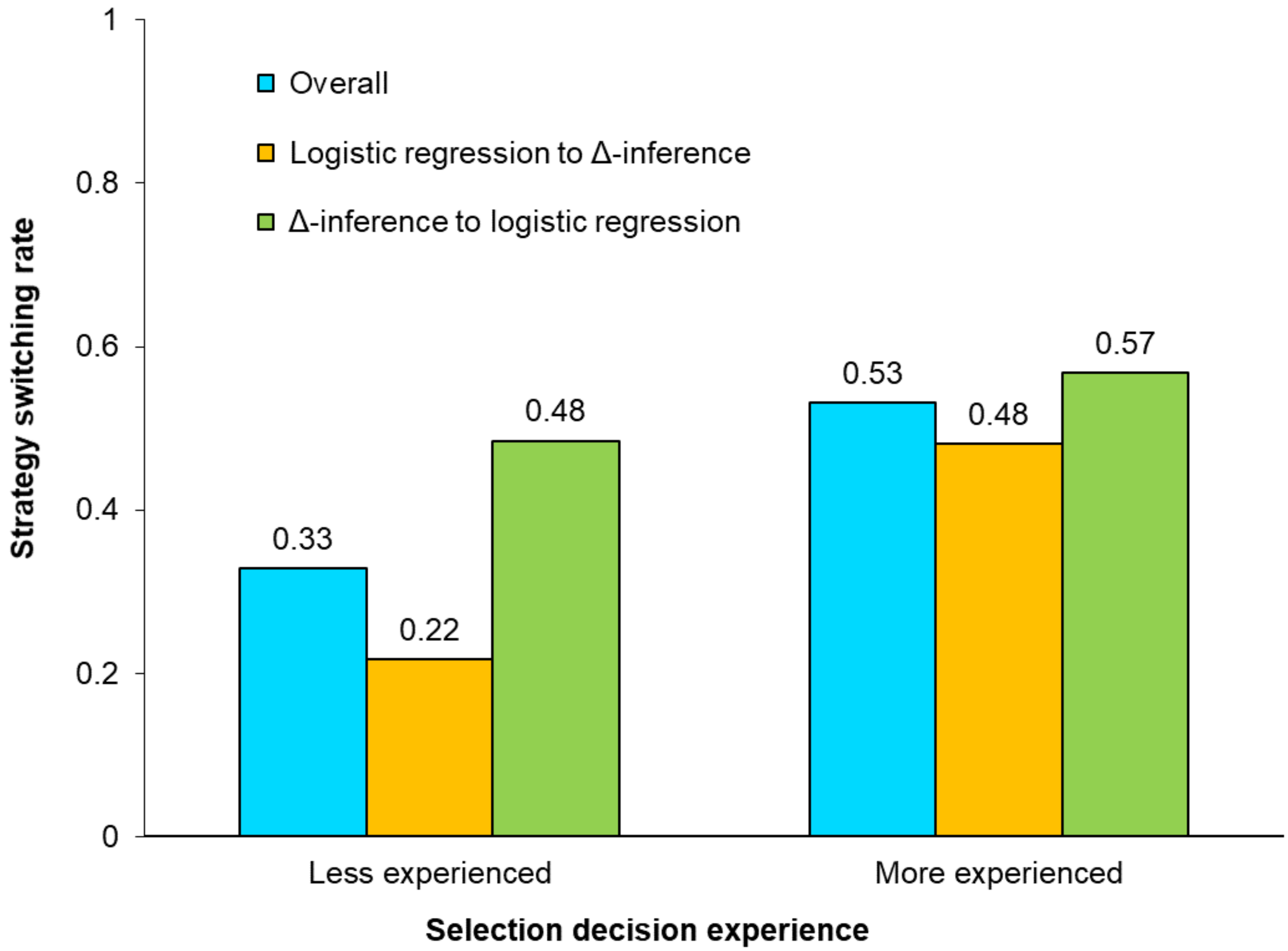


Figure 5. The rate of strategy switching from the receptionist condition to the analyst condition for the less and the more experienced participants, Study 3.



APPENDIX A: INSTRUCTIONS BEFORE EACH EXPERIMENTAL CONDITION IN STUDY 3

The Receptionist Condition

In this part of this study, we will ask you to make decisions on 105 pairs of candidates for a receptionist position. Please carefully read the position description below before moving on to the next part of the study in which you will make your decisions.

Position: Receptionist
Responsibilities:

- Manage front desk duties including:
  - Greet and attend to walk-in guests on their appointment and queries
  - Exchanging of visitor passes
  - Usher guests to meeting rooms
  - Handle general queries related to meeting rooms and booking of facilities
- Attend to incoming calls to main office hotline and handle caller's enquiries
- Re-direct calls as appropriate and take adequate messages when required
- Assist in the preparation of refreshments for meetings and meeting rooms (includes the setup of video conferencing equipment, laptop, etc.)
- Collate and prepare monthly statistical report for submission to management
- Provide general office administration support
- Any other duties as assigned

The Analyst Condition

In this part of this study, we will ask you to make decisions on 105 pairs of candidates for a lead data analyst position. Please carefully read the position description below before moving on to the next part of the study in which you will make your decisions.

Position: Lead Data Analyst
Responsibilities:

- Work with business departments and other technical team to gather data assets to support a single source of truth for all data across departments
- Design and build logical data model to meet business capabilities and technical requirements from different source systems and databases
- Analyse potential areas where existing data model, data policy and procedures require change, or where new ones need to be developed, especially regarding future business capabilities
- Gather data requirements, design and implement data integration, data quality, data cleansing and other ETL-related projects
- Perform ETL programming activities with scripts, packages and mappings using SAS data management solution
- Build dashboards and reports using Qlik solution to provide business and operational to the departments
- Use statistical methods to analyse customer data trends and generate useful business reports
- Provide primary operational support for information architecture, data factory and data analysis

SUPPLEMENTARY MATERIALS FOR "ECOLOGICAL RATIONALITY: FAST-AND-FRUGAL HEURISTICS FOR MANAGERIAL DECISION MAKING UNDER UNCERTAINTY"

Additional Analyses and Results for Study 1

In this section, we report and discuss the findings of additional analyses we conducted in Study 1. Table S1 shows first the frugality (i.e., average number of cues searched) of Δ-inference. As can be seen, Δ-inference searched less information to make a decision with fewer learning opportunities and on average did not search more than 1.5 cues. Table S1 also shows the magnitudes of overfitting by Δ-inference and logistic regression. Overfitting is defined as the difference between a model's accuracy in the training/learning cases and in the prediction cases. As the table shows, the two models' magnitudes of overfit depended on sample size: logistic regression overfit more when sample size was very small, both models overfit only slightly when sample size was very large, and Δ-inference overfit more, though by only small margins, when sample size was moderate.

Table S2 shows the prediction accuracy and overfitting of three models: Δ-inference, logistic regression, and the OLS (ordinary least squares) regression. To make the models comparable, the criterion variable in all three models was the binary decision between two job candidates. The new model, OLS regression, performed very similarly to logistic regression in both prediction accuracy and overfitting. This is not surprising, because both are linear models predicting a binary criterion variable.

One advantage of OLS regression over logistic regression, however, is that OLS regression can utilize the continuous form of the criterion variable—in our case, the specific difference between two job candidates' FJP ratings—in the training phase to fine-tune its parameters. Δ-inference can also benefit from this additional information to estimate more precisely cues' validities and then their search orders in the training phase. Table S3 shows these two models' prediction accuracy when using the continuous criterion variable and the improvements over using the binary criterion variable. For each model, there were indeed improvements, and Δ-inference stood to benefit even more than OLS regression in all conditions and especially when sample sizes were relatively small. These results show further how good Δ-inference can potentially be even as a simple and frugal heuristic.

Another analysis we did is about the prediction accuracy of logistic regression and Δ-inference when each used only the GMA and CON cues to predict FJP, given that the USIP cue's validity was very low in this data set (see Table 1). The results, shown in Table S4, indicate that (a) both models became more accurate by not considering the USIP cue; (b) both benefitted more when sample size was smaller; and (c) Δ-inference generally benefitted more than logistic regression, especially when sample size was very small, and this occurred with Δ-inference needing to search for even less information (see the last column in Table S4). In general, these results provide another example of the "less-is-more" phenomenon and show that considering more information does not necessarily lead to better predictions. That said, we chose not to report these findings in the main text, because (a) the improvements from considering three cues to two cues were generally not large in prediction accuracy and, in the case of Δ-inference, frugality; (b) this could be a special case because of the particularly low validity of USIP in the data set; and (c) the results are less compatible with those reported in Studies 2 and 3. In Study 3, we did compare three-cue with two-cue models again on the behavioral data and found that three-cue models generally perform better there (see below).

The final additional analysis we conducted is to run three popular machine learning algorithms in the Study 1 data set and compare their prediction accuracies to those of Δ-inference and logistic regression. The three algorithms are: LASSO regression, random forest, and support vector machine (SVM). Their prediction accuracies are shown in Table S5. The results show that when sample size was large (n = 1,000), LASSO and random forest performed similarly to logistic regression, trailing Δ-inference only very slightly. When sample size was smaller, however, these two were even worse than logistic regression, let alone Δ-inference. SVM performed better than the other two machine learning algorithms when sample size was very small, but its performance failed to reach the same levels as the other two when sample size was large. In sum, the three machine learning algorithms, which required much fine-tuning during programming and substantially more computational power to get the results, were generally outperformed by Δ-inference and logistic regression, two much simpler strategies, in the Study 1 data set. Given the great uncertainty/unpredictability in the task, we had anticipated that complex algorithms like those three would not predict much better than Δ-inference in the data set. The fact that they failed to outperform Δ-inference in a single condition—even logistic regression in most of the conditions—still came as a surprise to us. "Less-is-more" is once again being demonstrated here; and amidst the frenzy of machine learning algorithms in this time, our results serve as a cautionary case against the blind belief in those algorithms' ability to predict the uncertain future.

Model Performance in Simulated Lexicographic Environments

In Study 2, we simulated 1,728 linear environments and compared the prediction performance of logistic regression and Δ-inference in those environments. How would the two perform in lexicographic environments? To explore, we simulated eight environments, in each of which paired-comparison decisions were generated following the Δ-inference heuristic. Specifically, we first generated three cues with a multivariate Normal distribution, in which the variance of each marginal distribution (i.e., distribution of each cue) was 1 and the intercue correlation ρ between each pair of cues was uniform. Without loss of generalizability, we assumed that cues were searched in the order of GMA, CON, and SIP. Each time a cue was searched, the values of the two candidates on the cue were made available and their difference was calculated. The difference was then compared to a preset threshold value Δ; if the difference exceeded Δ, a decision was made; otherwise, the search continued to the next cue.

The eight lexicographic environments were results of eight combinations of three factors: (1) the uniform pairwise cue correlation ρ, with ρ = .10 or .40; (2) the intended threshold Δ, with Δ = .20 or .50, both being z-scores; and (3) the magnitude of random noise on Δ. Random noise was generated from a uniform distribution defined by lower and upper bounds; and for each cue, a noise value was independently generated and added to the intended Δ. The lower bound of the noise distribution was fixed at −0.20, so that the final Δ (e.g., the intended Δ plus the noise) would not become negative in the low-threshold environments; however, we varied the upper bound with two levels: 0.20 and 1.0.

As in the main text, we tested the prediction accuracy of logistic regression and Δ-inference using cross-validation with three sample sizes: 30, 100, and 1,000, and applying a fixed "60-40" split. In each environment, the two strategies' performances were based on 10,000 random samples in the n = 30 and n = 100 conditions and 1,000 in the n = 1,000 condition. Different from the main text, though, the main measure we used here was the relative frequency that *Δ-inference* achieved higher prediction accuracy than logistic regression across all samples in a sample-size condition. Because the results in the n =100 condition fell in between those in the n =30 and n = 1,000 conditions, we only show results from the latter two in Figure S1.

First of all, Δ-inference performed better than logistic regression (i.e., the relative frequency above .50) in all eight environments. However, its advantages dwindled with smaller sample sizes, similar to what occurred to logistic regression in the linear environments (see Table 3). Smaller sample sizes made the environments noisier and learning more difficult, increasing the chance that a non-generating model would outperform the generating one. Second, Δ-inference generally performed better than logistic regression in environments with a lower intercue correlation and/or a lower level of random noise, regardless of the threshold level.

Third and quite interestingly, the effect of threshold level apparently depended on sample size: When sample size was large, Δ-inference outperformed logistic regression more in the high-threshold environments than in the low-threshold environments; however, when sample size was small, the opposite pattern occurred. Large sample size makes the discovery of the true underlying model more likely, no matter whether the environments are linear or lexicographic. In the latter, when Δ is high, it is more likely that a decision is not made by the first cue but deferred to the second and the third cues step-by-step, which makes it more difficult for a compensatory strategy such as logistic regression to mimic the process. When Δ is low, similar to a J-shaped linear environment, a large proportion of decisions are made by the first cue, which makes it easier for logistic regression to mimic the process by assigning a much higher beta weight to the first cue. Therefore, Δ-inference should in theory outperform logistic regression more when sample size is large. When sample size is small, as we discussed before, the large amount of noise in the environment makes learning very difficult and distorts model-comparison patterns. Indeed, out of the eight environments, Δ-inference performed the worst (relatively) in high-correlation-high-threshold environments where random noise on threshold was high and sample size was small.

In previous simulation studies of ecological rationality, the simulated environments were overwhelmingly linear, likely due to researchers' familiarity with such environments and the linear combination rules. How models perform in environments where decisions are generated by lexicographic rules has not been investigated. Our analysis reported here provides some first insights on what may happen, in terms of model comparison, in lexicographic environments and what environmental properties may affect models' relative performance.

Additional Analyses and Results for Study 3

In Study 3, we had a special treatment for response times (RTs) that were abnormally long and short for each participant. What would happen if those RTs were left as they were? Would this affect the modeling results? Table S6 shows the proportions of participants who were classified as using either logistic regression or Δ-inference with and without the RT treatment. Without the RT treatment, the proportion of identified Δ-inference users increased slightly in both experimental conditions, but the general pattern remained.

Figures 4 and 5 in the main text show the main model-testing results in Study 3. Figures S2 and S3 report results without the RT treatment that correspond to those shown in Figures 4 and 5, respectively. With regard to the type of participants who used Δ-inference, Figure S2 shows the same pattern as in Figure 4; that is, participants with more experience in selection decisions and with a skewed cue importance distribution adopted Δ-inference more frequently, and this pattern occurred in both experimental conditions. With regard to strategy switching, Figure S3 shows the same pattern as in Figure 5; that is, the more experienced participants were more likely to switch strategies between the two experimental conditions than the less experienced. Therefore, all the main results in Study 3 pertaining to model testing remained without the RT treatment. Table 4 in the main text shows that RTs in only a very small proportion of trials were deemed as abnormally long or short. Leaving those RTs untreated did affect some aspects of the results but did not change the main result patterns and conclusion in Study 3.

The second additional analysis we conducted for Study 3 is about cue importance. Besides participants' subjective importance ratings on the cues, we also measured cue importance with two other measures: (a) for participants classified as using logistic regression, we derived a cue's importance by the frequency of its beta weight being statistically significant ($p < .01$), and (b) for those using Δ-inference, we calculated the cue's average search order. Table S7 shows the results from all three measures separated by experimental condition and the type of strategy participants classified to use. There were remarkable consistencies among the three measures. Specifically, the average ranks of cue importance were identical by all three measures in the analyst condition; in the receptionist condition, there was only one inconsistency about the ranks of the second and the least important cues between the cue search orders identified with Δ-inference and the subjective importance ratings.

The next additional analysis we did for Study 3 was to test how well two-cue models could describe participants' behavioral data. Different from the two-cue analyses done for Study 1 where the two cues were always GMA and CON, the two cues were selected for each participant here based on the following procedure: We first calculated the bivariate correlation between each cue and a participant's decisions in the training cases, ordered them by their absolute magnitudes, and then excluded the last-ranked cue from consideration. After this screening step, we tested four models, the two-cue and three-cue versions of LR and Δ-inference, using the same method as how we tested the two three-cue models alone. Therefore, the competition was not only about LR versus Δ-inference, but also the three-cue version of each against its two-cue counterpart. The results are shown in Table S8.

With regard to three-cue versus two-cue models, the three-cue logistic regression was better than the two-cue logistic regression in each experimental condition; however, for Δ-inference, the two types were almost equally good. Sorting models by the number of cues they considered, we see from the bottom two rows of Table S8 that the three-cue models had overall advantages over the two-cue models in both experimental conditions. This is the main reason why we decided not to report and discuss the two-cue models in the main text.

Sorting models by their underlying processes, the results were almost identical to when we only compared the three-cue versions of logistic regression and Δ-inference (see Table 4): In the receptionist condition, logistic regression was still the strategic choice of a slight majority of participants (52%; and it was 51% when considering three-cue models only), and in the analyst condition, logistic regression was still the clear majority choice (61%; and it was 62% when considering three-cue models only). This near non-change of the model-comparison results was another reason why we chose not to report the two-cue modeling result in the main text.

Our last additional analysis for Study 3 is about the exclusion of participants. Recall that we excluded three types of participants (23 out of 166) in the study to ensure data quality. How robust are our results if we instead included everyone for analyses? We suspected that not much would change given the relatively small proportion of the excluded participants. Figures 4 and 5 in the main text show the main model-testing results in Study 3. Figures S4 and S5 report results with all 166 participants that correspond to those shown in Figures 4 and 5, respectively. With regard to the type of participants who used Δ-inference, Figure S4 shows the same pattern as in Figure 4; that is, participants with more experience and/or with a skewed cue importance distribution adopted Δ-inference more frequently. With regard to strategy switching, Figure S5 shows the same pattern as in Figure 5; that is, the more experienced participants were more likely to switch strategies than the less experienced. Therefore, the main results in Study 3 pertaining to model testing remained without excluding the 23 participants.

Table S1. The frugality of Δ-inference and the magnitudes of overfitting by Δ-inference and logistic regression in Study 1.

	% of training cases	Frugality of Δ-inference	Overfit (learning minus prediction accuracy)		
			Logistic regression	Δ-inference	Overfit more
n = 30 (10,000 samples)	50	1.222	0.191	0.098	LR
	60	1.237	0.164	0.097	
	80	1.259	0.133	0.093	
n = 100 (10,000 samples)	50	1.319	0.071	0.072	Δ-inference
	60	1.334	0.059	0.068	
	80	1.359	0.047	0.060	
n = 1,000 (1,000 samples)	50	1.476	0.009	0.009	Roughly equally small
	60	1.483	0.008	0.007	
	80	1.499	0.006	0.006	

Table S2. Comparisons in prediction accuracy and overfitting among three models in Study 1.

	% of training cases	Prediction accuracy			Overfit (learning minus prediction accuracy)		
		OLS regression	Logistic regression	Δ-inference	OLS regression	Logistic regression	Δ-inference
$n = 30$ (10,000 samples)	50	0.560	0.564	0.595	0.188	0.191	0.098
	60	0.569	0.570	0.597	0.161	0.164	0.097
	80	0.576	0.575	0.599	0.130	0.133	0.093
$n = 100$ (10,000 samples)	50	0.596	0.598	0.609	0.072	0.071	0.072
	60	0.600	0.601	0.609	0.061	0.059	0.068
	80	0.607	0.608	0.613	0.046	0.047	0.060
$n = 1,000$ (1,000 samples)	50	0.622	0.622	0.624	0.008	0.009	0.009
	60	0.623	0.622	0.627	0.007	0.008	0.007
	80	0.623	0.623	0.627	0.007	0.006	0.006

Table S3. The prediction accuracy and the improvements in prediction accuracy of OLS regression and Δ-inference with the continuous criterion variable in Study 1.

	% of training cases	Continuous criterion		Improvement over binary criterion	
		OLS regression	Δ-inference	OLS regression	Δ-inference
n = 30 (10,000 samples)	50	0.573	0.599	0.013	0.035
	60	0.576	0.600	0.007	0.030
	80	0.590	0.601	0.014	0.026
n = 100 (10,000 samples)	50	0.606	0.612	0.010	0.014
	60	0.608	0.613	0.008	0.012
	80	0.613	0.616	0.006	0.008
n = 1,000 (1,000 samples)	50	0.625	0.626	0.003	0.004
	60	0.625	0.629	0.002	0.007
	80	0.625	0.629	0.002	0.006

Table S4. The prediction accuracy of logistic regression and Δ-inference and frugality of Δ-inference with two and three cues in Study 1.

	% of training cases	Logistic regression			Δ-inference			Frugality of Δ-inference	
		2-cue	3-cue	2-cue benefit	2-cue	3-cue	2-cue benefit	2-cue	3-cue
n = 30 (10,000 samples)	50	0.572	0.564	0.008	0.610	0.595	0.016	1.176	1.222
	60	0.576	0.570	0.006	0.611	0.597	0.014	1.190	1.237
	80	0.583	0.575	0.008	0.612	0.599	0.013	1.216	1.259
n = 100 (10,000 samples)	50	0.603	0.598	0.005	0.616	0.609	0.007	1.275	1.319
	60	0.607	0.601	0.006	0.617	0.609	0.008	1.291	1.334
	80	0.611	0.608	0.003	0.618	0.613	0.005	1.307	1.359
n = 1,000 (1,000 samples)	50	0.624	0.622	0.002	0.625	0.624	0.001	1.413	1.476
	60	0.624	0.622	0.001	0.627	0.627	0.000	1.423	1.483
	80	0.624	0.623	0.002	0.627	0.627	0.000	1.435	1.499

Table S5. Comparisons in prediction accuracy among five models in Study 1.

	% of training cases	Δ-inference	Logistic regression	LASSO regression	Random forest	SVM
n = 30 (10,000 samples)	50	0.595	0.564	0.535	0.532	0.549
	60	0.597	0.570	0.542	0.537	0.555
	80	0.599	0.575	0.549	0.543	0.557
n = 100 (10,000 samples)	50	0.609	0.598	0.575	0.570	0.574
	60	0.609	0.601	0.582	0.576	0.575
	80	0.613	0.608	0.592	0.585	0.575
n = 1,000 (1,000 samples)	50	0.624	0.622	0.621	0.618	0.601
	60	0.627	0.622	0.623	0.619	0.602
	80	0.627	0.623	0.624	0.622	0.609

Table S6. The proportions of participants in Study 3 classified as using LR and Δ-inference with and without the RT treatment.

	Model	Receptionist condition	Analyst condition
With RT treatment	LR	0.51	0.62
	Δ-inference	0.49	0.38
Without RT treatment	LR	0.47	0.59
	Δ-inference	0.53	0.41

Table S7. Cue importance by different measures in Study 3.

Measure	Cue	Receptionist condition		Analyst condition	
		Using LR	Using Δ-inference	Using LR	Using Δ-inference
Average rank of a cue by participants' subjective importance ratings	GMA	2.31	2.53	1.56	1.62
	CON	1.60	1.60	2.01	1.97
	SIP	2.09	1.86	2.43	2.42
Frequency of a cue being statistically significant ($p < .01$) in LR	GMA	0.08		0.67	
	CON	0.60		0.63	
	SIP	0.55		0.35	
Average search order of a cue in Δ-inference	GMA		2.16		1.69
	CON		1.50		2.00
	SIP		2.34		2.31

Table S8. The proportions of participants in Study 3 classified using both two-cue and three-cue versions of LR and Δ-inference.

		Receptionist condition	Analyst condition
LR	Two cues	0.20	0.20
	Three cues	0.33	0.41
	Total	0.52	0.61
Δ-inference	Two cues	0.24	0.18
	Three cues	0.23	0.21
	Total	0.48	0.39
Two-cue models total		0.44	0.38
Three-cue models total		0.56	0.62

Figure S1. The relative frequency that Δ-inference achieved higher prediction accuracy than LR in the eight lexicographic environments. *n* = sample size.

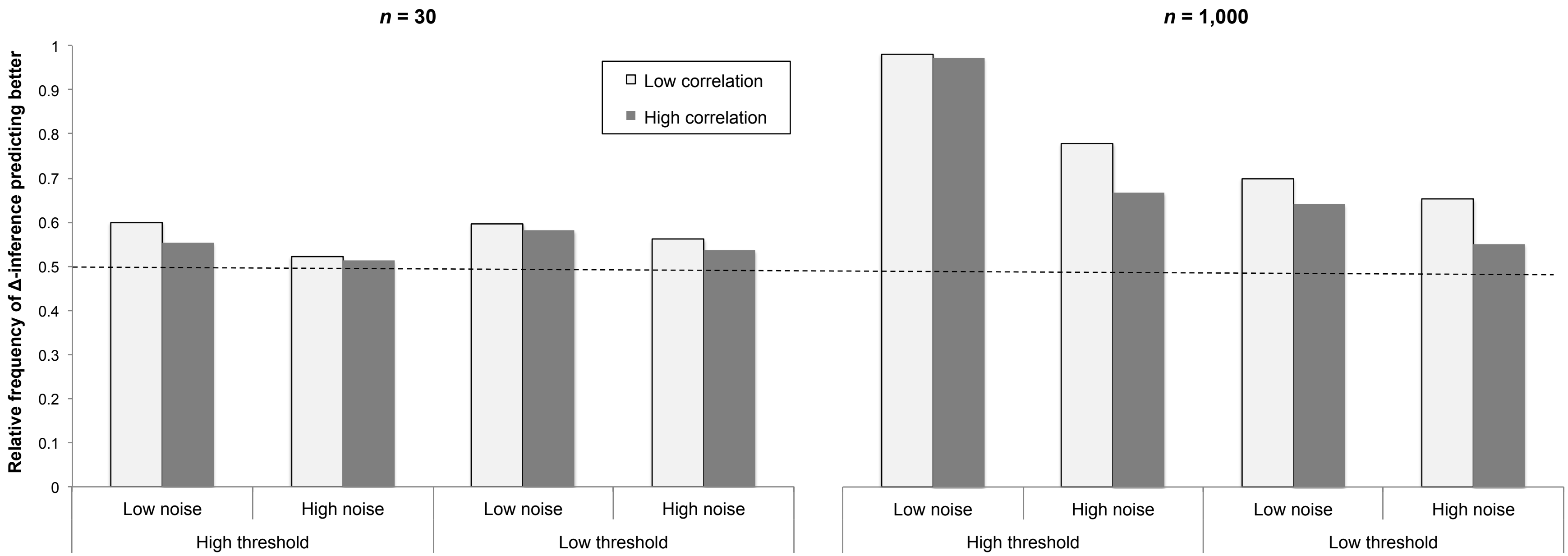
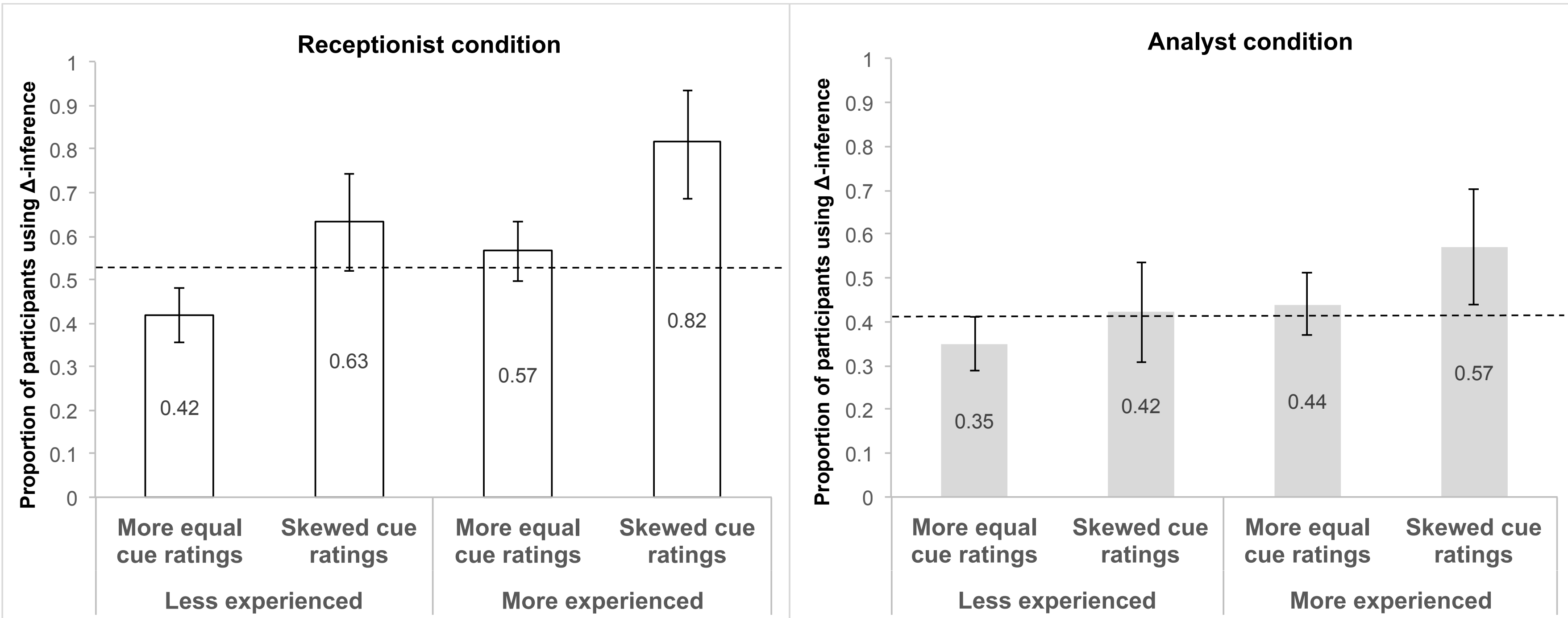


Figure S2. Results without the RT treatment that correspond to those reported in Figure 4 in the main text. The dotted line in each panel indicates the overall proportion of participants who used Δ-inference in each job condition. Errors bars indicate standard errors.



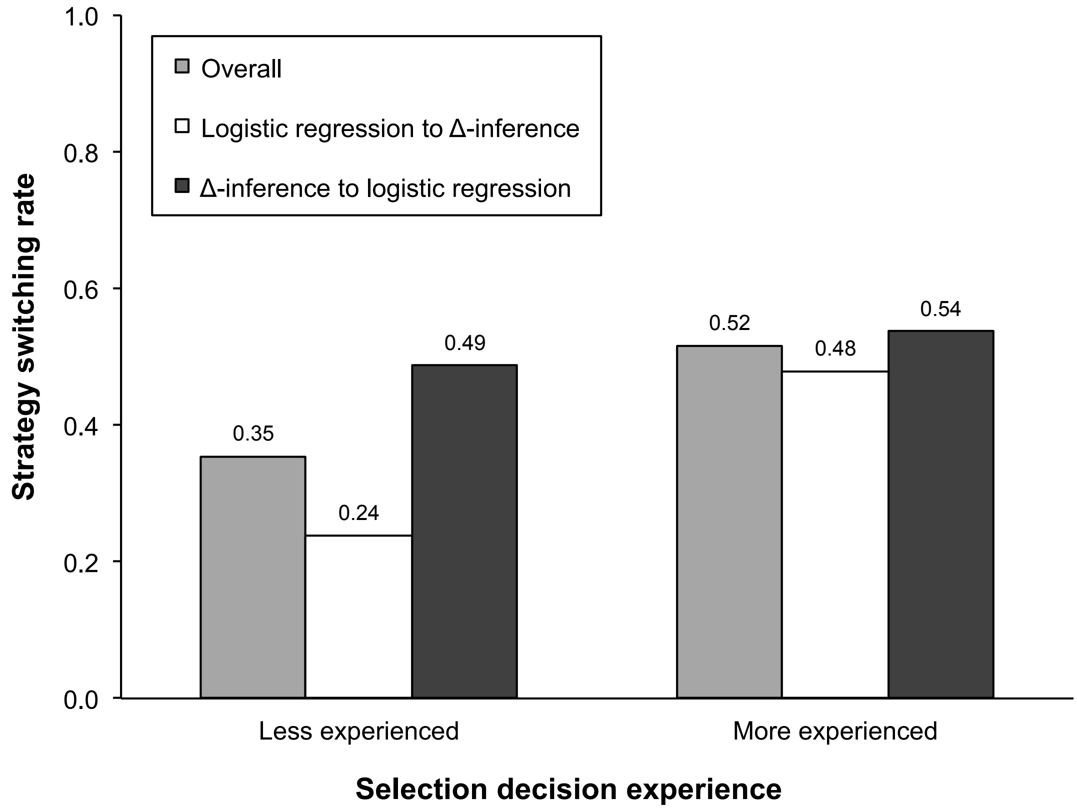
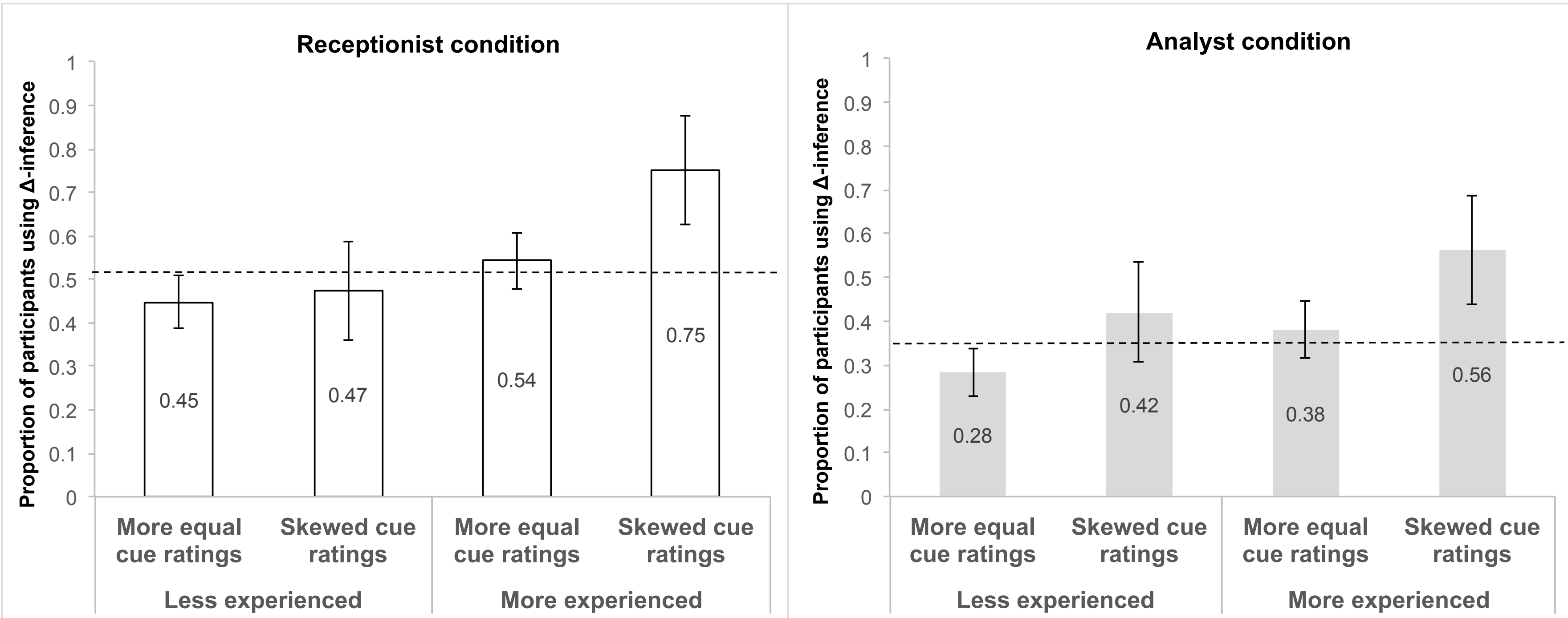


Figure S3. Results without the RT treatment that correspond to those reported in Figure 5 in the main text.

Figure S4. Results of all participants, including the 23 excluded ones, that correspond to those reported in Figure 4 in the main text. The dotted line in each panel indicates the overall proportion of participants who used Δ-inference in each job condition. Errors bars indicate standard errors.



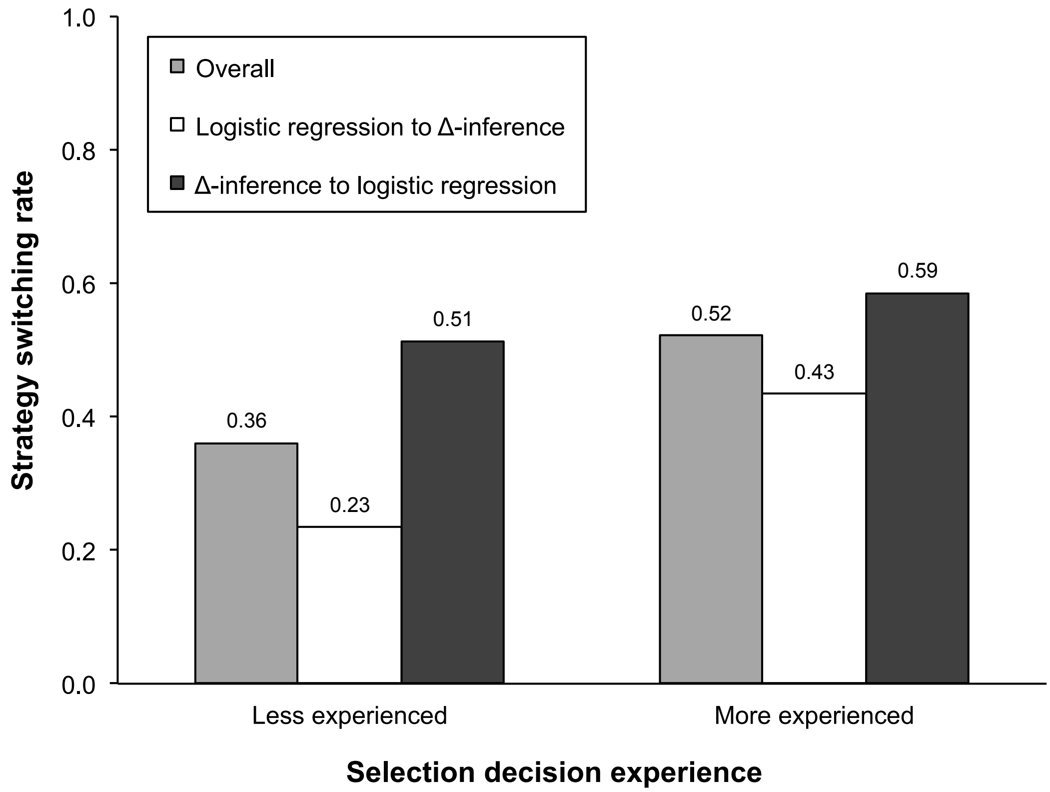


Figure S5. Results of all participants, including the 23 excluded ones, that correspond to those reported in Figure 5 in the main text.

APPENDIX A: INSTRUCTIONS BEFORE EACH EXPERIMENTAL CONDITION IN STUDY 3

The Receptionist Condition

In this part of this study, we will ask you to make decisions on 105 pairs of candidates for a receptionist position. Please carefully read the position description below before moving on to the next part of the study in which you will make your decisions.

Position: Receptionist
Responsibilities:

- Manage front desk duties including:
    - Greet and attend to walk-in guests on their appointment and queries
    - Exchanging of visitor passes
    - Usher guests to meeting rooms
    - Handle general queries related to meeting rooms and booking of facilities
- Attend to incoming calls to main office hotline and handle caller's enquiries
- Re-direct calls as appropriate and take adequate messages when required
- Assist in the preparation of refreshments for meetings and meeting rooms (includes the setup of video conferencing equipment, laptop, etc.)
- Collate and prepare monthly statistical report for submission to management
- Provide general office administration support
- Any other duties as assigned

The Analyst Condition

In this part of this study, we will ask you to make decisions on 105 pairs of candidates for a lead data analyst position. Please carefully read the position description below before moving on to the next part of the study in which you will make your decisions.

Position: Lead Data Analyst
Responsibilities:

- Work with business departments and other technical team to gather data assets to support a single source of truth for all data across departments
- Design and build logical data model to meet business capabilities and technical requirements from different source systems and databases
- Analyse potential areas where existing data model, data policy and procedures require change, or where new ones need to be developed, especially regarding future business capabilities
- Gather data requirements, design and implement data integration, data quality, data cleansing and other ETL-related projects
- Perform ETL programming activities with scripts, packages and mappings using SAS data management solution
- Build dashboards and reports using Qlik solution to provide business and operational to the departments
- Use statistical methods to analyse customer data trends and generate useful business reports
- Provide primary operational support for information architecture, data factory and data analysis

Shenghua Luan (luansh@psych.ac.cn) is a Professor at the Institute of Psychology, Chinese Academy of Sciences. He received his PhD from the University of Florida. His research interests include heuristics in judgment and decision making, wisdom of crowds, cooperation, sports and business forecasting, and managerial decision making.

Jochen Reb (jreb@smu.edu.sg) is an Associate Professor of Organisational Behaviour and Human Resources at the Lee Kong Chian School of Business at Singapore Management University. He received his PhD from the University of Arizona. His research focuses on judgment and decision making, particularly in workplace applications such as performance appraisals or selection decisions, as well as the role of mindfulness at work, particularly in areas such as leadership, employee performance, and teams.

Gerd Gigerenzer is Director of the Harding Center for Risk Literacy at the Max Planck Institute for Human Development in Berlin. He earned his PhD from the University of Munich, Germany. His research focusses on decision making under uncertainty, fast-and-frugal heuristics and their ecological rationality, as well as tools for improving statistical thinking and risk communication.