Organizational Behavior and Human Decision Processes 141 (2017) 29–42

Contents lists available at ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

Fast-and-frugal trees as noncompensatory models of performance-based personnel decisions

Shenghua Luan [a,*], Jochen Reb [b,1]

[a] Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany
[b] Singapore Management University, 50 Stamford Road, 178899 Singapore, Singapore

ARTICLE INFO

Article history:
Received 22 June 2016
Revised 10 May 2017
Accepted 13 May 2017

Number of Reviews = 1

Keywords:
Fast-and-frugal trees
Cue-based decision making
Dynamic performance
Personnel decisions
Process models
Forced distributions
Ecological rationality

ABSTRACT

Employees' performance provides the basis for many personnel decisions, and to make these decisions, managers often need to integrate information from different performance-related cues. We asked college students and experienced managers to make a series of performance-based personnel decisions and tested how well weighting-and-adding, compensatory logistic regression and lexicographic, noncompensatory fast-and-frugal trees (FFTs) could describe participants' decision processes regarding both choices and reaction times. Results show that a significant proportion of the participants (i.e., nearly half of the college students and more than two-thirds of the experienced managers) applied FFTs to make such decisions, and that the majority of them adopted key features of FFTs adaptively in response to a manipulation of the required distributions of positive (bonus) or negative (termination) decisions. Overall, the process-oriented approach applied in our study provides insights on not only what cues managers use for performance-based personnel decisions, but also how they use these cues.

1. Introduction

Given the crucial role of human capital for organizational success, personnel decisions such as whom to fire, whom to promote, and whom to reward are among the most influential managerial decisions (Guion, 2011). Because employees' job performance provides the basis, at least in part, for such decisions, researchers have been studying how performance-related cues influence the decision process (e.g., DeNisi, Cafferty, & Meglino, 1984; Landy & Farr, 1980). An important recognition is that for various reasons (e.g., market fluctuation and personal development), employee performance is often dynamic, displaying short-term and long-term changes over time, and that cues of dynamic performance can strongly influence performance appraisals and performance-based decisions (e.g., Barnes, Reb, & Ang, 2012; Reb & Cropanzano, 2007).

The three well-studied cues of dynamic performance are the performance mean (i.e., the average performance level over an evaluation period), trend (i.e., the trajectory of performance changes), and variation (i.e., the degree to which the performance fluctuates). Fig. 1 shows an employee's performance profile in which these cues can be readily discerned. Previous research suggests that performance appraisals are highly correlated with performance mean and trend, whereas findings have been mixed regarding the influence of performance variation (e.g., Reb & Cropanzano, 2007; Reb & Greguras, 2010). Extending this research to personnel decisions, Barnes et al. (2012) showed that mean and trend—but not variation—of NBA players' performance were positively related to managers' decisions to increase a player's salary in a new contract.

Building on these and other related studies (e.g., Lee & Dalal, 2011), we aim to address two important questions that have not been well understood in research of performance-based personnel decisions. First, how do managers use dynamic performance cues to arrive at such decisions? And second, to what extent do managers' decision processes correspond to the characteristics of the task environment? Drawing on the work of Simon on bounded rationality (1955) and recent research on decision heuristics (e.g., Todd, Gigerenzer, & the ABC Research Group, 2012), we posit that a significant proportion of managers would use fast-and-frugal trees (FFTs; defined below), a type of noncompensatory, lexicographic heuristics, to make performance-based personnel decisions and that they could apply FFTs adaptively in different task environments.

* Corresponding author.
E-mail addresses: shluan@mpib-berlin.mpg.de (S. Luan), jreb@smu.edu.sg (J. Reb).

[1] The authors contributed equally to this work. The order of authorship is alphabetic.

http://dx.doi.org/10.1016/j.obhdp.2017.05.003
0749-5978/

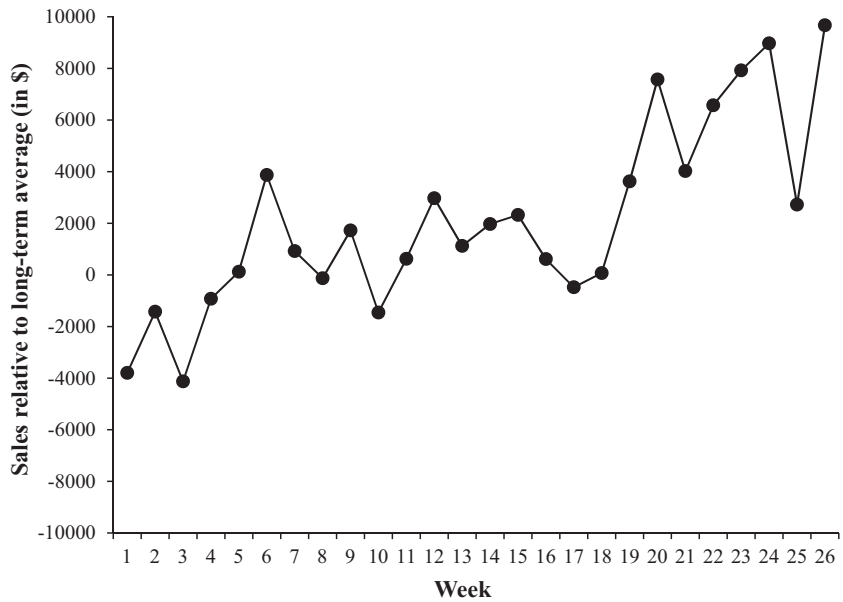

Fig. 1. A sample performance profile of an employee. Each point in the graph represents the employee's sales performance in a certain week, and the performance is quantified as the dollar amount the employee has made relative to the long-term average performance of all employees working for the company.

We investigated these questions in two studies with either college students (Study 1) or highly experienced managers (Study 2). In both studies, employees' performance in a certain period of time was displayed in charts similar to Fig. 1 and participants were asked to make a decision on bonus or termination for each employee. We varied three aspects of the performance—namely, the mean, trend, and variation—that could be used as cues for such decisions. In each study, we examined how well FFTs, in comparison to the compensatory logistic regression, could describe a participant's decision process, measured by the models' ability to predict both the choices and reaction times of the participant. Moreover, to study whether participants could adjust their decision processes adaptively, we manipulated the required distributions of bonus or termination decisions in different experimental conditions and tested how this would affect participants' decision processes.

In so doing, our research makes several noteworthy theoretical and methodological contributions. First, our prediction that managers use FFTs for performance-based decisions is novel in studies of dynamic performance. Analyzing data with either regression analysis or analysis of variance, previous studies have always assumed that managers integrate cues following a compensatory strategy by weighting and adding cue values. Whether this is what managers actually do has not been examined, nor have alternative, noncompensatory strategies been tested. Knowing the specific decision strategies managers apply will not only improve our understanding of *how* they integrate cues of dynamic performance to make decisions in addition to what cues they use, but also help us predict better what decisions managers would make and the importance of each cue in this process.

Second, the idea of "adaptive decision makers"—that people are capable of adapting their decisions strategies to the characteristics of the task environment—has been proposed and tested in many areas of decision making (e.g., Payne, Bettman, & Johnson, 1993; Simon, 1955; Todd et al., 2012), but received little attention in research of managerial decisions. Key to the success of an organization is the ability of its leaders and managers to apply strategies suitable for a task and be adaptive when the characteristics of the task have changed. Our study addresses this adaptiveness question in the context of personnel decisions, filling a critical gap in the literature.

Third, previous studies have found evidence for the use of FFTs in several domains of decision making (e.g., Dhami, 2003; Hertwig, Fischbacher, & Bruhin, 2013; Tan, Luan, & Katsikopoulos, 2017). Our study is the first to examine the possibility of FFTs for managerial decisions, a domain in which decision makers are argued to rely on heuristics to make many of their judgments and decisions (e.g., Artinger, Petersen, Gigerenzer, & Weibler, 2015; Hodgkinson & Healey, 2008). Furthermore, whereas past research has claimed that decision makers can adapt features of FFTs to different task environments (e.g., Luan, Schooler, & Gigerenzer, 2011), our study is the first to test this claim empirically.

Finally, we took a comparative approach in model testing by examining models with distinct assumptions and evaluated the descriptiveness of each model with respect to both choices and reaction times. These approaches are rarely applied in research on personnel decisions and managerial decision making more broadly, but can provide much insight on the underlying processes (e.g., Glöckner, 2009; Lewandowsky & Farrell, 2011). They are the main methodological contributions of our study.

1.1. Fast-and-frugal trees

FFTs are heuristics for binary decisions (i.e., decisions with two alternatives). As process models of decision making, FFTs make predictions not only about what cues will influence decisions but also how decision makers might use these cues. Formally, given m decision-related cues, an FFT is defined as "a decision tree that has $m + 1$ exits, with one exit for each of the first $m - 1$ cues and two exits for the last cue" (Luan et al., 2011, p. 320). An "exit" on an FFT points to the type of decision (e.g., award a bonus) made by a decision maker and is usually the outcome of meeting some specified condition set on a cue.

To illustrate how an FFT works, suppose that a manager is deciding whether to award a bonus to an employee upon seeing the performance profile shown in Fig. 1. The three cues that she could use to make the decision are the mean, trend, and variation of the employee's performance; and let us assume that the

manager deems the importance of the cues in that particular order. Using the FFT shown on the left side of Fig. 2, the manager first sets a criterion value on each cue that reflects what she considers "good enough," similar to the role of aspiration levels in the satisficing heuristic (Simon, 1955). If the performance mean of the employee is below the criterion (i.e., $<CR_M$), the manager will stop looking for other cues and decide not to award the bonus at this point; otherwise, she will check the next cue. If the employee has shown a sufficiently promising trend (i.e., $>CR_T$), the manager will make the decision to award and ignore the next cue; if not, she will check the variation cue and apply the criterion of that cue (i.e., CR_V) to make the decision.

1.2. Compensatory versus noncompensatory decision strategies

FFTs are noncompensatory strategies that differ qualitatively from compensatory strategies such as logistic regression (LR). To make decisions, the latter involve weighting and adding different cues in a manner that allows trade-offs: A less desirable value in one cue can be compensated by more desirable values in other cues in the process of assigning an overall value to a decision alternative. Compensatory decision strategies have been assumed and studied in many domains. For example, in risky choices, subjective expected utility theory and its many variants suggest that people weigh potential outcomes by their perceived probabilities and choose the option with the highest weighted sum (e.g., Kahneman & Tversky, 1979); in multi-attribute-multi-alternative decisions, the "gold" rule is to assign an importance rating to each attribute, figure out the utility of an attribute value, and multiply them to derive the expected utility of an alternative (e.g., Payne et al., 1993); and in decision analysis pertaining to classification, LR models are widely used (e.g., Green & Mehr, 1997).

In contrast to compensatory strategies, people adopting noncompensatory strategies such as FFTs do not decide by trading-off cue values, but instead search and consider cues in a certain order and stop whenever the value on a cue indicates a decision. The unconsidered cues have no effect on the decision outcome even if their values all point to the opposite direction. In hiring, for instance, a manager using a noncompensatory strategy may reject a candidate graduating from an unknown university despite the candidate's good grades and abundant working experience. Because noncompensatory strategies usually do not consider all available information and are simple to implement, they are often referred to as heuristics, such as the priority heuristic for risky choices (e.g., Brandstätter, Gigerenzer, & Hertwig, 2006), the elimination-by-aspect heuristic for multi-attribute-multi-alternative decisions (e.g., Tversky, 1972), and the take-the-best heuristic for paired-comparisons (e.g., Gigerenzer & Goldstein, 1996).[2]

In the context of our studies, a manager using the compensatory LR would estimate the regression coefficient for each cue, weight cues by their coefficients, add up the weighted components, and make a decision by comparing the sum with a criterion value. This makes it possible for the manger to arrive at a positive decision (e.g., award a bonus) on an employee who has mixed cue values (e.g., a low performance mean but an upward trend), as long as the overall evidence strength surpasses the decision criterion. This, however, may not be the case for another manager using the FFT shown on the left side of Fig. 2: If the performance mean is deemed too low (i.e., $<CR_M$), this manager will not make the "award" decision no matter how promising the trend is.

[2] In terms of decision outcomes, some noncompensatory models make the same predictions as linear models with specific configurations of predictor weights (e.g., Berg & Hoffrage, 2008). However, they will arrive at the decisions with very different processes, an aspect that is the central to our investigation.

The main goal of our studies was to examine which type of strategies, LR or FFTs, could describe participants' decision processes better. Most research on personnel decisions and managerial decision making has assumed that managers adopt compensatory strategies and analyze data accordingly. Although the possibility of managers using heuristics is often assumed and discussed (e.g., Artinger et al., 2015; Simon, 1947), there has been no study that compares compensatory and noncompensatory strategies directly regarding how well they can describe the underlying decision process.

There are two main reasons why we expect individuals to use FFTs when making performance-based decisions. First, noncompensatory strategies, including FFTs, impose lower computational demand on the cognitive system, making them more feasible for decision making in complex tasks such as performance-based decisions (e.g., Payne et al., 1993). This is consistent with Simon's argument for both human beings in general and managers in specific (1947, 1990) that decision-makers seek to *satisfice*, coming up with good enough solutions by using mostly heuristics for the tasks at hand. Empirical studies in many domains suggest that people indeed often decide with noncompensatory strategies similar to FFTs (e.g., Brandstätter et al., 2006; Bröder, 2011; Ford, Schmitt, Schechtman, Hults, & Doherty, 1989; Kohli & Jedidi, 2007; Lopes, 1995).

Second, the nature of FFTs fits the task. Specifically, compensatory, optimization-oriented strategies work better in "small worlds" in which everything is known and calculable (Savage, 1954). However, performance-based decisions are characterized by unclear utilities, unknown probabilities, and often multiple goals. These conditions severely restrict the effectiveness of compensatory strategies in finding the optimal solutions. Noncompensatory strategies, meanwhile, often lead to performance as good as or better than compensatory strategies in tasks laden with uncertainty, because it makes sense for decision-makers in such tasks to ignore noisy information in order to make robust predictions (e.g., Gigerenzer, 2008). Indeed, studies on FFTs have shown that they can achieve accuracy and payoffs similar to or better than those of compensatory strategies in both simulated and consequential real-world tasks, such as medical diagnosis and emergency triage (e.g., Green & Mehr, 1997; Luan et al., 2011; Martignon, Katsikopoulos, & Woike, 2008).

Consistent with the above arguments, empirical evidence suggests FFTs as good descriptive models. The evidence comes from both tasks in which cues are many—so that weighting and adding cues would be more challenging (e.g., Dhami, 2003; Dhami & Ayton, 2001)—and tasks with only a few cues available (e.g., Fific, Little, & Nosofsky, 2010; Tan et al., 2017). For instance, Dhami (2003) found that there were 25 cues that judges in London could use to decide whether to make a punitive decision in a bail hearing. Upon inspecting 342 decisions made by the judges, she discovered that the FFT shown on the right side of Fig. 2 could predict judges' decisions better than a linear model using all 25 cues.

Thus, we expect that FFTs could describe the decision processes of a significant proportion of participants in our studies.

1.3. The influence of task environment on decision processes

Simon's work on bounded rationality (1955, 1990) has been hugely influential in the study of human decision making. He summarized the essence of bounded rationality 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" (1990, p. 7). Following Simon, there has been much research on the interplay between decision strategies and task environments and whether people can adapt the use of a strategy to the demand of a given task

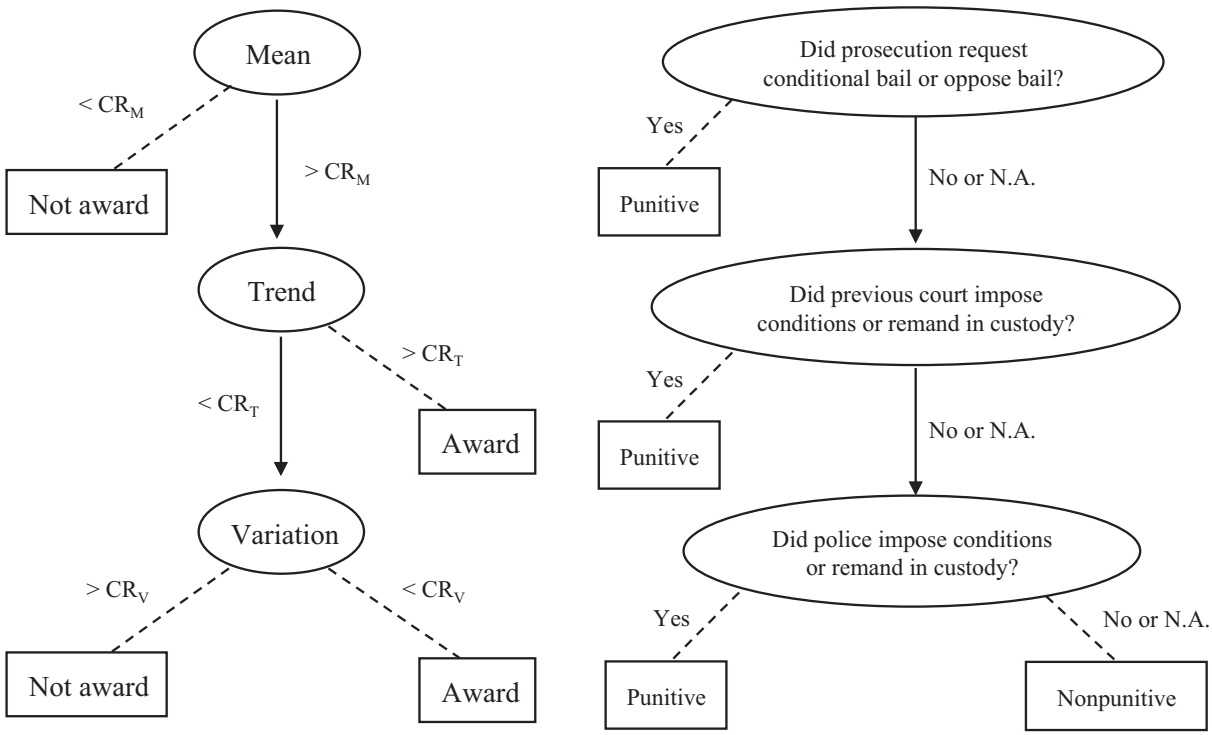

Fig. 2. Two fast-and-frugal trees (FFTs). Left: A hypothetical FFT that a manager might use to decide whether to award a bonus to an employee. Right: An FFT describing how judges in London might decide whether to make a punitive decision in a bail hearing (Dhami, 2003). CR = Criterion; M = mean; T = trend; V = variation; N.A. = not available.

(e.g., Gigerenzer & Selten, 2002; Payne et al., 1993; Todd et al., 2012). Studies of this sort, however, are largely absent in research of personnel decisions.

In the present study, we investigated whether participants could apply FFTs adaptively by manipulating the required distribution of positive and negative decision outcomes. Such requirements are common for both personnel decisions and appraisal systems such as forced distribution rating systems (e.g., Lawler, 2003; Scullen, Bergey, & Aiman-Smith, 2005). Whereas it is difficult to specify *a priori* how participants adopting an LR strategy would adjust their strategy (e.g., by applying systematically different cue weights) in response to this characteristic of the task, participants adopting FFTs could adjust two key features of the heuristic—the exit structure and the criterion values of the cues—in directions that are justifiably suitable for a specific distribution requirement.

With m cues, a decision maker can choose one of the 2^{m-1} exit structures that are possible for an FFT. The four exit structures in the case of three cues are shown in the top panel of Fig. 3. There, the letter N denotes an exit with a negative decision outcome from the perspective of an employee (e.g., not getting a bonus or being terminated from employment) and the letter P an exit with a positive outcome (e.g., getting a bonus or not being terminated). We name the four FFTs, from left to right, FFT_{NN}, FFT_{NP}, FFT_{PN}, and FFT_{PP}, on the basis of the first two exits in the tree.

Luan et al. (2011) found that other things being equal, a decision maker's tendency to make negative decisions is strongest with an FFT_{NN} and reduces gradually from FFT_{NP}, to FFT_{PN}, to FFT_{PP}. In general, FFT_{NN} and FFT_{NP} are "conservative" FFTs that are biased to make negative decisions more frequently, while FFT_{PN} and FFT_{PP} are "liberal" FFTs that have the opposite tendency. Thus, when the task environment calls for a higher percentage of negative decisions, applying a conservative FFT will be more suitable, and the higher the percentage of negative decisions, the more suitable an FFT_{NN}. Conversely, when a higher percentage of positive decisions are needed, liberal FFTs (i.e., FFT_{PN} and FFT_{PP}) will be more suitable.

Another, independent way to respond to different decision distributions is to adjust the cue criterion values on an FFT, especially that of the first cue, CR_1. In general, a higher and stricter CR_1 leads to a smaller percentage of employees passing the examination of the first cue (Luan et al., 2011). Thus, decision makers should adopt a higher CR_1 when the task calls for a lower percentage of positive decisions. Adjusting the criterion values of other cues is also possible, but the effect of such adjustments will depend on the value of CR_1, because these cues are considered later. For this reason, we focus on CR_1 in the present study.

In sum, if participants could indeed apply FFTs adaptively, we expect that as the required percentage of negative decisions gets *lower*, they would show a stronger tendency to adopt FFTs with liberal exit structures (i.e., FFT_{PN} or FFT_{PP} for a 3-cue FFT and FFT_P for a 2-cue FFT; see more below) and with a lower criterion value on the first cue (i.e., CR_1).

1.4. Overview of studies and models

We conducted two studies with college students (Study 1) and highly experienced managers (Study 2). In Study 1, participants were randomly assigned to one of four between-subjects conditions: 10% and 40% bonus (i.e., they were required to award a bonus to 10% or 40% of all employees they evaluated) and 10% and 40% termination. In Study 2, to examine more directly whether participants could adapt their decision processes to changes in the task environment, we manipulated the required decision distribution *within-subjects* with two conditions: 25% bonus and 25% termination. In both studies, an employee's performance was presented to participants in charts similar to Fig. 1.

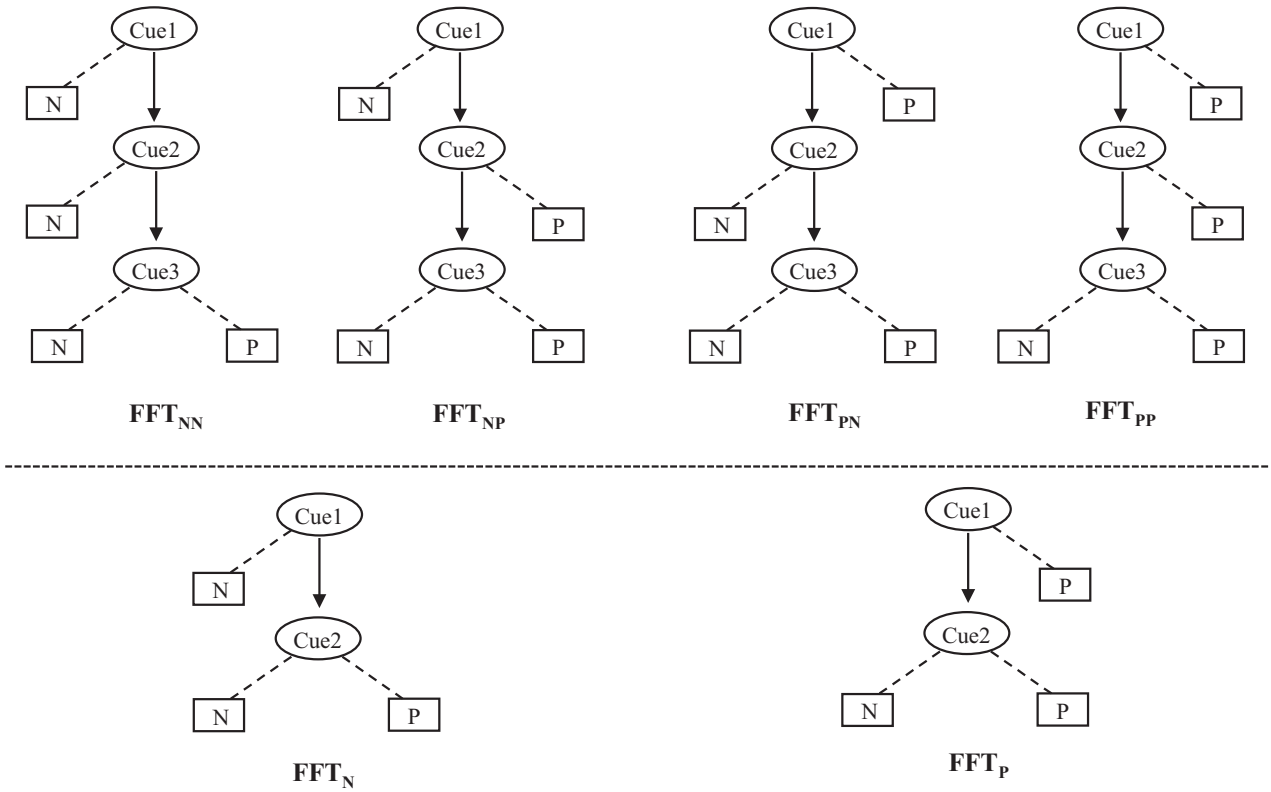

Fig. 3. Schematic representations of fast-and-frugal trees (FFTs). Top: 3-cue FFTs with four different exit structures. Bottom: 2-cue FFTs with two different exit structures. The exit labels N and P stand for negative (e.g., not getting a bonus or being terminated from employment) and positive (e.g., getting a bonus or not being terminated) decision outcomes from the perspective of an employee. The FFTs are named according to the exits in the first $m - 1$ cues, where m is the total number of cues.

In each study, we tested and compared how well two LR models, LR_3 cues and LR_2 cues, and two FFTs, FFT_3 cues and FFT_2 cues, could describe each participant's decision process. A 3-cue model, for both LR and FFT, took three cues, mean, trend, and variation, into consideration. A 2-cue model, however, only considered mean and trend, which were found much more influential than variation in some previous studies of dynamic performance (e.g., Barnes et al., 2012). For FFT_2 cues, because only the exit in the first cue is adjustable (see the lower panel in Fig. 3), it has only two possible exit structures, of which FFT_N is the conservative one and FFT_P is the liberal one. The four models were compared with regard to their ability to predict both choice and reaction time. Reaction time (RT) has long been used as a major process measure. Examining a model's ability to predict RT or RT patterns has been a common practice in cognitive modeling and often provides critical evidence for or against certain models (e.g., Fific et al., 2010; Luce, 1986; Sternberg, 1969).

In our studies, a model was evaluated based on the multiple-measure maximum likelihood (MM-ML) method proposed by Glöckner (2009). In essence, the method tests how well a model can describe multiple aspects, including choice and RT, of a participant's behavior by estimating the conditional likelihood of the data given the model. However, different from Glöckner's original procedure, we compared models not in terms of how well they fitted all trials of a participant's data; instead, we estimated parameters of each model based on half of the trials using the MM-ML method, derived the likelihood of the model in the second half of the trials with the estimated parameters, and compared models by their "cross-validated" likelihoods. Details of our model testing and comparison procedure can be found in Appendix A. The goal of the procedure was to identify which model could best *predict* a participant's choices and RTs simultaneously.

2. Study 1

2.1. Method

2.1.1. Participants

One hundred and twelve undergraduate students from a management university in Singapore participated in this study in exchange for course credit. Forty-eight percent were male, the average age was 21.1 years ($SD = 1.4$), and 89% were enrolled in a business degree program.

2.1.2. Design

Participants were instructed to assume the role of a regional supervisor in a company and to make decisions on a group of employees working in sales. Performance of an employee, in terms of the sales (in dollars) she or he made relative to the company's long-term average, in each of the past 26 weeks was shown in a graph (see a sample in Fig. 1). This graph was the employee's performance profile and provided the basis for the decision made by the participant. Each participant made decisions on 200 employees. We manipulated the required distributions of positive and negative decisions through instruction, resulting in four between-subjects conditions with 28 participants in each condition.

2.1.3. Procedure and materials

All participants completed the experiment individually on computers situated in a laboratory. After inputting demographic information, participants received a general introduction on the purpose of the experiment and were told that the sales performance of each employee would be displayed in a graph. Specific

instructions followed for each condition. In the two bonus conditions, the instructions were as follows:

> Your company gives a bonus to some of its employees. Based on performance, you need to decide whether to give each employee a bonus. Because of limited financial resources and because a bonus serves as an incentive and as recognition for good performance, not every employee can receive a bonus. In fact, you are required by company policy to give a bonus to approximately 10% [or 40%] of your employees, although slightly more or less than 10% [or 40%] is okay.

In the two termination conditions, participants received the following instructions:

> Your company is facing a difficult economic environment and needs to downsize its number of employees. Based on performance, you need to decide for each employee whether or not to terminate this person. Given current conditions faced by the organization, it is not possible to retain all employees. In fact, you are required by company policy to terminate 10% [or 40%] of your employees, although slightly more or less than 10% [or 40%] is okay.

Underlining was used in the instructions to highlight the required percentage.

After reading the instructions, participants proceeded to the profile-viewing stage. There, each participant saw the performance profiles of all 200 employees without making any decision. Four profiles, randomly selected from the pool of 200 without replacement, were shown side by side on a single screen. Participants could control the viewing time on each screen and proceed to the next screen at any time. The purpose of this procedure was to give participants an overview of how the employees compared to each other and which cues might be relevant to the impending decisions. The procedure also served to make our study more realistic, because real-life managers usually have some general understanding of their subordinates' performance before making decisions about them. After viewing all the profiles, participants entered the decision stage where they needed to make a decision on each employee. The profiles were now displayed on the screen one at a time.

The 26 points displayed in each profile were generated through a computer program with the procedure explained in Appendix B. In a nutshell, the intended values of the three cues (i.e., mean, trend, and variation) for a profile were drawn randomly and independently from three normal distributions. These procedures made the 200 profiles displayed to each participant differ from those displayed to others. Upon completion of the experimental session, which took about 40 min, participants were thanked and debriefed.

2.1.4. Data

We recorded each participant's choice and RT in each trial. Because it took participants some practice to get used to the settings of the experiment and because they tended to become less focused at the end of the experimental session, we did not include data from the first and last 10 trials in our analyses (see also Brandstätter et al., 2006); thus, for each participant, we performed analyses on 180 trials. Among the 180 trials, there were some in which the RTs were abnormally long. We suspect that participants were either performing some activity not related to the experiment or simply taking breaks in those trials. To reduce their effects on the RT analysis, we treated any RT longer than 10,000 ms as 10,000 ms. Such trials occurred on average 2.9% of the time, and the average RT after the treatment was 2763 ms ($SD = 2116$) per trial. We also tried cutoff values other than 10,000 ms and found that they had little effect on the major results of the study.

Values of the 26 points displayed in an employee's profile were also recorded. Using these numbers, we calculated values of the three cues of a profile as the following: First, we took the average of the numbers as the mean; then, we fitted the points with a linear function, with the slope of the function representing the trend; finally, the value of variation was the variance of the residuals (i.e., the differences between the actual values of the points and those predicted by the linear function).

2.2. Results

2.2.1. What decision processes were participants using?

Following the model testing and selection procedure described in Appendix A, we identified the model that had the highest cross-validated likelihood as the best model describing a participant's decision process. Among the 112 participants, the frequencies that LR_2 cues, LR_3 cues, FFT_2 cues, and FFT_3 cues were identified as the best model were 30%, 21%, 18%, and 31%, respectively (see Fig. 4). Overall, the decision processes of a significant proportion of participants could be best described by noncompensatory FFTs ($z = 10.37$, $p < 0.001$). In fact, roughly half of the participants (49%) used FFTs, while the other half could best be described by the compensatory LR models (see also Table 1).

A closer look at the data (see Table 1, first row) shows that averaged over participants in all experimental conditions, the LR models had higher accuracy in predicting participants' decisions in the second half of the trials (i.e., in cross-validation; see Appendix A for details) than the FFTs. Why, then, were the FFTs equally descriptive of participants' decision processes as the LR models in general? Recall that to identify the best model, our method took both the decisions *and* RTs of a participant into consideration. Thus, even though the LR models were more accurate in predicting the decisions, they were worse in predicting the RTs than the FFTs, resulting in an overall tie between the two model classes.

Table 1 also shows the combined frequency of the two models in each class as the best models in each experimental condition. The LR models were best models more frequently in the two 40% conditions, while the FFTs were best models more frequently in the two 10% conditions: $\chi^2(1) = 2.89$, $p = 0.089$ in a Chi-square test that compared the two 40% conditions against the two 10% conditions. As will be discussed below, the more balanced the required decision percentages (i.e., closer to 50–50), the more uncertain the decisions. Under such situations, the results in Table 1 suggest that participants were more inclined to adopt a compensatory strategy than a noncompensatory one.

2.2.2. Did participants adopt FFT features suitable for the task environment?

Besides examining which models would describe participants' decision processes better, we also tested whether participants adopted key features of FFTs suitable for the task environment. Specifically, when the required percentage of negative decisions gets lower, participants should have a stronger tendency to adopt FFTs with liberal exit structures and with a lower criterion value on the first cue (i.e., CR_1). Because FFT_3 cues was found to be an overall better model than FFT_2 cues (see Fig. 4), we focused on FFT_3 cues in this analysis. The results of the same analysis on FFT_2 cues are reported in the Supplementary Materials, and they show very similar patterns.

The analysis was conducted first under the assumption that all participants applied FFT_3 cues. This is equivalent to describing all participants' decision processes with a single model (i.e., FFT_3 cues) and examining how the two parameters of the model (i.e., exit structure and CR_1) varied across experimental conditions.

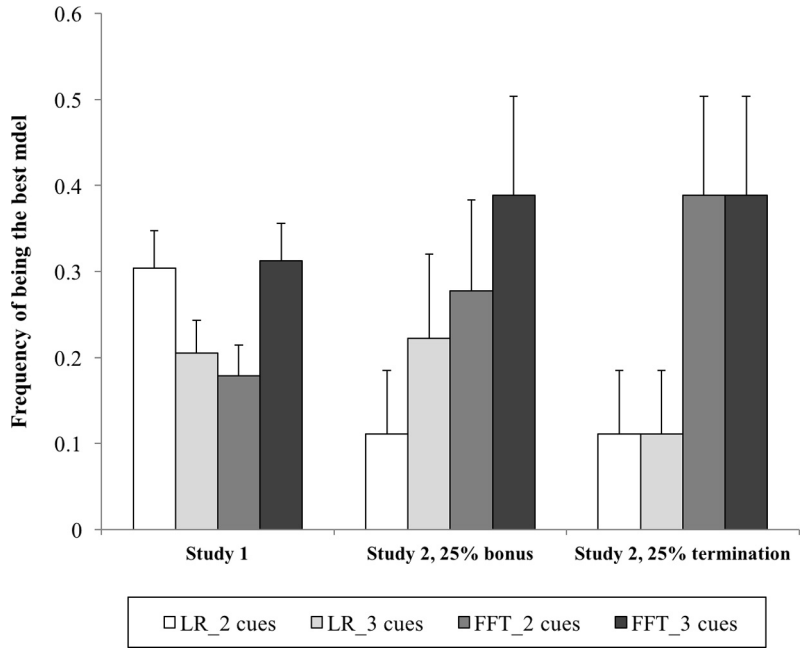

Fig. 4. The frequency of each model being identified as the best model. For Study 1, the figure shows the averaged results across the four between-subjects experimental conditions for the sake of brevity, and the detailed results in each condition can be found in Table 1. For Study 2, the results from the two within-subjects conditions are shown separately. Error bars indicate + 1 *SE*.

Table 1
Model performance and the average reaction time of participants in Study 1.

Condition	Best model (%)		Average choice prediction accuracy		Average RT (in ms)
	LRs	FFTs	LRs	FFTs	
All conditions	50.9	49.1	0.840	0.822	2763
10% bonus	46.4	53.6	0.887	0.872	2275
40% bonus	53.6	46.4	0.790	0.772	2665
40% termination	64.3	35.7	0.803	0.783	3260
10% termination	39.3	60.7	0.880	0.860	2851

Note. RT = Reaction time; LRs = LR_2 cues and LR_3 cues; FFTs = FFT_2 cues and FFT_3 cues.

The results are shown in Panels A and B of Fig. 5. With respect to exit structure (Panel A), the proportion of participants who adopted a liberal exit structure indeed increased gradually when the required percentage of negative decisions became progressively lower from the 10% bonus condition (i.e., 90% of the decisions were negative) to the 10% termination condition; and the reverse pattern held for the conservative FFTs.

Regarding CR_1, we separated cases in which participants were estimated to check the mean cue first ($N = 51$) from those in which the trend cue was checked first ($N = 61$). For each group, we calculated the average CR_1 across all participants in an experimental condition. As Panel B of Fig. 5 shows, regardless of which cue was searched first, the average CR_1 decreased gradually from the 10% bonus condition to the 10% termination condition. ANOVA tests show that both decreasing patterns were statistically significant, $F(3,47) = 38.77$, $p < 0.001$, partial $\eta^2 = 0.71$ for the mean-first cases and $F(3,57) = 19.35$, $p < 0.001$, partial $\eta^2 = 0.51$ for the trend-first cases. In general, the results for both the exit structure and CR_1 of FFT_3 cues show evidence that participants could indeed adopt features of FFTs that were suitable for the characteristics of the task.

One problem for the above analysis is the assumption that all participants adopted FFT_3 cues, which was not empirically true (Fig. 4). Panels C and D in Fig. 5 show results parallel to Panels A and B, respectively, but only for participants for whom FFT_3 cues was identified as the best model ($N = 35$). They show very similar patterns as those in Panels A and B; that is, with a higher required percentage of negative decisions, a higher proportion of participants adopted a liberal exit structure and their adopted CR_1 became lower. This analysis, however, has its own drawback: Because the total sample size was now fairly limited, the results may not be of high reliability.

In sum, we found that participants did adopt FFT features that were suitable for task environments differing in the required distributions of negative and positive decisions.

2.2.3. How did the required decision distribution affect model predictions?

In addition to testing how the required decision distributions affected participants' adoptions of key FFT features, we also examined their effect on the general descriptive ability of a model, an issue that has received little research attention to date. Table 1 (rows 2–5) shows the average accuracy of the LR models and the FFTs, respectively, in predicting participants' decisions in each experimental condition. Each model type was more accurate in the two 10% conditions than in the two 40% conditions, $F(1, 110) = 42.24$, $p < 0.001$, partial $\eta^2 = 0.28$ for the LR models and $F(1, 110) = 33.15$, $p < 0.001$, partial $\eta^2 = 0.23$ for the FFTs in ANOVA

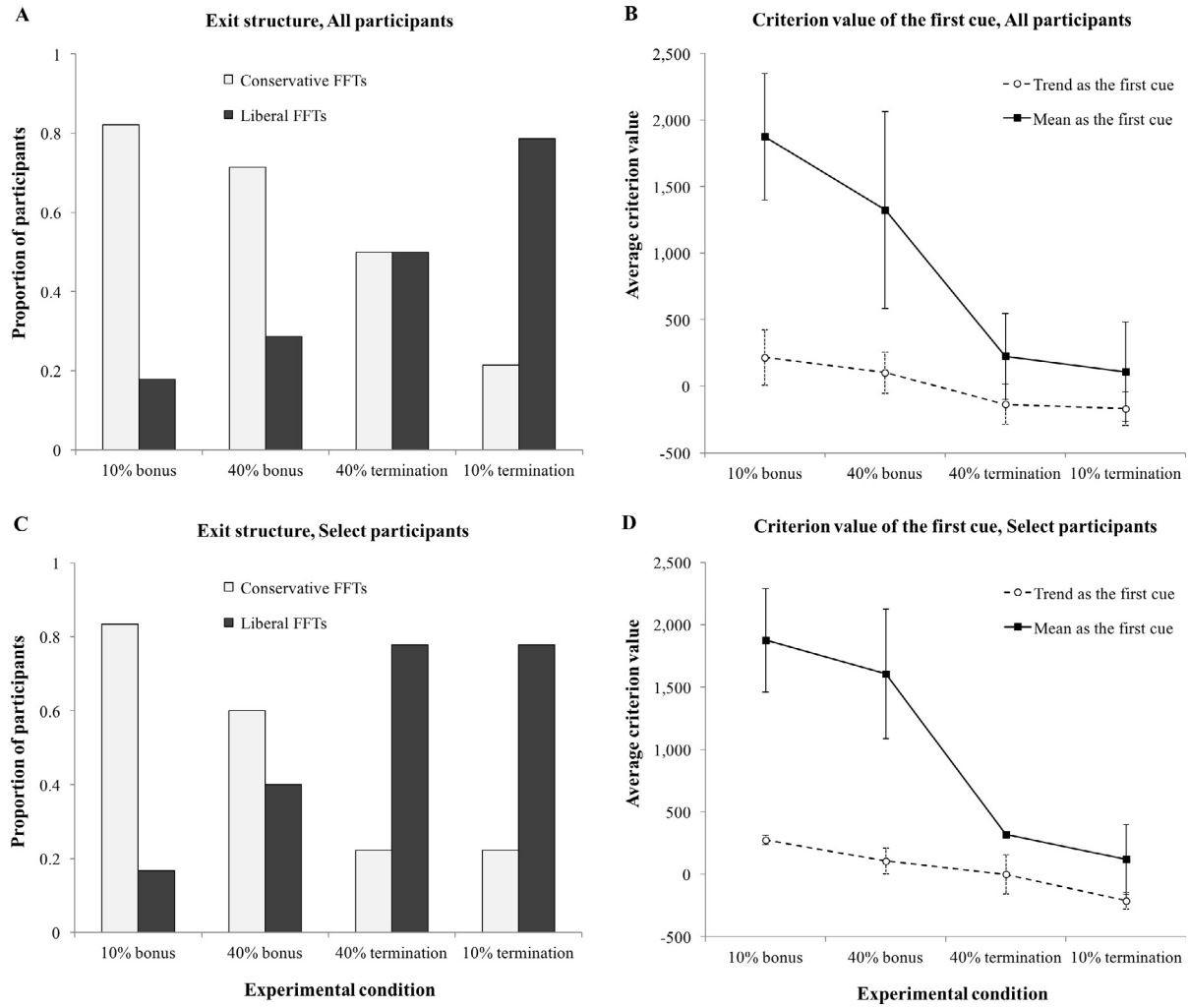

Fig. 5. Differences in two features of FFT_3 cues across the experimental conditions in Study 1. (A) Proportions of all participants (N = 112) whose decision processes could be best described by FFT_3 cues with either conservative (i.e., FFT_{NN} and FFT_{NP}) or liberal (i.e., FFT_{PN} and FFT_{PP}) exit structures. (B) Criterion value of the first cue, averaged over all participants, when either performance mean or trend was estimated to be the first cue searched. (C & D) The matching results to A & B, respectively, for participants for whom FFT_3 cues was identified as the best model (N = 35). Error bars indicate + 1 *SD*, but note that there were no error bars displayed for one data point in Panel D because there was only one participant in that specific situation.

tests that treated the two conditions with the same required percentages as one general condition. We speculate that a higher level of uncertainty in the more balanced percentage conditions (i.e., closer to 50–50) might be the main reason for this result.

Uncertainty exists in most nontrivial decisions, and it is quantified in information theory by a measure called "entropy" (Shannon, 1948). For binary choices, the highest level of entropy occurs when the required decision percentages or base rates are equal, and the more unbalanced the base rates, the lower the entropy.[3] Therefore, uncertainty was inherently higher in a 40% condition than in a 10% condition in our study. For individuals using either FFTs or LR to decide, it is assumed that they try to reduce the initial level of uncertainty by considering indicative cues and applying some rules to separate the two decision categories. However, the separation is more challenging in a 40% condition, in which the mental representations of the two categories tend to overlap more and rules of separation tend to fluctuate trial by trial to a greater extent (e.g., Benjamin, Diaz, & Wee, 2009; Erev, 1998). This should lead to not only a lower level of prediction accuracy by a model but also prolonged RTs by the participants. As shown in Table 1, the average RT was indeed longer in the 40% conditions than in the 10% conditions, $F(1, 110) = 4.56$, $p = 0.035$, partial $\eta^2 = 0.040$.

This result supports our speculation that a higher level of uncertainty might be the reason behind the lower prediction accuracy of a decision model in the 40% conditions. That being said, more studies are needed to examine the generalizability of this phenomenon and to continue exploring reasons for its occurrence.

2.2.4. What is the role of the variation cue in the decision process?

Fig. 4 shows that LR_2 cues was the best model more frequently than LR_3 cues, suggesting that when participants integrated cues compensatorily, most of them did not consider the variation cue. This is consistent with the finding in some previous studies of dynamic performance (e.g., Barnes et al., 2012). However, the result was quite the opposite for the FFTs, for which FFT_3 cues was the best model more frequently than FFT_2 cues. Thus, for participants who adopted a noncompensatory strategy, most of them did consider the variation cue.

[3] Entropy is defined as: $-[P \times \mathrm{Log}_2(P) + (1 - P) \times \mathrm{Log}_2(1-P)]$, in which P is the base rate of one of the two decision options. The highest entropy is reached when P is 0.50 (entropy = 1). The basic idea is that when there is no other way to make a decision but by guessing on the base rates, a person's decision accuracy will be lower when entropy is higher.

2.3. Discussion

Past research has shown that dynamic performance cues such as performance mean and trend influence appraisal ratings and performance-based decisions. Taking a process-oriented approach and drawing on methodologies from cognitive model testing, we examined how decision makers integrate information of dynamic performance cues in personnel decisions. Our results show that noncompensatory FFTs could describe the decision processes of nearly half of the participants, challenging previous studies' assumption that only compensatory strategies are adopted and demonstrating the importance of considering both compensatory and noncompensatory strategies in understanding personnel decision processes.

We further examined what FFT features participants adopted when applying FFTs in different experimental conditions that varied on the required distribution of negative and positive decisions. The results show that both the exit structure and the criterion value of the first cue adopted by participants differed across the conditions, and these differences were consistent with what was expected if they would behave adaptively. Hence, not only did many participants decide using FFTs, but they also applied key features of this heuristic in directions suitable for the demand of a task, providing evidence for the "adaptive decision makers" argument (e.g., Payne et al., 1993; Simon, 1990; Todd et al., 2012) in the context of personnel decisions.

In addition to its effects on the adopted features of FFTs, the required decision distribution also affected the prediction accuracy of a decision model systematically: Both FFTs and LR were more accurate when the required decision percentages were more extreme than when they were more balanced. To the best of our knowledge, there has been no study reporting such a finding. Therefore, even though we have an explanation (i.e., an inherently higher level of uncertainty in the 40% conditions), both the finding and the explanation need to be further examined in future studies.

We also found that the variation cue was used more frequently by participants adopting FFTs than those adopting LR. There could be many reasons for this result. For example, integrating three cues compensatorily could add too much computational demand to many participants, so that they opted to process only the two most important cues instead (e.g., Payne et al., 1993). For FFTs, however, the demand for processing the third cue is much less; thus, most participants could "afford" considering the variation cue. To pin down the exact reason for this result is beyond the scope of the current study. Nonetheless, its implication is important: Whether and to what extent the variation cue plays a role in performance-based personnel decisions may depend on the kind of strategy, compensatory or noncompensatory, an individual applies in the decision-making process.

3. Study 2

The participants of Study 1 were "novice" managers who had little real experience of supervising and evaluating others. What strategies would more experienced decision makers use then? They might be more capable of processing multiple cues simultaneously, enabling them to adopt and apply a compensatory strategy with relative ease. However, there is also evidence in research of expert decision-making showing that experts actually have a stronger tendency to process cues noncompensatorily and heuristically than the less experienced (e.g., Garcia-Retamero & Dhami, 2009; Pachur & Marinello, 2013). The main goal of Study 2 was to test this matter by recruiting highly experienced managers and finding out what strategies they tend to use.

Moreover, we applied a between-subjects design in Study 1 to investigate whether participants could adjust features of FFTs adaptively in response to the changing task environment. A within-subjects design, however, should be better suited to address this adaptiveness question. Therefore, we asked participant in this study to make decisions in not one but two conditions that differed in the required decision distribution.

3.1. Method

3.1.1. Participants

Eighteen managers working in eight different industries, such as finance, consulting, and manufacturing, participated in this study on a voluntary basis. Four were female, 12 identified themselves as having an upper-management position, the average age was 49.7 years ($SD = 5.9$), and they reported having supervised and evaluated lower-level employees for 22.1 ($SD = 5.6$) and 19.5 ($SD = 7.3$) years, respectively.

3.1.2. Stimuli, design, and procedure

The stimuli (i.e., employees' performance profiles) were generated by the same procedure as in Study 1 (see Appendix B). The experimental procedure was the same, as well. However, there were a couple of changes in the design: First, we used a within-subjects design with two conditions, 25% bonus and 25% termination, and counterbalanced the condition orders among the participants; and second, in each condition, each participant made decisions on 108 employees. Because of the within-subjects design, we had to reduce both the number of conditions and the total number of trials in a condition in consideration of fatigue and participants' limited time availability.

3.1.3. Data

For the same reasons as in Study 1, we excluded participants' data in the first and last four trials in each experimental condition from analysis. Because the average RT for participants in this study was much longer than that in Study 1, instead of using 10,000 ms as the cutoff, we treated RTs longer than 15,000 ms as 15,000 ms, which occurred on average 4.2% of the time. After this treatment, the average RT was 4736 ms ($SD = 3506$) and 4934 ms ($SD = 3569$) per trial for the bonus and termination conditions, respectively.

3.2. Results and discussion

What strategies did the experienced managers adopt? Fig. 4 shows the frequency of each decision model being identified as the best model in each of the two experimental conditions. The combined frequency of FFTs being the best models was much higher than that of the LR models in the 25% bonus condition (67% vs. 33%), and the difference was even more pronounced in the 25% termination condition (78% vs. 22%). Therefore, in comparison to the college students in Study 1, the experienced managers were even *more* likely to apply a noncompensatory FFT to make decisions.

Because FFT_3 cues was overall more prevalent than FFT_2 cues, we focused on FFT_3 cues to examine whether participants adjusted key FFT features adaptively in response to the changing decision distribution requirement. The matching, and similar, results for FFT_2 cues can be found in the Supplementary Materials. Moreover, because of the limited sample size in this study and the more dominant presence of FFTs as the best models for the participants, our analysis was conducted on the assumption that all participants adopted FFT_3 cues to make decisions.

To apply FFT_3 cues adaptively in this study, participants were expected to adopt a conservative FFT (i.e., FFT_{NN} or FFT_{NP}) in the

25% bonus condition and a liberal FFT (i.e., FFT_{PN} or FFT_{PP}) in the 25% termination condition. The left side of Fig. 6 shows that the majority of participants indeed adopted the expected exit structure in each experimental condition. In addition to exit structure, participants could also adjust the criterion value of the first cue (CR_1) between the two conditions. The right side of Fig. 6 shows the differences in CR_1 between the 25% bonus and 25% termination conditions (former minus latter) for 14 participants. Results of the other four participants were not available because their estimated cue orders changed between conditions, making the estimated criterion values unsuitable for comparison.[4] Among the 14 participants, 10 were estimated to search the mean cue first and four to search the trend cue first, and 12 of the 14 participants adjusted CR_1 in the expected direction; that is, they adopted a higher (stricter) CR_1 in the 25% bonus condition than in the 25% termination condition.

The within-subjects design of this study allowed us to examine how *each* participant adjusted the exit structure and CR_1 between the two experimental conditions. Table 2 shows how many participants fell in each adjustment category based on the adaptiveness criteria described above. In general, this individual-level analysis indicates that the large majority of participants (i.e., 15 of 18) made adaptive adjustments on at least one FFT feature when the decision distribution requirement changed. Compared to Study 1, these within-subjects comparisons support even more strongly the argument that managers adaptively apply FFTs for personnel decisions under changing task environments.

In sum, results of this study show that (a) compared to the college students in Study 1, a higher proportion—and the clear majority—of the experienced managers decided using noncompensatory FFTs; and (b) most of them adjusted key FFT features adaptively when the required decision distribution changed.

4. General discussion

Decision making is an integral part of management, and personnel decisions are among the most important decisions managers need to make (e.g., Guion, 2011; March & Simon, 1958). Drawing on research on decision heuristics (e.g., Luan et al., 2011; Todd et al., 2012) and methods in cognitive modeling (e.g., Geisser, 1993; Glöckner, 2009), we took a process-oriented approach in the present study, testing and comparing how well noncompensatory FFTs and compensatory LR models could describe the process underlying performance-based personnel decisions. Moreover, building on Simon's notion of bounded rationality that views rational decision making as the result of adapting strategies to the task environment, we also investigated whether decision makers could apply key features of FFTs adaptively in response to changes in the task environment when making performance-based personnel decisions.

We found that nearly half of the college students in Study 1 applied FFTs to make decisions and at least two-thirds of the experienced managers in Study 2 did so as well, showing that noncompensatory strategies are commonly used for performance-based personnel decisions. Moreover, manipulating the required distribution of positive and negative decisions, a common aspect of personnel decisions, we found that most participants adopted suitable FFT features (i.e., exit structure and criterion value of the first cue) in and across different task conditions. Finally, our results show that many participants, especially the ones applying FFTs, did consider the variation cue in making their decisions, clarifying the role of this cue in the decision-making process.

In what follows, we first discuss contributions of our findings to relevant research areas. We then argue for the importance of studying process models in management research and also provide recommendations for management practice in light of our findings. After pointing out limitations of our study, we conclude by suggesting some directions for future research on heuristics in managerial decisions.

4.1. Dynamic performance-based personnel decisions

The most direct contribution of our study is that it extends the existing work in dynamic performance appraisal and decisions. Employee performance changes over time, and past research has shown that dynamic performance cues such as performance mean and trend can significantly influence managers' summary ratings and decisions on the employees (e.g., Barnes et al., 2012; Reb & Cropanzano, 2007). However, the research has told us little about how managers use these cues to arrive at their decisions. Our studies addressed this limitation by directly comparing two types of models that assume qualitatively different processes and testing which model could better predict a participant's decisions and RTs.

Overall, contrary to previous studies' assumption that decision makers process cues in parallel and integrate cue information in a compensatory way, a substantial proportion of our participants actually decided using a sequential and noncompensatory strategy. Whether the process of cue information integration is mainly compensatory or noncompensatory has been a long-standing debate in economics, psychology, and marketing (e.g., Brandstätter et al., 2006; Payne et al., 1993; Todd et al., 2012). This debate has rarely appeared in management research, likely because of the variable-oriented approach—that is, trying to find out *what* cues or variables affect outcomes (Mohr, 1982)—predominantly applied in the field. Taking instead a process-oriented approach (i.e., trying to find out *how* individuals judge and decide), our study brought this debate into performance-based personnel decisions, shedding light on the processes underlying these important managerial decisions.

Moving forward, it will be useful to explore the environmental and personal factors that may affect individuals' strategy preferences in making performance-based personnel decisions. Our studies show that more experienced participants had a stronger tendency to apply noncompensatory heuristics, consistent with findings from other domains (e.g., Garcia-Retamero & Dhami, 2009; Pachur & Marinello, 2013). However, departing from findings in risky choices (e.g., Brandstätter et al., 2006), we found that participants generally preferred compensatory strategies when there was a higher level of uncertainty in the decisions (Table 1). More studies are needed to test the generalizability of these results and the effects of other potentially relevant variables, such as task complexity and the possibility of learning (e.g., Todd et al., 2012), on individuals' strategy preferences.

4.2. The ecological rationality of decision strategies

Previous research has examined the influence of task characteristics on people's use of information. For example, applying a variable-oriented approach, Reb and Greguras (2010) examined differences in the amount of variance explained by dynamic performance cues in tasks with varied purposes. They found that performance mean had a stronger influence (i.e., explaining more variance) on appraisals when appraisals were made for administrative purposes, whereas trend had a stronger influence when

[4] Interestingly, all four participants switched from searching the mean cue first in the 25% bonus condition to searching the trend cue first in the 25% termination condition. A study by Reb and Greguras (2010) found that mean mattered more for past-oriented rating purposes while trend mattered more for future-oriented purposes. Future research could examine in more detail whether it is indeed the case that experienced decision makers tend to deem mean or trend as the most important cue, respectively, for bonus (more past-oriented) and termination (more future-oriented) decisions.

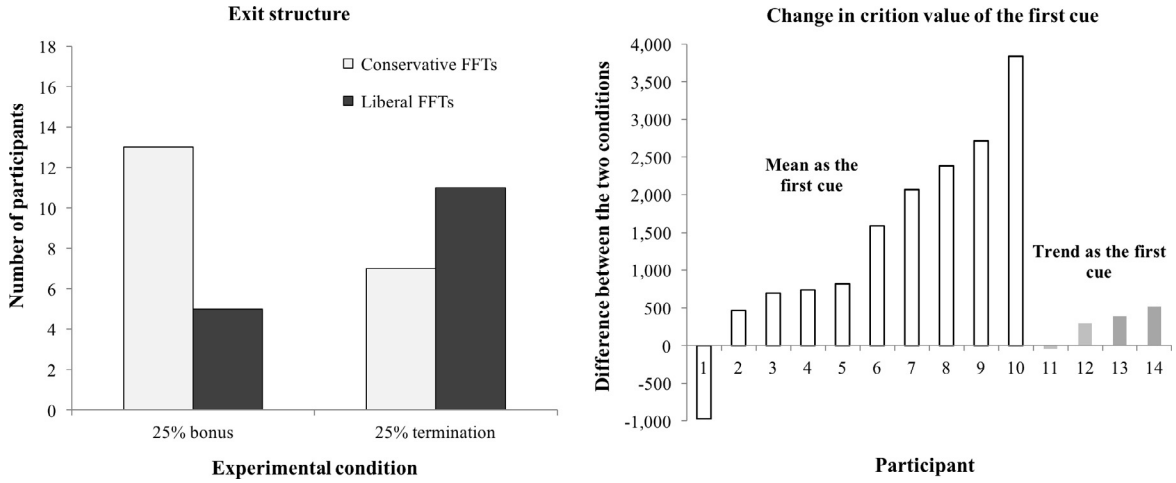

Fig. 6. Differences in two features of FFT_3 cues between the experimental conditions in Study 2, assuming that all participants adopted FFT_3 cues. Left: Number of participants whose decision processes could be best described by FFT_3 cues with either conservative (i.e., FFT_{NN} and FFT_{NP}) or liberal (i.e., FFT_{PN} and FFT_{PP}) exit structures. Right: Change in the criterion value of the first cue (i.e., the estimated value in the 25% bonus condition minus that in the 25% termination condition) for 14 participants. Note that the differences for the "trend-first" participants were smaller because trend was quantified in a smaller unit than mean.

Table 2
Participants classified by their adaptive adjustments of two FFT features between experimental conditions of Study 2.

Category	Number of participants
Adjusted both features adaptively	4
Adjusted one feature adaptively and the other unchanged	9
Adjusted one feature adaptively and the other unable to judge	2
Adjusted one feature adaptively but the other non-adaptively	2
Adjusted both features non-adaptively	1

Note. The two features are the exit structure and the criterion value of the first cue of FFT_3 cues. FFT = fast-and-frugal trees.

appraisals were for developmental purposes. Our study extends the understanding of how task environment may affect the use of dynamic performance cues in two major ways. First, we show that in addition to the amount of variance explained, changes in the task environment can also affect decision processing, such as the features of FFTs adopted by decision makers; and second, we found that the adjustments decision makers made were generally adaptive to the task environment.

Like organisms in the natural environment, organizations and managers need to adopt strategies suitable to the business environment in order to gain a competitive edge. However, few studies in managerial decision making have attempted to understand the influence of task characteristics through the perspective of adaptation. The main reason, we suspect, is the lack of good theories and normative analyses of model–environment interactions. To that end, arguments, findings, and methodological tools from the ecological rationality program can be quite useful (e.g., Gigerenzer & Selten, 2002; Todd et al., 2012). In essence, the program follows Simon's bounded rationality framework but emphasizes more the effects of environment on human behavior. It argues that good performance arises when a strategy matches well with characteristics of the environment and that comparative model testing is needed to find out which strategy is better suited for a given task.

In managerial decision making, it is often difficult to judge the quality of a decision and in turn the performance of a certain strategy. Therefore, the focus is usually on examining the descriptiveness of a strategy. Even so, adaptive arguments can still be made based on careful analyses and previous findings and comparative model testing needs to be conducted to see which strategy or strategies managers actually apply. The present study can be viewed as an application of these approaches of ecological rationality in the context of performance-based personnel decisions. Given the insights we have gained with these approaches, we believe that they should be applied more widely in the research of managerial decision making.

4.3. FFTs as models of managerial decisions

It has been argued that heuristics are more likely to work well and be applied in tasks nested in complex, dynamic, and competitive environments (e.g., Gigerenzer, 2008; Simon, 1955, 1957). Past studies of FFTs have shown that they are both prescriptively useful and descriptively valid in a variety of such domains, including medical (e.g., Green & Mehr, 1997), legal (e.g., Dhami, 2003), social (e.g., Hertwig et al., 2013), and military (e.g., Keller & Katsikopoulos, 2016). Our study contributes to the growing research on FFTs by testing them in a domain that is in every sense as challenging as the others: personnel decisions. Building on the methodologies applied in our and other studies of FFTs, the validity of FFTs can be further tested in other types of managerial decisions.

Within the domain of personnel decisions, it would be interesting to examine FFTs' ability to capture selection decisions. Whom to hire constitutes one of the most important managerial decisions, and organizations devote tremendous resources to recruiting, including large monetary amounts (e.g., US $124 billion in 2011 according to a report by Bersin & Associates; Leonard, 2011). For important positions in particular, a typical recruiting process involves various technology-assisted and face-to-face stages, including resume screening, assessment centers, and interviews (Hough & Oswald, 2000). From our perspective, given their sequential nature, these stages can be viewed as cues and the whole process can be conceptualized as an FFT, whose exits lead to either offering an applicant the position (i.e., a positive decision) or rejecting the applicant (i.e., a negative decision). The study of FFTs as models of the selection process could offer novel insights into how selection decisions are made, as well as address issues of validity and utility that have been central to selection research (e.g., Sackett & Lievens, 2008; Schmidt & Hunter, 1998).

4.4. Toward process models of managerial decision making

We took a process-oriented approach in this study, aiming to find out how different models could describe participants'

underlying decision processes. Our approach differs in several aspects from other process-oriented approaches in management research. First, whereas we frame our process models as descriptive of human behavior, many process models in the field are normative and prescriptive, trying to specify the various rational stages of decision making (e.g., Russo & Schoemaker, 2012). Such models draw heavily on the notions of *homo economicus* and utility maximization and should be better considered as *as-if* models (e.g., Berg & Gigerenzer, 2010); that is, these models do not pay attention to how individuals calculate utilities but only assume that they behave *as if* they do so. Second, some models focus on organizational processes, such as communication and coordination among different actors (e.g., the garbage can model; Cohen, March, & Olsen, 1972), rather than the mental processes of managers. Thus, they are process models of management at a different level from ours. And third, previous process-oriented studies tended either to be theoretical in nature or to use qualitative methods, including in-depth case studies, interviews, or grounded theory building, to study processes. This is consistent with a distinction that treats variable-oriented research as characterized by quantitative methods and process-oriented research as characterized by qualitative methods (Mohr, 1982). Transcending this rather narrow distinction, our research demonstrates the value of quantitative methods (e.g., the MM-ML-based modeling testing procedure) in process-oriented research.

In general, process models try to describe the underlying processes of individual behavior by taking human cognitive abilities into consideration, being as precise as possible, and generating falsifiable hypotheses. Because they look into the transitional states between input and output and make explicit assumptions about how these processes unfold, a richer set of hypotheses can be derived from process models than as-if models that focus only on explaining behavioral outcomes. Generating more hypotheses that can be falsified by a wider range of data may make it more difficult for a process model to find support. However, the credibility of the model will be strengthened if the data do support the hypotheses. Either way, something valuable can be learned about the model and the processes underlying the phenomena it tries to describe.

That being said, we acknowledge that depending on the goal of a study, both process-oriented and variable-oriented approaches can be useful and should complement each other. However, given the relative dearth of process-oriented research in management, we argue that more such research would be particularly important to advance our understanding of managerial decision making.

4.5. Practical implications

Findings from our study and previous research on heuristics have at least two practical implications for management. First, just as the experienced managers in Study 2 did, professionals and experts often approach decisions of serious consequences in ways consistent with heuristics, considering only a small number of cues and processing them hierarchically (e.g., Dhami, 2003; Dhami & Ayton, 2001; Garcia-Retamero & Dhami, 2009; Pachur & Marinello, 2013). What cues they look at and how they prioritize these cues for processing represent a large part of their expertise, which is often the result of years of working in the field and many costly trial-and-error experiences. Therefore, organizations could benefit by spending more effort on discovering what heuristics their experienced managers use and understanding how, why, and when those heuristics work (e.g., Artinger et al., 2015). Such knowledge can be an invaluable asset to an organization as well as highly instructive for training purposes.

Second, heuristics such as FFTs generally have transparent structures and are easy to understand. For instance, once it is clear that managers use the FFT shown on the left side of Fig. 2 to make bonus decisions, employees will know what aspects of their performance matter (i.e., the cues) and what levels of performance are considered sufficient (i.e., the cue criteria). It has been shown that the lack of policy transparency can reduce trust in an organization from both its employees and the public (e.g., Norman, Avolio, & Luthans, 2010; Palanski, Kahai, & Yammarino, 2011). Thus, adopting policies with heuristic structures, or at least framing them in that way, may help organizations communicate their policies more effectively and improve their perceived trustworthiness.

4.6. Limitations and future directions

Just as with other research, ours is not without limitations. One limitation is that we asked participants to make a large number of decisions based on hypothetical performance profiles. We adopted this approach to ensure that there would be sufficient variance in each cue so that we could measure its effect on decisions and that there would be sufficient data for us to conduct model comparisons at the individual level. This approach is common, and to a large extent even necessary, in cognitive modeling (e.g., Glöckner, 2009; Lewandowsky & Farrell, 2011). However, it does raise the question of external validity of whether the less cognitively demanding FFTs might have an advantage given the limited time available for each decision. We are to some extent reassured by the finding that the experienced managers in Study 2 were even more likely to use FFTs than the novices in Study 1. This is consistent with other research showing that heuristics are not only used when more complex methods are not feasible (e.g., Payne et al., 1993) but also by experts as their preferred decision strategies (e.g., Klein, 2003; Pachur & Marinello, 2013). Nevertheless, future research could use other methods, such as in-depth interviews or think-aloud protocol analysis, to triangulate our findings (e.g., Mohr, 1982).

Moreover, in examining LR and FFTs as models for performance-based decisions, we built on previous research on dynamic performance evaluation that focuses on the three cues of performance mean, trend, and variation. There is no doubt that for a different managerial decision task, a different set of cues will emerge. For example, managers' internal capital allocation decisions may be influenced by cues such as the total number of business units over which allocations are distributed and the investment opportunities, profitability, and growth of different units (e.g., Bardolet, Fox, & Lovallo, 2011). FFTs, in essence, depict possible ways in which managers may integrate cue information. By maintaining the basic structure but changing the cue contents, we believe that FFTs could be applied to model various managerial decisions, including those related to supply, production, marketing and sales, human resources, financing, and strategy. Furthermore, one common challenge for cue-based models is that not all cues can be conveniently expressed in quantitative form (e.g., the personality of an employee). However, this may be less of a challenge for FFTs, because the exit conditions set on the cues can be either numerical or logical. Future studies could explore how managers make decisions using qualitative cues, testing how well FFTs and other models describe the corresponding decision processes.

Finally, FFTs are only one type of "fast-and-frugal" heuristics studied in the ecological rationality paradigm. There are many others, such as take-the-best (Gigerenzer & Goldstein, 1996), Δ-inference (Luan, Schooler, & Gigerenzer, 2014), the recognition heuristic (Goldstein & Gigerenzer, 2002), and the $1/N$ heuristic (Hertwig, Davis, & Sulloway, 2002). They have been applied to a multitude of tasks, ranging from inference of city populations to parental investment to sports forecasting. Given the assortment of activities required in management, we believe that there must be tasks in which these heuristics are plausible and useful; thus,

it would be a promising direction for future research to test them as descriptive as well as prescriptive models for a variety of managerial tasks.

Appendix A. Method of model testing and comparison

Many criteria have been used for model testing and comparison (see reviews by Myung, Forster, & Browne, 2000; and Shiffrin, Lee, Kim, & Wagenmakers, 2008). In this research, we used maximum likelihood as the comparison criterion, following a modified procedure of the multiple-measure maximum likelihood (MM-ML) method by Glöckner (2009). For the sake of brevity, we will not describe the MM-ML method in detail here but refer interested readers to Glöckner's article. In a nutshell, MM-ML fits model parameters by maximizing the likelihood of observed data, which in our studies were participants' choices and RTs, and compares models in terms of their fitted maximum likelihoods.

In Glöckner's original method, a model's maximum likelihood is corrected for the number of free parameters and the number of observations based on which the parameters are estimated. The result is the Bayesian Information Criterion (BIC), and whichever model has the lowest BIC is selected. In theory, BIC approximates the marginal probability of a model (e.g., Shiffrin et al., 2008). However, its practical implications are not very clear, especially when data are limited and model estimations come with high degrees of uncertainty. In our studies, we adopted a different procedure, deriving a model's parameters using the MM-ML method but comparing models on how well they predicted unknown data.

Key to our procedure is the split-half cross validation (e.g., Geisser, 1993; Stone, 1974). Specifically, for each participant's data that consisted of n trials, we first estimated the parameters of a model in the first $n/2$ trials (i.e., the learning set) using the MM-ML method. We then applied the model with the estimated parameters to the second half of the trials (i.e., the generalization set) and calculated the likelihood of the data given the model there. The model with the highest cross-validated likelihood was chosen as the model most likely for a participant. The advantage of this procedure is its clear relevance to prediction: Instead of penalizing models with more free parameters like in BIC, the models are compared in their ability to predict unknown data, a criterion mattering more to practice than fitting, regardless how many free parameters they have.

For LR_2 cues, six parameters were estimated, including three linear parameters (i.e., one constant and two beta weights), one error rate in applying the model, and two parameters for RTs (i.e., the mean and standard deviation). For FFT_2 cues, eight parameters were estimated: cue order (i.e., mean or trend as the first cue), exit structure (i.e., FFT_N or FFT_P), criterion values for the two cues, error rate, and three RT parameters (i.e., mean, standard deviation, and a scaling parameter; see Glöckner, 2009). The two three-cue models each had one more parameter than their two-cue counterpart: the beta weight of the third cue for LR_3 cues and the criterion value of the third cue for FFT_3 cues. Note that the MM-ML method assumes that for an LR model, there are no true cross-trial differences in a participant's RTs and the observed differences are caused by random factors whose effects are summarized by the standard deviation of the observed RTs; however, for an FFT, the method assumes that there *are* true cross-trial RT differences because individuals are expected to search cues sequentially and to stop searching as soon as the exit condition for a decision is met; as a result, they may stop at different cues in the search hierarchy from trial to trial, leading to different RTs across trials. We followed the same assumptions in our model testing.

Appendix B. Procedure for generating points in a performance profile

In both Studies 1 and 2, the 26 points in the profile of an employee i were generated through a computer program with the following procedure:

Step 1. *Generating 26 original numbers* O_{1-26}. The intended mean of these numbers, m_i, was randomly drawn from a normal distribution $N(0,1)$, and the intended variation, v_i, was randomly drawn from another normal distribution $N(2,0.25)$. O_{1-26} were then randomly drawn from a new normal distribution with the mean at m_i and variance at v_i.

Step 2. *Generating 26 numbers with a certain trend* T_{1-26}. The intended trend t_i was randomly drawn from a normal distribution $N(0,0.25)$ and was multiplied with 26 constants that ranged from -12.5 to 12.5 with an increasing step of 0.5. The 26 products resulting from this operation were then added to O_{1-26} to create T_{1-26}.

Step 3. *Generating 26 numbers shown in the profile* P_{1-26}. T_{1-26} were multiplied by 1000. If a resulting number was higher than 10,000 or lower than $-10{,}000$, it was rounded to 10,000 or $-10{,}000$, respectively. The numbers resulting from this procedure were P_{1-26}.

All profiles were generated according to this procedure for each participant and were stored in a database once the experiment started.

Appendix C. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.obhdp.2017.05.003.

References

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

Bardolet, D., Fox, C. R., & Lovallo, D. (2011). Corporate capital allocation: A behavioral perspective. *Strategic Management Journal, 32*, 1465–1483.

Barnes, C. M., Reb, J., & Ang, D. (2012). More than just the mean: Moving to a dynamic view of performance-based compensation. *Journal of Applied Psychology, 97*, 711–718.

Benjamin, A. S., Diaz, M. L., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychological Review, 116*, 84–114.

Berg, N., & Gigerenzer, G. (2010). As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas, 18*, 133–165.

Berg, N., & Hoffrage, U. (2008). Rational ignoring with unbounded cognitive capacity. *Journal of Economic Psychology, 29*, 792–809.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Choices without tradeoffs. *Psychological Review, 113*, 409–432.

Bröder, A. (2011). The quest for take-the-best: Insights and outlooks from experimental research. In G. Gigerenzer, R. Hertwig, & T. Pachur (Eds.), *Heuristics: The foundations of adaptive behavior* (pp. 364–382). New York, NY: Oxford University Press.

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly, 17*, 1–25.

DeNisi, A. S., Cafferty, T. P., & Meglino, B. M. (1984). A cognitive view of the performance appraisal process: A model and research propositions. *Organizational Behavior and Human Performance, 33*, 360–396.

Dhami, M. K. (2003). Psychological models of professional decision making. *Psychological Science, 14*, 175–180.

Dhami, M. K., & Ayton, P. (2001). Bailing and jailing the fast-and-frugal way. *Journal of Behavioral Decision Making, 14*, 141–168.

Erev, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychological Review, 105*, 280–298.

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

Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes, 43*, 75–117.
Garcia-Retamero, R., & Dhami, M. K. (2009). Take-the-best in expert-novice decision strategies for residential burglary. *Psychonomic Bulletin and Review, 16*, 163–169.
Geisser, S. (1993). *Predictive inference: An introduction*. New York, NY: Chapman & Hall.
Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science, 3*, 20–29.
Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650–669.
Gigerenzer, G., & Selten, R. (Eds.). (2002). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT 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.
Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109*, 75–90.
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.
Hertwig, R., Davis, J. N., & Sulloway, F. J. (2002). Parental investment: How an equity motive can produce inequality. *Psychological Bulletin, 128*, 728–745.
Hertwig, R., Fischbacher, U., & Bruhin, A. (2013). Simple heuristics in a social game. In R. Hertwig, U. Hoffrage, & the ABC Research Group (Eds.), *Simple heuristics in a social world* (pp. 39–66). New York, NY: Oxford University Press.
Hodgkinson, G. P., & Healey, M. P. (2008). Cognition in organizations. *Annual Review of Psychology, 59*, 387–417.
Hough, L. M., & Oswald, F. L. (2000). Personnel selection: Looking toward the future—Remembering the past. *Annual Review of Psychology, 51*, 631–664.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.
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*, 1063–1073.
Klein, G. (2003). *The power of intuition*. New York, NY: Currency-Doubleday.
Kohli, R., & Jedidi, K. (2007). Representation and inference of lexicographic preference models and their variants. *Marketing Science, 26*, 380–399.
Landy, F. J., & Farr, J. (1980). Performance ratings. *Psychological Bulletin, 87*, 72–107.
Lawler, E. E. (2003). Reward practices and performance management system effectiveness. *Organizational Dynamics, 32*, 396–404.
Lee, H., & Dalal, R. S. (2011). The effects of performance extremities on ratings of dynamic performance. *Human Performance, 24*, 99–118.
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.
Lopes, L. L. (1995). Algebra and process in the modeling of risky choice. In J. R. Busemeyer, R. Hastie, & D. Medin (Eds.), *Decision making from the perspective of cognitive psychology* (pp. 177–220). New York, NY: Academic Press.
Luan, S., Schooler, L. J., & Gigerenzer, G. (2011). A signal detection analysis of fast-and-frugal trees. *Psychological Review, 118*, 316–338.
Luan, S., Schooler, L. J., & Gigerenzer, G. (2014). From perception to preference and onto inference: An approach-avoidance analysis of thresholds. *Psychological Review, 121*, 501–525.
Luce, R. D. (1986). *Response times*. New York, NY: Oxford University Press.
March, J. G., & Simon, H. A. (1958). *Organizations*. New York, NY: Wiley.
Martignon, L., Katsikopoulos, K. V., & Woike, J. K. (2008). Categorization with limited resources: A family of simple heuristics. *Journal of Mathematical Psychology, 52*, 352–361.
Mohr, L. B. (1982). *Explaining organizational behavior: The limits and possibilities of theory and research*. San Francisco, CA: Jossey-Bass.
Myung, I. J., Forster, M., & Browne, M. W. (Eds.). (2000). Special issue on model selection. *Journal of Mathematical Psychology, 44*.
Norman, S. M., Avolio, B. J., & Luthans, F. (2010). The impact of positivity and transparency on trust in leaders and their perceived effectiveness. *The Leadership Quarterly, 21*, 350–364.
Pachur, T., & Marinello, G. (2013). Expert intuitions: How to model the decision strategies of airport customs officers? *Acta Psychologica, 144*, 97–103.
Palanski, M. E., Kahai, S. S., & Yammarino, F. J. (2011). Team virtues and performance: An examination of transparency, behavioral integrity, and trust. *Journal of Business Ethics, 99*, 201–216.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, England: Cambridge University Press.
Reb, J., & Cropanzano, R. (2007). Evaluating dynamic performance: The influence of salient gestalt characteristics on performance ratings. *Journal of Applied Psychology, 92*, 490–499.
Reb, J., & Greguras, G. J. (2010). Understanding performance ratings: Dynamic performance, attributions, and rating purpose. *Journal of Applied Psychology, 95*, 213–220.
Russo, J. E., & Schoemaker, P. J. (2002). *Winning decisions: Getting it right the first time*. New York, NY: Crown Business.
Sackett, P. R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology, 59*, 419–450.
Savage, L. J. (1954). *The foundations of statistics*. New York, NY: Wiley.
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*, 262–274.
Scullen, S. E., Bergey, P. K., & Aiman-Smith, L. (2005). Forced distribution rating systems and the improvement of workforce potential: A baseline simulation. *Personnel Psychology, 58*, 1–32.
Shannon, C. E. (1948). A mathematical theory of communications. *Bell System Technical Journal, 27*, 623–656.
Shiffrin, R. M., Lee, M. D., Kim, W., & Wagenmakers, E.-J. (2008). A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science, 32*, 1248–1284.
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. *Quarterly Journal of Economics, 59*, 99–118.
Simon, H. A. (1957). *Administrative behavior: A study of decision-making processes in administrative organizations* (2nd ed.). New York, NY: Macmillan.
Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*, 1–20.
Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315.
Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B, 36*, 111–147.
Tan, J. H., Luan, S., & Katsikopoulos, K. V. (2017). A signal-detection approach to modeling forgiveness decisions. *Evolution and Human Behavior, 38*, 27–38.
Todd, P. M., & Gigerenzer, G. the ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. New York, NY: Oxford University Press.
Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79*, 281–299.