

**“Take-the-Best” and other simple strategies: Why and when they  
work “well” in binary choice**

Robin M. Hogarth & Natalia Karelaia\*  
Universitat Pompeu Fabra, Barcelona

[robin.hogarth@upf.edu](mailto:robin.hogarth@upf.edu)

[natalia.karelaia@upf.edu](mailto:natalia.karelaia@upf.edu)

October 2003

---

\* Robin M. Hogarth is ICREA Research Professor, and Natalia Karelaia is a graduate student at Universitat Pompeu Fabra, Barcelona, Spain. This research was financed partially by a grant from the Spanish Ministerio de Ciencia y Tecnología. We thank Ulrich Hoffrage and Laura Martignon for penetrating comments on an earlier version and Marc LeMenestrel for a most illuminating conversation on the topic of the paper.

For correspondence, please contact Robin M. Hogarth at Universitat Pompeu Fabra, Department of Economics and Business, Ramon Trias Fargas 25-27, 08005, Barcelona, Spain. (Tel: +34 93 542 2561; Fax: +34 93 542 1746).

## **Abstract**

The effectiveness of decision rules depends on characteristics of both rules and environments. A theoretical analysis of environments specifies the relative predictive accuracies of the lexicographic rule “take-the-best” (TTB) and other simple strategies for binary choice. We identify three factors: how the environment weights variables; characteristics of choice sets; and error. For cases involving from three to five binary cues, TTB is effective across many environments. However, hybrids of equal weights (EW) and TTB models are more effective as environments become more compensatory. In the presence of error, TTB and similar models do not predict much better than a naïve model that exploits dominance. We emphasize psychological implications and the need for more complete theories of the environment that include the role of error.

**Keywords:** Decision making, Bounded rationality, Lexicographic rules

**JEL classification:** D81, M10.

**“Take-the-Best” and other simple strategies: Why and when they  
work “well” in binary choice**

Imagine that you are facing a binary choice. You must decide which of two alternatives, A or B, is “better” in the sense of having more of a specific criterion. Examples could include choosing between two job candidates, two stocks, two restaurants, which route to take on a trip, and so on. In other words, this is a familiar judgment and decision making task. Imagine further that the information on which you can make your judgment is limited to several ( $k$ ) binary cues, i.e., cues that indicate presence or absence of an attribute relevant to the task. Thus, if the number of cues ( $k$ ) is, say, three, option A can be characterized by the vector or cue profile

$$A = \{x_{a1}, x_{a2}, x_{a3}\} \quad (1)$$

where the  $x_{aj}$  can only take the values of 0 or 1 ( $j = 1, \dots, 3$ ).

Similarly, option B can be characterized by the vector or cue profile

$$B = \{x_{b1}, x_{b2}, x_{b3}\} \quad (2)$$

where the  $x_{bj}$  can only take the values of 0 or 1 ( $j = 1, \dots, 3$ ).

Finally, assume that the environment determines the correct answer, A or B, by weighting the cues according to some function. In what follows, we assume an additive function<sup>1</sup> where the sum of all the weights is equal to 1, i.e., let  $\beta_j$  be the

weight given to attribute  $j$  ( $j = 1, \dots, k$ ) such that  $\sum_{j=1}^k \beta_j = 1$ .

In many studies, simple lexicographic rules have demonstrated remarkably accurate performance in binary choice when compared to statistical benchmarks (Gigerenzer & Goldstein, 1996; Gigerenzer, Todd & the ABC Group, 1999;

---

<sup>1</sup> As noted by Dawes and Corrigan (1974), many nonlinear functions can be well approximated by linear functions and particularly when the former are conditionally monotonic with respect to the criterion variable. The presence of error or “noise” also makes linear approximations more “optimal.”

Martignon & Hoffrage, 2002). Of particular interest is the rule known as “take-the-best” (henceforth TTB) which works as follows. First, the model assumes knowledge of the differential ability of the cues to predict the criterion, i.e., the cue validities (in this case, assume that the order is  $x_{.1}$ ,  $x_{.2}$ ,  $x_{.3}$ ). Second, choice between A and B is made if the first cue ( $x_{.1}$ ) can discriminate between the options; if the first cue cannot, the second cue is used to make the choice; and so on. Finally, if none of the cues can discriminate, choice is made at random. From a cognitive viewpoint, this rule can be easily implemented. It does not require many mental operations nor, in many cases, examining much information (often just the first cue). However, it does require ordering the cues by their validities. (For further details, see Gigerenzer & Goldstein, 1996; Gigerenzer, Todd et al., 1999.)

TTB has been presented as an example of a “fast and frugal” heuristic, an element of the “adaptive toolbox” of bounded rationality (Gigerenzer & Goldstein, 1996; Gigerenzer & Selten, 2001). In addition to demonstrations of its predictive ability, several experimental studies have addressed whether and when people actually use TTB-like mental strategies (see, e.g., Hoffrage & Rieskamp, 1999; 2002; Bröder, 2000; Bröder & Schiffer, 2003; Newell & Shanks, 2003; Newell, Weston, & Shanks, 2003). Overall, there is evidence that people do use TTB-like processes but, not to the exclusion of other strategies.

In this paper, we emphasize that the performance of response strategies or decision rules depends on characteristics of *both* the rules and of the environments in which they operate (Brunswik, 1952; Simon, 1956). A complete theory of psychological functioning needs to specify both. However, whereas investigators have had little difficulty in specifying rules, the specification of task environments has proven more problematic. Our goal is to illuminate this issue and our approach is

theoretical. It involves specifying abstract characterizations of tasks and noting how different models would be expected to perform in these environments.

In conceptualizing environments for binary choice, we emphasize three dimensions. One is the type of function used by the environment to decide which alternative is correct. The second is the type of distribution of cue profiles in the choice set. The third is the role of error. This can be located in the application of the model, the environment, or both.

The paper is organized as follows. We first define the different models we consider. Second, we examine their theoretical performance under both non-compensatory and compensatory weighting functions in environments characterized by lack of error. This is done separately for models involving three, four and five cues.<sup>2</sup> Third, we investigate one aspect of error in models: namely, failure to apply the model appropriately, i.e., to respect the ecological ordering of cues in TTB. Fourth, we consider how different distributions of cue profiles affect the relative performance of simple models and illustrate this using the 20 datasets of Czerlinski, Gigerenzer, and Goldstein (1999). Fifth, the importance of error the environment may contain is highlighted by lack of agreement between theoretical predictions of model performance based on characteristics of distributions and actual empirical results. We therefore use simulation to examine the role of error. Finally, we discuss our results from both psychological and prescriptive perspectives.

In brief, we show that at a theoretical level TTB does work “well” as a model of binary choice. But to understand how “well” requires specifying appropriate benchmarks. The normative standard involves models such as Bayesian networks, multiple regression or exemplar-based approaches (cf., Chater, Oaksford, Nakisa, &

---

<sup>2</sup> We limit our analysis to three, four, and five cues for two reasons. One is to reduce analytical complexity. The second is that three, four, and five cues seem sufficient to understand what people can actually do within limited information processing constraints.

Redington, 2003). Whereas such comparisons are interesting, we do not believe they are the most illuminating. There are two reasons.

First, the advantage of simple models in the tradition of the “adaptive toolbox” (Gigerenzer & Selten, 2001) is that they can be used in many situations where people lack the experience necessary to develop more sophisticated processes.<sup>3</sup> Second, when dealing with small samples, it is well-known that regression analysis (and other optimizing techniques) produces parameter estimates with large standard deviations such that predictions to further samples are subject to much error. In these cases, regression analysis and similar tools become “straw men” that lack meaning (see, e.g., Einhorn & Hogarth, 1975).

Instead, we compare TTB to other simple models, some of which incorporate features of TTB. For example, we explore one model that combines features of both TTB and equal weighting – called EW/TTB. Moreover, we show that this model improves the predictive ability of TTB in certain environments and yet, when the number of variables is small, does not require much additional information processing.

To establish a yardstick for simple models, we propose that all reasonable models of binary choice should exploit dominance. This leads to the following benchmark. Choose according to dominance. If there is no dominance, choose at random. As we show, this strategy – that we call DOMRAN – actually predicts quite well in the kinds of environments studied by Gigerenzer and his colleagues. In particular, when data are “noisy” its performance does not fall far behind that of TTB.

---

<sup>3</sup> As argued by Chater et al. (2003), it is clear that many basic physical and psychological processes can be well modeled by what most would classify as complex normative models. However, note that most of these processes (as an example, consider perception) have evolved over many years of evolution thereby implying much data in their “development.” Paradoxically, many “higher order” mental processes are used in situations involving scarce data which effectively preclude using complex normative models (see also Todd & Gigerenzer, 2000).

Thus in interpreting the performance of TTB and similar models, it is important to investigate how they predict in cases that cannot be decided by dominance. What is the marginal predictive significance compared to the standard set by DOMRAN?

### **The different models**

In comparing the different models that we investigate, it is of interest to note: (1) what knowledge they require about the variables (i.e., the cues) such as relative importance and, if so, how accurate this needs to be; (2) how many cues must be examined to make a decision; (3) whether explicit calculations are required; (4) the number of comparisons to be made; and (5) whether random choice is used to break ties. Table 1 provides an overview of such characteristics. As will be seen, some of the models are combinations of different models.

*DOMRAN.* This exploits dominance. If one alternative dominates another, it is chosen; if not, choice is made at random. The psychological inspiration is provided by the work of Montgomery (1983) who has documented how people seek to find and exploit dominance and may even distort information so that dominance can be “justified.” In general, we suspect that screening for dominance occurs frequently and thus this model provides a useful lower bound in terms of a “reasonable” simple strategy.

-----  
Insert Table 1 about here  
-----

*EW.* In the equal weighting or “tallying” model (Gigerenzer, Todd, et al., 1999), each variable is given the same weight and the alternative chosen has the larger weighted sum. In practice, since variables take the values of 0 or 1, this is equivalent to summing the variables of each alternative and choosing the larger sum. When the

sums are equal, choice is made at random. This model involves examining all data, making two sums, and one comparison. It has proven useful in predicting many phenomena when people do not know the relative weights to give to variables (Dawes & Corrigan, 1974; Dawes, 1979; Einhorn & Hogarth, 1975).

*TTB*. See description above.

*EW/TTB hybrids*. As will be seen below, in many cases EW results in ties. The hybrid models operate in two phases. In the first, EW is used – on all or a subset of the variables (to be specified). If EW favors one alternative, it is chosen. Otherwise, choice is made by TTB.

It is important to note that in all of the models, the signs of the correlations between cues and criterion are assumed to be known, i.e., the variables are scaled such that a cue value of “1” implies a greater value on the criterion than a cue value of “0.” In addition, we do not deal with cases involving missing values of cues.

Finally, although all these models appear to be different, we shall show below (see Discussion) that, when decisions are made by examining sequentially differences between cue values of the alternatives, all of these models can be thought of as belonging to the same general psychological process. The critical distinctions between the models lie in the rules used to stop the decision process.

### **Non-compensatory and compensatory functions**

We first characterize environments by the types of functions used to classify choices. In doing this, we follow the lead of Martignon and Hoffrage (1999; 2002) who have distinguished between non-compensatory and compensatory functions. Specifically, Martignon and Hoffrage define by non-compensatory any weighting scheme or function that has the property that, when weights are ordered from largest

to smallest, each weight is larger than the sum of all weights that are smaller than it, i.e.,

$$\beta_j > \sum_i \beta_i, \text{ for any } i > j, j = 1, \dots, k - 1. \quad (3)$$

Martignon and Hoffrage define all other functions as *compensatory*. Thus, for three cues,  $\beta_1 > (\beta_2 + \beta_3)$  in the non-compensatory case whereas  $\beta_1 \leq (\beta_2 + \beta_3)$  in the compensatory case (assuming that  $\beta_1 > \beta_2 > \beta_3$ ). An important theoretical result proven by Martignon and Hoffrage (1999; 2002) is that TTB is the optimal model for choice when weighting functions comply with their definition of *non-compensatory* environments.<sup>4</sup> Thus, the bulk of our attention will be focused on what happens when weighting functions are compensatory. Moreover, we provide separate analyses for cases involving three, four, and five cues.

### Different cue environments

#### *The 3-cue environment*

In empirical tests of TTB and other models conducted by the ABC Group, the basic task involves seeing how models predict between all possible pairs of a set of choice alternatives. Thus, given  $n$  alternatives, each characterized by binary vectors of length  $k$ , predictions are made for the  $n(n-1)/2$  possible pairs of alternatives. Thus, with 30 alternatives there are 435 pairs to predict, with 40 alternatives, 780 pairs, and so on. However, even though there may be many pairs to predict, it should be clear that the distinct cue profiles that characterize alternatives are limited by the number of  $k$  binary cues. More specifically, the number of distinct cue profiles is  $2^k$  such that with three cues there are eight distinct profiles, with four cues, 16 profiles, with five

---

<sup>4</sup> It is important to emphasize that this is a *theoretical* result, i.e., the  $\beta_j$ 's are environmental parameters.

cues, 32 profiles, and so on. This means that, in large samples of alternatives, many predictions must involve cases involving identical cue profiles (so-called “repeats”) and that the distribution of cue profiles among the alternatives affects results. This, as we shall show below, is an important insight.

To illustrate the effects of different cue profiles, consider the case of three cues and the eight different profiles that can result from these cues. These are shown in the top left section of Table 2 where the distinct cue profiles are given the labels A, B, C, D, E, F, G, and H. (Profile A is (1, 1, 1); profile B is (1, 1, 0); and so on.) Furthermore, assume that the variables have been ordered in importance, that is  $\beta_1 > \beta_2 > \beta_3$ .

-----  
 Insert Table 2 about here  
 -----

Now consider the vectors of the arithmetical differences between the two vectors representing the attributes of the alternatives. In the case of A and B, above, this is

$$A - B = \{x_{a1}-x_{b1}, x_{a2}-x_{b2}, x_{a3}-x_{b3}\} \tag{4}$$

where each element of the vector can take values of 1, 0, or -1 depending on the characteristics of the alternatives. These are shown to the right of the eight distinct profiles. Thus, the first set of difference vectors under the letter A shows the differences between A and the seven other profiles (B to H); those under B the differences between B and its successors (C to H); those under C the differences between C and its successors (D to H); and so on.

The difference vectors provide a simple way of assessing the predictions of different models for each combination of cue profile types. First, if all the elements of a difference vector are non-negative and at least one is positive, the first cue profile

dominates the second. Thus, as can be seen, cue profile A dominates all the other profiles – B through H. Similarly, B dominates D but not E, and so on. Cases involving dominance are indicated by a “d” in the matrix on the right hand side of Table 2.

Second, consider cases where the difference vectors contain negative elements such as the B-C pairing that takes the values 0, 1, -1 and restrict attention to weighting functions that are strictly compensatory, i.e., where  $\beta_1 > \beta_2 > \beta_3$ . Indeed, in this paper, we only consider strict inequalities.<sup>5</sup> TTB chooses B over C based on the first difference that appears here between values of the second variable. Moreover, this prediction is consistent with any model where  $\beta_2 > \beta_3$ . Thus, TTB also predicts this case correctly when the weighting function is strictly compensatory. Continuing to examine all cases that do not involve dominance, we use the fact that  $\beta_1 > \beta_2 > \beta_3$  to determine consistency between the choices of TTB and any strictly compensatory weighting function. These consistent cases are marked by a “c” in the appropriate places on the right hand side of Table 2.

Third, by the same logic, it is clear that TTB does not predict the D-E pairing correctly in the compensatory case where  $\beta_1 < (\beta_2 + \beta_3)$ . This is indicated by marking a “w” in the appropriate cell on the right of Table 2.

Fourth, we also indicate the letter “t” in cases where the EW model predicts ties, as in B-E and C-E.

Finally, Table 2 summarizes TTB’s predictions between all cue profile types for strictly compensatory weighting functions. There are 19 cases involving dominance. Thus, all models considered in this paper would make the same

---

<sup>5</sup> The effect of this may be to underestimate marginally the performance of TTB (see Appendix).

predictions for these cases. Of the remaining nine cases, TTB predicts eight correctly and makes one error.

The left hand side of Table 3 shows the accuracy of the various models for three cues when the population of cue profiles consists of just one of each type (i.e., A through H), for both non-compensatory and compensatory weighting functions. First, TTB is 100% accurate for non-compensatory functions – as proven by Martignon and Hoffrage (1999; 2002). (To simplify reading of tables, we adopt the practice of highlighting the largest figures in relevant comparisons in **bold**.)

-----  
Insert Table 3 about here  
-----

Second, TTB makes one error for compensatory functions where it achieves an overall accuracy rate of 96%.

Third, EW makes one error for non-compensatory functions and no errors for compensatory functions. However, as can be seen from Table 2, there are six cases where EW predicts ties for compensatory functions where decisions have to be made at random – hence its expected predictive accuracy of 89%. Thus the trade-off between the performances of TTB and EW is one sure error versus six cases that are decided by chance.

Fourth, EW/TTB makes one error with non-compensatory functions and no errors with compensatory functions. More specifically, it makes the same error as EW with the non-compensatory function but all ties are correctly resolved by the TTB mechanism. For the compensatory functions, EW/TTB has perfect performance because, first, it correctly predicts the D-E case for which TTB makes an error and, second, all of the EW ties are again correctly resolved by the TTB mechanism.

To summarize, *in the 3-cue case TTB is optimal for non-compensatory functions and EW/TTB is optimal for compensatory functions*. Moreover, these optimality statements can be made about TTB and EW/TTB irrespective of distributions of cue profiles precisely because they never make mistakes.<sup>6</sup>

It is important to emphasize this result from a psychological viewpoint. In many situations, people will not know whether they are being confronted by non-compensatory or compensatory functions. Thus, until they learn the characteristics of the environments they face, one would expect to see the use of different decision rules. In 3-cue environments, EW/TTB is not a difficult strategy to execute.

#### *The 4-cue environment*

The operational definitions of compensatory and non-compensatory functions are quite straightforward in the 3-cue case. However, defining compensatory functions by violations of the condition for non-compensatory functions leads to several distinct classes of the former in the 4-cue case. Specifically, if – for four cues – we define non-compensatory by the conditions that, first,  $\beta_j > \sum_i \beta_i$ , for any  $i > j$ ,  $j = 1, \dots, k - 1$ , and second, that  $\beta_1 > \beta_2 > \beta_3 > \beta_4$ , there are several compensatory functions that violate the first condition to different extents.

For instance, if we specify that  $\beta_1 < \beta_2 + \beta_3$  (which in turn implies that  $\beta_1 < \beta_2 + \beta_3 + \beta_4$ ), this can be accompanied by either  $\beta_2 < \beta_3 + \beta_4$  or  $\beta_2 > \beta_3 + \beta_4$ . In fact, as shown in the Appendix, there are five different classes of weighting functions that span the parameter space of compensatory environments – see Figure 1. As in the 3-cue case, we only consider strict inequalities in dividing up the parameter space.

---

<sup>6</sup> The expected predicted performances of these optimal models will not necessarily always be 100%. This can occur when the distribution of cue profiles contains one or more “repeats” of the same profile. We consider this issue in more detail below.

To illustrate differences between the weighting functions, Table 4 provides numerical examples. As can be seen, CF1, CF2, and CF3 are close to “non-compensatory” and CF5 can accommodate distributions of weights that vary from the first being much larger than the others to a set of almost equal weights.

-----  
Insert Table 4 and Figure 1 about here  
-----

With four cues, there are 120 distinctive profile pairings. The right hand side of Table 3 characterizes the performance of the different models. Compared to the 3-cue case, we have an additional hybrid model labeled EW-3/TTB. This is a modification of EW/TTB that works as follows. In the first stage, the decision maker uses an equal weighting model on the *three* most important variables (i.e., omitting the fourth). If this points to a decision, it is taken. If there is a tie, it is resolved by TTB.<sup>7</sup>

In this 4-cue world, 54% of the distinctive pairings involve dominance such that the expected performance of DOMRAN is 77%, i.e.,  $54\% + 0.5(100\% - 54\%)$ . TTB makes, of course, no errors in the non-compensatory case and is unique in this respect. CF4 and CF5 are unable to provide unambiguous choices for three and five cue profile pairings respectively (see Appendix, Table A1). Operationally, these cases have been treated as ties which all models are assumed to predict correctly with probability of 0.5.

Overall, with populations of unique cue profile pairings, the pattern of results for the 4-cue case matches that of three cues. For non-compensatory (CF1) and close to non-compensatory functions (CF1, CF2, CF3), TTB makes the least numbers of errors. As the functions become more compensatory (CF4, CF5), it is the EW/TTB

---

<sup>7</sup> Why, the reader may ask, do we not also define a EW-2/TTB model? The reason is that this latter model makes predictions that are identical to those of TTB.

models that perform relatively better. In particular, under CF5, EW makes no errors such that the TTB contribution to EW/TTB is the correct allocation of all EW ties.

*The 5-cue environment*

The 5-cue environment exhibits the same general trends. However, it is more complex. There are 496 distinctive profile pairings; 23 different types of compensatory functions (see Appendix, Figure A1); and many more cases where functions imply ambiguous predictions (see Appendix, Table A2). Once again, we draw attention to the numerical examples of parameters in Table 4. Functions CF1 through CF17 are close to “non-compensatory” and, even with CF23 it is possible to have the weight of the first variable much larger than the others.

-----  
Insert Table 5 about here  
-----

Table 5 presents the expected predictive accuracies of the different models for the 5-cue case for non-compensatory functions and 12 of the 23 compensatory functions. The models are the same as in the 4-cue case except that the EW/TTB hybrids include a version based on the first four most important cues. At the foot of Table 5 we also indicate the number of ambiguous choices for each set of weighting functions. These become much larger as the parameters indicate more compensatory environments.

TTB is 100% correct with the non-compensatory function (as must be the case) but its performance drops off in relative terms as the functions become more compensatory. It makes the smallest number of errors, as defined above, through CF16. EW-3/TTB has the best performance for CF18, CF19, and CF20, and the EW/TTB hybrids perform relatively well for the most compensatory functions: see EW-4/TTB for CF22, and EW/TTB for CF23.

### *Summary*

As shown by Martignon and Hoffrage (1999; 2002), TTB is optimal when environments are non-compensatory (by their definition). In addition, for the 3-, 4-, and 5-cue cases, TTB is one of the best strategies when environments consist of unique cue profile pairings. Moreover, even in fairly compensatory environments, TTB does well. However, as the environments become more compensatory, hybrid strategies such as EW/TTB become more effective in a relative sense. In these strategies, TTB intervenes when EW predicts ties. The EW/TTB hybrid is 100% accurate with 3-cues and the strategy of preference for the most compensatory strategies in the 4- and 5-cue cases, i.e., for CF5 in the 4-cue case, and CF23 in the 5-cue case.

From a psychological perspective, TTB is useful in both non-compensatory and compensatory environments. In the former, it is all that is needed for choice. In the latter, it acts as a tie-breaker for EW or “almost EW” models (e.g., when EW is calculated on a subset of the variables). From an empirical viewpoint, therefore, to the extent that people’s decision making is well adapted to their environments, one should expect to see the use of TTB in conjunction with other models.

Finally, we note that although the DOMRAN strategy has the lowest expected performance in all cases, in absolute terms its expected performance is quite high, i.e., 84% in the 3-cue case, 77% in the 4 cue-case, and 71% in the 5-cue case. As we shall demonstrate below, the fact that DOMRAN provides such a high “lower benchmark” is important for understanding the relative success of simple models for binary choice.

## Error in the application of models

All the simple models we investigate assume correct knowledge and use of the signs of the zero-order correlations between cues and criterion. In addition, TTB is assumed to know and use the relative sizes of the  $\beta$ -parameters associated with the cues.<sup>8</sup> What happens, therefore, when the cues used in the TTB process are not considered in the appropriate order, i.e., there is error in knowledge and/or application of the TTB model?

We begin by examining the 3-cue case since this is simpler than the 4- and 5-cue cases and suggests the direction of more complicated results.

With 3 cues, there are 3! (= 6) possible orderings of the cues. These are shown in Table 6 together with results of different models. (Once again, characters in **bold** indicate the best expected correct percentages within classes of parameters, non-compensatory and compensatory.) In the non-compensatory case, TTB remains the best strategy but only provided the most important variable is correctly identified as such. For all other orderings, EW/TTB has the best expected correct predictions. In the compensatory case, EW/TTB is better than the other strategies no matter the order in which variables enter the models. (However, note comments about EW below.)

At the foot of Table 6, we have also indicated the means of the different columns as well as the expected performance of models that are not affected by the order in which cues are examined. As noted previously, DOMRAN has expected performance of 84%. Indeed, this outperforms the last three cue orderings of all models with the exception of EW/TTB. EW achieves 86% and 89% for the non-compensatory and compensatory weighting functions, respectively. The means of the TTB columns equal the performance of what Gigerenzer, Todd et al. (1999) refer to

---

<sup>8</sup> Recall, however, that it does not require precise knowledge of the sizes of the  $\beta$ -parameters and, in this sense, the prior knowledge requirements are not necessarily that onerous.

as the MINIMALIST strategy. This is the performance that would be expected of a TTB model where the order of the variables entering the model is decided at random.<sup>9</sup> However, MINIMALIST fails to reach the expected performance level of EW which actually matches the mean of EW/TTB. Thus EW matches or exceeds EW/TTB in roughly half of the possible orderings (i.e., the lower orderings).

Parenthetically, one way to interpret the expected performance level of MINIMALIST is to consider situations where people cannot control the order in which cues are examined or, indeed, which cues will eventually become available. In these situations, decisions are made by using cues in the order in which they are accessed. Thus, if the environment essentially randomizes the ordering of cues, a TTB-like strategy will have the expected performance of MINIMALIST. However, if the more (less) important cues happen to be accessed first, the strategy will be more (less) effective than MINIMALIST.

-----  
Insert Tables 6 and 7 about here  
-----

As shown in Table 7, in the 4-cue case there are 24 different possible orderings of the variables. For clarity, we only present the results of three of the six possible weighting functions – NonCF, CF2, and CF5. Once again, we emphasize the best predictions within an order in **bold** characters. Overall, results mirror the 3-cue case. When the functions are non-compensatory or least compensatory, TTB performs best provided the most important variable enters the model first. Otherwise, EW/TTB performs best and is best across the range of orderings as the parameters become more compensatory (see CF5). DOMRAN achieves 77% in this population

---

<sup>9</sup> In the MINIMALIST strategy, variables enter the model in a random order for each binary choice. The assumption being made here is that the expected performance level of MINIMALIST is equal to that of the mean of TTB across all possible orderings, i.e., the expected performance of MINIMALIST is equivalent to that obtained by sampling different cue orderings at random.

and this is clearly a better score than achieved by different models that fail to identify the appropriate ordering of the variables. Interestingly, DOMRAN only exceeds EW/TTB in the most compensatory case (CF5) in the last (and most incorrect) ordering (77% vs. 76%). Finally, EW matches the mean of EW/TTB across orderings.

Table 8 presents the analogous results for 5-cue models. Given that there are 120 different ways in which the cues can enter models and 24 different weighting functions, we neither show the results of all cue orderings nor of all weighting functions. Instead, we illustrate the trends by showing significant subsets of combinations of functions and cue orderings. Overall, these are similar to the results of the 3- and 4-cue cases. First, for non-compensatory functions (NonCF) as well as lower levels of compensatory functions (CF1 to CF17), TTB performs best provided the most important cue enters the model first. When this does not occur, EW/TTB performs better. Second, EW/TTB is dominant across all cue orderings for the most compensatory set of weighting functions (CF23). Third, DOMRAN is superior to many of the combinations of cue orders and models where the cue orderings are inappropriate (with the exception of EW/TTB). And fourth, as before, MINIMALIST (the mean of TTB across cue orderings) is inferior to EW which is equal to the mean of EW/TTB.

-----  
Insert Table 8 about here  
-----

Overall, differences between the models across the different combinations of weighting functions are small. Moreover, provided the first two most important variables enter the models in the appropriate order, TTB and EW/TTB do quite well in an absolute sense, i.e., approximate expected success rates of 90% and above. However, the major result for populations of distinctive cue pairings is that TTB is

best provided the most important cue or variable does enter the model first *and* the weighting functions are not the most compensatory. Otherwise, EW/TTB should be preferred.

It is perhaps surprising that all models seem to have quite high expected correct predictions even when cue orderings are inappropriate. However, once again the DOMRAN model provides a good naïve benchmark with which to calibrate this impression. It is consistently superior to many results achieved with incorrect cue orderings.

### **Different distributions of cue profiles**

The above results are conditioned on populations consisting of unique cue profiles. However, characteristics of choice sets are an important dimension of environmental variability. In particular, we would expect both the general level of predictive ability as well as relative differences between models to depend on characteristics of the distributions of cue profiles in given populations. We consider three main factors that we illustrate by the three distributions shown in Table 9.

-----  
Insert Table 9 about here  
-----

First, distributions can differ in the number of dominating cue profiles. In general, the greater the proportion of dominating cue profiles, the greater is the expected performance of all models. In Table 9, Distribution III has a lower proportion of dominating profiles than the other distributions. Hence, DOMRAN (as well as the other models) performs less well here than in the other distributions.

Second, when there are repeats of the same cue profile, all models would only be expected to discriminate correctly between such cases at a rate of 50%.<sup>10</sup> Thus, the general level of predictability between two populations depends, in part, on the number of repeated profiles in each. Specifically, repeats lower overall expected performance. In Table 9, Distribution III contains several repeats.

Third, the conflict implicit in the difference vectors has more impact on the relative success of some models than of others. For instance, TTB always makes mistakes for the D-E choice in the 3-cue case with the compensatory weighting function (see Table 2), but EW – and thus EW/TTB – does not. Hence, the presence or absence of D-E choices in a population can affect the relative success of these decision rules. As a case in point, D-E conflict is present in Distributions I and III but absent from Distribution II. Note, in particular, that TTB has predicted performance of 100% correct in Distribution II but 96% and 80% in Distributions I and III, respectively.

More generally, distributions or “choice environments” can be described as being “TTB-friendly” or “TTB-unfriendly” for compensatory functions depending on the absence or presence of cue profile pairings that TTB classifies incorrectly. Thus, Distribution II in Table 9 can be described as TTB-friendly (there are no D-E pairings) whereas Distribution III is TTB-unfriendly.

Whether a distribution is TTB-friendly or TTB-unfriendly can be characterized by asking how it varies from a uniform distribution (e.g., Distribution I in Table 9) in terms of the number of errors made by TTB. Specifically, we describe choice environments with uniform distributions of cue profiles as “TTB-neutral.”

---

<sup>10</sup> In a world without error, repeat profiles must have identical values on the dependent variable. Thus, it could be argued, it does not matter which profile is selected because each must be “correct.” In this work, however, we take a more conservative approach and assume that one of the two alternatives is indeed correct. Thus, models can only discriminate the correct alternative by chance, i.e., with probability of 0.50.

Thus, if the expected number of TTB errors in a distribution is less (more) than that expected on the basis of a uniform distribution, it will be described as TTB-friendly (TTB-unfriendly). As an example, consider Distribution III in Table 9. This has 16 observations such that a uniform “equivalent” would have two observations of each cue profile type. This uniform distribution would have two D observations and two E observations and, consequently, make four TTB errors (i.e., there are  $2 \times 2$  D-E pairings). In Distribution III, note that TTB makes 12 errors (i.e., there are  $2 \times 6$  D-E pairings). Because 12 is greater than four, we describe Distribution III as TTB-unfriendly. Distribution II, on the other hand, is TTB-friendly because  $0 < 1$ . In short, when the number of predicted TTB errors is smaller (greater) than expected on the basis of a uniform distribution, we describe the distribution as TTB-friendly (TTB-unfriendly).

Parenthetically, we note that the predictive success of lexicographic models such as TTB has sometimes been attributed to correlation between the cues. However, this is not a complete explanation. As indicated above, TTB is quite successful in TTB-neutral environments in which the intercorrelations between cues are zero, i.e., uniform distributions of distinctive cue profiles (see Tables 3 and 5 through 8). What is critical to the performance of TTB is the presence or absence of the specific cue pairings that it predicts incorrectly, i.e., whether the distributions are TTB-unfriendly or TTB-friendly.

### *Some empirical distributions*

The data in Table 9 were constructed for illustrative purposes. What can be said about “real” data? To examine the characteristics of different distributions, we

use 20 datasets created by Czerlinski et. al (1999).<sup>11</sup> First, we ignore the empirical criterion variables and examine the characteristics of the datasets by cue profiles. What proportions of the choices in each dataset involve dominance and repeats? To what extent are these datasets TTB-friendly or TTB-unfriendly?

Second, we use the actual distributions of cue profiles to predict the performance of TTB and other simple models assuming both non-compensatory and compensatory weighting functions, i.e., given the distributions of cue profiles, and assuming no error, we calculate the expected predictive performance of the models.

Third, we contrast these theoretical predictions with actual predictions by the models of the criteria in the 20 datasets.

We emphasize that the theoretical predictions are made assuming no error – either in the environment or in the models (i.e., knowledge and application). In the subsequent section, we consider effects of error.

Table 10 reports characteristics of the 20 datasets that we have split into three groups according to numbers of cues (three, four, and five). The 5-cue set actually includes datasets that had more than five cues. However, in each of these we have only considered the five most important cues (determined by examining cue validities across all data). First, in addition to numbers of observations, we report the number of cases where all models make identical choices, i.e., for cases involving dominance and repeated cue profiles (in the second and third columns of the table). The total number of common choices (i.e., the sum of dominant pairs and repeats) is large, varying from 39% to 96% with an average of 67%. In other words, the predictions of models can only be distinguished on about one-third of these data. We shall return to this point below.

---

<sup>11</sup> We are very grateful to these researchers for making their data available on their website. See <http://www-abc.mpib-berlin.mpg.de/sim/Heuristica/environments/>

Second, Table 10 illustrates the extent to which the datasets are TTB-friendly or TTB-unfriendly. That is, for each dataset we calculate the ratio of the number of TTB errors that would be expected in a uniform distribution of cue profiles (of size equal to the actual distribution) with the theoretical number of TTB errors implied by the actual distribution.

We begin by considering the 3-cue sets. The four 3-cue datasets are all TTB-friendly. Indeed, for three of the distributions, the ratios of expected errors are infinite because there are no expected TTB errors in the actual distributions. For these distributions, therefore, TTB is expected to perform as well as EW/TTB for compensatory functions. This is shown in the upper part of Table 11 that details predictions of the models for both non-compensatory and compensatory functions.

-----  
Insert Tables 10 and 11 about here  
-----

There are two 4-cue distributions. However, to assess whether a distribution is TTB-friendly or TTB-unfriendly, specific compensatory weighting function must be used. Here we use CF2 and CF5. As can be seen, one distribution (“Oxidant”) is TTB-friendly, whereas the other (“Land rent”) is not. The effect of this can be seen in the theoretical predictions in Table 11. Predictions for “Oxidant” for TTB are uniformly high across all weighting functions. For “Land rent,” on the other hand, the TTB prediction under the most compensatory weighting scheme (CF5) is much lower than under the other schemes (79% vs. 96% and 92%).

The 5-cue datasets have a mix of TTB-friendly, TTB-unfriendly, and TTB-neutral distributions. Once again, classification of TTB-friendly or TTB-unfriendly depends on specifying particular weighting functions. In this case, we illustrate CF9

and CF23. In terms of theoretical predictions (Table 11), there is a tendency for TTB predictions to be lower under CF23 than CF9 with the reverse occurring for EW/TTB.

Finally, note that the averages of the DOMRAN predictions are fairly high thereby suggesting that strategies that exploit dominance should be quite predictive. At a process level, this is important because TTB makes identical predictions to what can be called a DOM/TTB strategy, i.e., use dominance to choose; if this fails, use TTB.

How do the theoretical predictions in Table 11 compare with predictions for the actual data? Figure 2 reports mean predictive accuracies of the models across all 20 datasets on holdout samples using 1,000 replications. Specifically, for each dataset we randomly sampled a proportion of the possible choices, fit parameters as appropriate (e.g., calculating cue validities in TTB), and then used these parameters to predict the remaining choices in the dataset (i.e., the holdout sample).<sup>12</sup> We replicated this process 1,000 times and used different proportions of fitting and holdout samples – a 50/50 split and a 20/80 split.

-----  
Insert Figure 2 about here  
-----

Figure 2 reveals four major trends. First, the differences between TTB, EW/TTB and EW are small (this is also true of the results of all the datasets that have been averaged). Second, the accuracy levels of the empirical predictions of all models except DOMRAN (and to a lesser extent EW) are some 20% below that of the theoretical predictions. Third, and as might be expected, predictability is somewhat greater in the 50/50 split than in the 20/80 split.

---

<sup>12</sup> Once again, we limited the number of cues in any dataset to five.

Fourth, DOMRAN is the least successful of the models. However, the difference between DOMRAN and the other models is small (at most between 4% and 5% in predictive accuracy). Indeed, compared to the other models, the differences between the theoretical DOMRAN predictions in Table 11 (average of 77%) and those actually realized (68% for the 50/50 split – Figure 2) are much smaller. Perhaps the surprising story of these data is not that TTB is the best of the simple models (an important finding), but that the naïve DOMRAN benchmark does so well.

Finally, the important role of error in the empirical datasets is highlighted by the fact that, across the 20 datasets, there is no correlation between the theoretical DOMRAN predictions (Table 11) and the cross-validated predictive accuracies of DOMRAN in the 50/50 split ( $r = 0.043$ , ns).

### **Understanding the role of error**

As noted above, error can be thought of as being located in models or in the environment. Errors in models can result from lack of knowledge (e.g., not knowing the correct order in which to consult variables in TTB) or execution (e.g., people might have appropriate knowledge but make errors in using models) or both. In addition, error in the environment can affect error in models. For example, it is probably more difficult for people to learn the relative sizes of the  $\beta$ -parameters in noisy as opposed to error-free environments.

Imagine the kinds of error that might occur within the linear model environments considered in this paper. First, consider the model of the error-free environment

$$y_i = \sum_{j=1}^k \beta_j x_{ij} \quad j = 1, \dots, k \quad (5)$$

where  $y_i$  is the dependent variable, the  $\beta_j$ 's are the weighting function parameters, and the  $x_{ij}$  can take the values of 0 or 1. Within this framework, one can conceive of at least three types of error: (1) noise that affects the distribution of  $y_i$  conditional on values of  $\sum_{j=1}^k \beta_j x_{ij}$ ; (2) noise that affects the weighting parameters, e.g., the  $\beta_j$ 's are not identical for all observations; and (3) errors in the independent variables, e.g., such that the  $x_{ij}$  cannot always be perceived accurately. Finally, these errors may not be well behaved in the sense of having, say, constant variances.

It is beyond the scope of this paper to investigate all these sources of error. We therefore consider only the first which we specify as follows

$$y_i = \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i \quad (6)$$

where  $\varepsilon_i$  is a normally distributed random error term  $(0, \sigma^2)$ . However, as suggested above, the use of TTB in the “real world” can induce error in knowledge in that, because the  $\beta_j$  parameters are unknown, their relative sizes can only be estimated on the basis of samples. The accuracy of such knowledge therefore depends on both the size of the sample used for estimation and the importance of error in the environment.

We conducted several simulations with 4-cue models. For different populations of cue profiles, we used the following methodology. (1) We used equation (6) to create  $y_i$  values for the population of cue profiles with particular specifications of the  $\beta_j$  parameters and the error term. (2) We sampled at random 50% of the  $y_i$  values of these populations and calculated goodness of fit for each of the simple models. (In essence, this involved ordering variables by the cue validities estimated in the samples.) (3) We then used the fitted values to predict the remaining

50% of the population. (4) We repeated steps (2) and (3) a further 99 times. (5) The whole process – steps (1) through (4) – was done 100 times re-generating errors from a normal distribution so that we finally obtained estimates of the predictive abilities of the models based on 10,000 trials (i.e., 100 x 100).

-----  
Insert Tables 12 and 13 about here  
-----

Table 12 provides information about the different populations used in the simulations as well as the extent to which the relative sizes of cue validities estimated at step (2) matched the relative sizes of the underlying, ecological  $\beta$ -parameters. We sampled from five different populations, using three different sets of  $\beta$ -parameters (NonCF, CF2, and CF5), and with three levels of random error (none or “0” with  $\sigma = 0$ , medium or “M” with  $\sigma = 0.40$ , and large or “L” with  $\sigma = 1.00$ ). However this was not done in a factorial manner.

The data in Tables 12 and 13 are presented in four panels: A, B, C, and D. The population distribution in panel A was the uniform distribution of all 16 distinctive cue profiles for the 4-cue case. In this panel, we used both NonCF and CF5 weighting parameters and the three levels of random error. To interpret the sizes of these errors, we estimated regression equations from our samples (regressing the  $y_i$  on the  $x_{ij}$ ) and noted how much variance was not “accounted for,” i.e.,  $(1-R^2)$  – see the right hand side of Table 12.

The populations in panel B were created by splitting the uniform population in panel A into two: one population being TTB friendly ( $n=8$ ); and the other TTB-unfriendly ( $n=8$ ). Once again, we simulated using three levels of error (0, M, L) but only used one set of  $\beta$ -parameters corresponding to CF5.

In panels C and D we created populations with the characteristics of the cue profile distributions of the “oxidant” and “land rent” data analyzed by Gigerenzer et al. (1999) – see Table 10. For both of these populations, we used two sets of  $\beta$ -parameters, CF2 and CF5, and three levels of error (0, M, L).

Table 12 shows that, as the level of environmental error increases, there is a decrease in the probability of correctly matching the relative sizes of cue validities with the corresponding ecological  $\beta$ -parameters. Moreover, even in the no error condition, only one figure is above 50% (the 58% for the TTB-friendly distribution in panel B), and the other estimates vary between 14% and 35%. As to identifying the number of times that the most important variable is identified as such, there is considerable success in the no error condition and particularly when weighting functions are more non-compensatory, e.g., NonCF vs. CF5 and CF2 vs. CF5. However, the mean rate of correct identification in the medium error condition is only about 50%.

Table 13 presents the results of cross-validated predictions. The overall conclusion is that error has a large impact on the predictive abilities of the different models. Across all panels, note first that, in the presence of large error, all models make essentially random predictions, i.e., close to 50%. With medium levels of error, differences between the performances of the different models are small with the exception of the TTB-friendly data (in panel B) where DOMRAN is some 8% below TTB and EW/TTB. Otherwise, the predictive ability of DOMRAN is never more than 5% below the performance of the other models.

The largest differences between the models occur when there is no error. In the non-compensatory case of panel A, TTB performs at 91% and DOMRAN at 70%; for the CF5 parameters EW/TTB performs at 89% and DOMRAN at 73%. There are

also large differences for the TTB-friendly case in panel B but not for the TTB-unfriendly case. Indeed, in this latter case, none of the models are too successful.

As noted above, the condition with no error captures the effects of lack of knowledge, i.e., when the relative sizes of cue validities estimated in small samples do not match their ecological counterparts. For example, from the theoretical analysis in Table 3 (that assumes a perfect match between relative sizes of cue validities and  $\beta$ -parameters), the predictions for TTB and EW/TTB are 100% and 93%, respectively, for NonCF, and 94% and 98%, respectively, for CF5. In the simulation results – see Table 13, panel A – we observe the same order of differences between the models but at some 10% below the theoretical results shown in Table 3. Interestingly, DOMRAN has a predicted 77% success rate in Table 3 but achieves somewhat less in the simulation experiment, i.e., 70% and 73%. Similar conclusions apply to the results in panels C and D.

To conclude, two types of error affect the predictive ability of TTB and TTB-dependent models (such as EW/TTB) in these simulations. One is imperfect knowledge that results in not using cues in their appropriate order. The second is noise in the relation between cues and the criterion. Moreover, the level of such noise in the environment affects lack of knowledge in the model.

## **Discussion**

This paper has investigated the important issue of why and when simple decision rules such as TTB are effective in binary choice. Our discussion is organized as follows. We first summarize our results. Second, we consider how characteristics of models and environments interact to produce variation in choices. In doing so, we emphasize the need to develop more adequate theories of decision environments and,

in particular, the role of error. Finally, we discuss the use of TTB as a prescriptive model.

### *Principal findings*

Our main results can be summarized as follows. First, in environments characterized by populations of all distinctive pairings of alternatives, no error, and known relative sizes of cue validities, TTB is an effective strategy – even in the presence of quite compensatory weighting functions. On the other hand, EW/TTB is optimal for the more compensatory functions. In the 3-cue case, EW/TTB is optimal for all compensatory functions and this holds irrespective of whether the choice set does or does not consist of all distinctive pairings of alternatives.

Second, when – in the same kinds of environments – errors are made in the relative sizes of cue validities, TTB typically remains the most effective strategy provided the most important cue is identified as such. When this is not the case, EW/TTB should be preferred even though TTB is still quite effective. However, to place this latter finding in context, recall that both EW and DOMRAN frequently perform better than TTB when this uses incorrect cue orderings.

Third, both the absolute and relative expected performances of models are affected by characteristics of sets of choice alternatives. Specifically, the numbers of dominance pairs and repeats in a distribution of cue profiles affect overall levels of expected predictive accuracy – increasing in the former and decreasing in the latter. In addition, all simple models considered here make the same predictions for all dominance pairs and have the same expected performance with respect to repeats. Thus, differences between the models only occur in subsets of data. Within these subsets, performance of the models is differentially sensitive to the presence or

absence of specific pairings of cue profiles (such as the D-E pairing in the 3-cue case). It is always possible to “engineer” environments that are more or less “friendly” to different models.

Fourth, we reanalyzed the data of Czerlinski et al. (1999) from two perspectives. In one, we used characteristics of the distributions of cue profiles to predict expected performance of the models assuming environments with no error and, in the case of TTB, correct orderings of cue validities. For these environments, TTB generally had expected predictive accuracy exceeding 90% with EW and EW/TTB not being quite as effective (with some exceptions for EW/TTB). The theoretical performance of DOMRAN was around 77% (averaging across all datasets). However, when we put these models to cross-validated predictive tests using the empirical criterion values, TTB, EW/TTB, and EW all had similar performance and were only superior to DOMRAN by 3% or 4% (across datasets, DOMRAN averaged 68%).<sup>13</sup> Error in the real-world datasets caused significant degradation in predictive ability and these analyses led to two important insights. The first was the need to understand the role of error. The second was the relative predictive success of the naïve DOMRAN model which proved to be a meaningful benchmark for the other “simple” models.

Fifth, we also demonstrated in simulation experiments that, in the presence of error, the performance of DOMRAN does not lie far behind that of the other models. In addition, we noted that error in the environment affects errors of knowledge, i.e., failure to identify the correct relative sizes of cue validities when using TTB.

---

<sup>13</sup> At the same time, it should be noted that all models made the same predictions for some two-thirds of the datasets (see Table 10). Thus, the superiority of TTB and the other models over DOMRAN was achieved on just one-third of all predictions.

*Environments and models*

Psychological understanding is enhanced to the extent that theory identifies relevant characteristics of models and environments *and* their interactions. In this paper, we have highlighted three dimensions of environments: the function “used” by the environment to weight variables; characteristics of sets of choice alternatives; and the role of error.

The models we investigated differed in terms of both knowledge and processes, i.e., prior knowledge of relative importance of variables, amount of information to consult, use of calculations, and numbers of comparisons (see Table 1). Despite these differences, it is instructive to consider similarities. In particular, all the models can be thought of as belonging to the same *general process that examines differences between cues in a sequential manner*. This is illustrated by the flowchart in Figure 3.

-----  
Insert Figure 3 about here  
-----

To amplify, consider how the difference between two cues is processed. In box 1, a cue is identified and appropriately scaled (i.e., as 0 or 1). In box 2, the difference,  $\delta_j$ , between the values on the cue for each alternative (A and B) is calculated. If  $\delta_j \neq 0$ , it is possible to stop the process (box 3) and make a choice (box 4), i.e., choose A if  $\delta_j = 1$ , and B if  $\delta_j = -1$ . (The process, however, does not need to stop at box 3.) If,  $\delta_j = 0$ , a new variable can be selected (i.e., going back to box 1), or the process can be stopped by way of boxes 6 and 7.

To explain these latter boxes, return to box 3 and the decision to continue the process. This leads to adding  $\delta_j$  to a running total in box 6 (set to zero before each

new choice).<sup>14</sup> The next decision to stop the process can be made at box 7. If yes, the decision taken (box 8) depends on the sum in box 6. If  $\sum_{j=1}^k \delta_j \neq 0$ , the decision is A if  $\sum_{j=1}^k \delta_j > 0$ , and B if  $\sum_{j=1}^k \delta_j < 0$  (box 9). If, however  $\sum_{j=1}^k \delta_j = 0$ , the decision is either made at random or the process can start again (i.e., as in EW/TTB).

As the flowchart illustrates, it is not costly for people to switch between models at a cognitive level. Indeed, the decision as to which model to use can be made while engaged in the process itself. Also, recall that all models make identical choices for those subsets of the choice alternatives that involve dominance or repeats. What matters is how the models treat the remaining choices.

To illuminate this issue, it is instructive to focus on the comparison of TTB with EW. Whereas TTB always makes choices by treating some variables as being more important than others, EW predicts ties between certain pairs of alternatives and is forced to choose between these pairs at random. However, to perform better than EW on these cases, TTB does not always need to be correct; its success rate only needs to exceed 50%. (A little knowledge is better than none.) On the other hand, on occasions when TTB is mistaken, EW sometimes makes correct decisions. In creating the EW/TTB composites, therefore, the advantages of both models can be achieved in the more compensatory environments. Indeed, as shown in the 3-cue case, the EW/TTB composite produces optimal performance and also tends toward this in the 4- and 5-cue cases as the weighting functions become more compensatory.

TTB differs from the other models in two major respects: it imposes an order in which cues are examined, and it can exit the process before consulting all

---

<sup>14</sup> We assume here the presence of a mental counter that sequentially updates the number of variables in favor of, say, alternative A by 1, and those in favor of alternative B by -1.

information (i.e., the models differ in their “stopping” rules). When the environmental weighting function is non-compensatory, stopping the process “early” is sensible precisely because subsequent cues cannot change the decision. However, as the environment becomes more compensatory, more information should be examined.

As noted above, environments can be created that are more or less “friendly” toward different models in terms of how they affect relative predictive performance (see also Shanteau & Thomas, 2000). For example, an environment where EW faced many (few) ties would be “unfriendly” (“friendly”) to EW. Similarly, we defined environments that were “TTB-friendly” or “TTB-unfriendly” by the extent to which they contained pairs leading to less or more errors made by TTB compared to the number of TTB errors that would be made in a uniform distribution of all possible pairings of alternatives with the same number of observations (cf. Table 10).

This leads to an implication and a question. The implication is that in interpreting predictive success, it is important to characterize whether environments are friendly or unfriendly to different models. The question – and it is important – is to understand the types of environments that people encounter in their decision making activities. For example, to what extent do the datasets compiled by Czerlinski et al. (1999) characterize the kinds of choices that people face in their natural ecologies? We simply do not know.

TTB both operates on binary cues and avoids the conflicts inherent in making trade-offs in choice. These features raise the intriguing issue of when TTB-like models are likely to be used in everyday life. The original work by Gigerenzer and Goldstein (1996), for example, suggested situations where recognition (yes or no) provided the first cue. More generally, it is intriguing to speculate that TTB-like processes will be triggered when the attributes of choice are characterized by the

presence (1) or absence (0) of qualitative states for which trade-offs are difficult to make. In particular, if the cues are affective in nature, TTB provides both a way to avoid possible emotional trade-offs and, in many circumstances, an effective decision tool.

From a psychological perspective, the work undertaken here can be seen as consistent with the notion that people use different strategies for binary choice. Although, as we noted, different models can sometimes be thought of as subroutines within a more general process (see Figure 3). We find it unlikely that people will always use the same heuristic but – and as evidence suggests – they will adapt the heuristics they use to the structures of the task they experience, e.g., consulting more information in more compensatory environments (cf. Payne, Bettman & Johnson, 1993).

Finally, the importance of error in the environment was highlighted by our use of Czerlinski et al's (1999) empirical datasets. However, since we do not know the true ecological models that generated the 20 datasets, it is difficult to understand exactly how error affected empirical predictions relative to theoretical predictions based on distributional characteristics. Indeed, we noted that there can be several different types of error within the theoretical linear worlds that we investigated and that some of these can interact with errors in models (e.g., knowledge – see the simulation results, above). Much more research is needed to illuminate these issues.

### *Prescriptive considerations*

In most of the environments examined in this work, TTB has been shown to be an effective, simple model of choice. To what extent, therefore, should it be prescribed as a way to choose?

Assume first that underlying assumptions are met, i.e., that the zero-order correlations between cues and criterion are known as well as the relative importance of variables. In this case, key issues center on the extent to which the environmental weighting function is compensatory and characteristics of the cue profiles.

Given sufficient resources, e.g., time, we first recommend checking for dominance. Indeed, with few cues this may be a simpler strategy than accessing relative cue validities from memory. Moreover, exploiting dominance can imbue the decision maker with appropriate confidence. Failing this, our recommendation is to use TTB or EW/TTB (and certainly in the 3-cue case). Briefly, if the decision maker feels uncomfortable about relying on TTB alone (e.g., she senses that the environment is compensatory), then the choice should also be examined using EW/TTB. If there is uncertainty about which variable is most important, then EW/TTB is the model to follow.

As a general point, it should be noted that all TTB errors (i.e., with compensatory weighting functions) occur because the decision process stops too soon, i.e., it fails to consider enough pertinent information. In the 4-cue case, for example, no errors ever occur if the process is decided by the third or fourth most important variable. Thus, the reluctance people express to base important decisions on only one or two cues may be a consequence of having experienced errors in the past when all available information was not consulted. Where people's intuitions may lead them astray is in assessing when choices require more information and when they do not.

For decisions taken under time pressure, people should exploit the fact that TTB has a high success rate. (In particular, in using this strategy they will automatically exploit dominance even though they may never know this.) Two further issues concern feedback and the relative importance of decisions. With good

feedback, people can learn to make appropriate responses. Failing accurate feedback or being faced with important decisions under time pressure, however, they are not powerless. Specifically, they can “rehearse” similar decisions (e.g., through simulations) and then use this knowledge to know what to do in real situations, e.g., what would happen if TTB or another heuristic were used in similar circumstances?

Of course, not all binary choice situations have the simple structures (e.g., linear weighting functions and binary cues) of the situations examined in this paper. There is a need for more research to examine generalizations of TTB-like models in more complex environments, i.e., involving different types of non-linear weighting functions, continuous cues, and so on. In addition, it will be important to investigate sensitivity to other errors in knowledge, e.g., concerning the zero-order correlations between cues and criterion.

## References

- Bröder, A. (2000). Assessing the empirical validity of the “Take-The-Best” heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26 (5), 1332-1346.
- Bröder, A., & Schiffer, S. (2003). Take the best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, 132 (2), 277-293.
- Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: The University of Chicago Press.
- Chater, N., Oaksford, M., Nakisa, R., & Redington, M. (2003). Fast, frugal, and rational: How rational norms explain behavior. *Organizational Behavior and Human Decision Processes*, 90, 63-86.
- Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G Gigerenzer, P. M. Todd and the ABC Research Group. *Simple heuristics that make us smart* (pp. 97-118). New York: Oxford University Press.
- Dawes, R. M. (1979). The robust beauty of improper linear models. *American Psychologist*, 34, 571-582.
- Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, 81, 95-106.
- Einhorn, H. J., & Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational Behavior and Human Performance*, 13, 171-192.
- Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650-669.

- Gigerenzer, G., & Selten, R. (2001). Rethinking rationality. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 1-13). Cambridge, MA: MIT Press.
- Gigerenzer, G., Todd, P. M., and the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision making work? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd and the ABC Research Group. *Simple heuristics that make us smart* (pp. 119-140). New York: Oxford University Press.
- Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, 52, 29-71.
- Montgomery, H. (1983). Decision rules and the search for a dominance structure: Towards a process model of decision making. In P. Humphreys, O. Svenson, & A. Vari (Eds.), *Analysing and aiding decision processes* (pp. 343-369). Amsterdam: North Holland.
- Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing “one-reason” decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(1), 53-65.
- Newell, B. R., Weston, N. J., & Shanks, D. R. (2003). Empirical tests of a fast-and-frugal heuristic: Not everyone “takes-the-best.” *Organizational Behavior and Human Decision Processes*, 91, 82-96.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.
- Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics, and

how can we tell? In G Gigerenzer, P. M. Todd and the ABC Research Group. *Simple heuristics that make us smart* (pp. 141-167). New York: Oxford University Press.

Rieskamp, J., & Hoffrage, U. (2002). *The use of simple heuristics: Inferences and preferences under time pressure*. Manuscript, Max Planck Institute for Human Development, Berlin.

Shanteau, J., & Thomas, R. P. (2000). Fast and frugal heuristics: What about unfriendly environments? *Behavioral and Brain Sciences*, 23, 762-763.

Simon, H. A. (1956). Rational choice and the structure of environments. *Psychological Review*, 63, 129-138.

Todd, P. M., & Gigerenzer, G. (2000). Précis of *Simple heuristics that make us smart*. *Behavioral and Brain Sciences*, 23, 727-780.

Table 1 -- Prior information and cognitive operations required by different models for binary choice

	<u>Prior information*</u>	<u>Amount of information to consult</u>	<u>Calculations</u>	<u>Number of Comparisons</u>	<u>Random choice if tie</u>
DOMRAN	None	All	None	Equal to number of variables	Yes
EW	None	All	Yes, two sums	One	Yes
TTB	Rank-order of importance of variables	Variable	None	Variable -- from one to number of variables	Yes
EW/TTB hybrids	Rank-order of importance of variables	All	Yes, two sums	One, if choice by EW. Otherwise, more.	Yes

\* For all models, the decision maker is assumed to know the sign of the zero order correlation between cues and the criterion.

Table 2 -- Basic analysis of 3-cue case

<u>Profiles</u>			<u>Difference vectors (column minus row)</u>							<u>Classification of difference vectors</u>								
<u>cues</u>			<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>G</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>G</u>		
<u>Profile type</u>	<u>x1</u>	<u>x2</u>	<u>x3</u>															
<u>A</u>	1	1	1	<u>B</u>	0 0 1					<u>B</u>	d							
<u>B</u>	1	1	0															
<u>C</u>	1	0	1	<u>C</u>	0 1 0	0 1 -1				<u>C</u>	d	t,c						
<u>D</u>	1	0	0															
<u>E</u>	0	1	1	<u>D</u>	0 1 1	0 1 0	0 0 1			<u>D</u>	d	d	d					
<u>F</u>	0	1	0															
<u>G</u>	0	0	1	<u>E</u>	1 0 0	1 0 -1	1 -1 0	1 -1 -1		<u>E</u>	d	t,c	t,c	w				
<u>H</u>	0	0	0															
				<u>F</u>	1 0 1	1 0 0	1 -1 1	1 -1 0	0 0 1	<u>F</u>	d	d	c	t,c	d			
				<u>G</u>	1 1 0	1 1 -1	1 0 0	1 0 -1	0 1 0	0 1 -1	<u>G</u>	d	c	d	t,c	d	t,c	
				<u>H</u>	1 1 1	1 1 0	1 0 1	1 0 0	0 1 1	0 1 0	0 0 1	<u>H</u>	d	d	d	d	d	d

Legend:

d = dominance relation

c = TTb predicts correctly in compensatory case ( $\beta_1 > \beta_2 > \beta_3$  &  $\beta_1 < (\beta_2 + \beta_3)$ )

w = TTb predicts incorrectly in compensatory case

t = EW predicts tie

No. of cases

19

8

1

28

Table 3 -- Expected predictive accuracy (%'s) for 3- and 4-cue cases

<u>Models</u>	<u>3-cue cases</u>		<u>4- cue cases</u>					
	<u>NonCF</u>	<u>CF</u>	<u>NonCF</u>	<u>CF1</u>	<u>CF2</u>	<u>CF3</u>	<u>CF4*</u>	<u>CF5*</u>
DOMRAN	84	84	77	77	77	77	77	77
EW	86	89	81	83	82	84	85	87
<i>No of errors**</i>	1	0	9	7	8	6	4	0
TTB	<b>100</b>	96	<b>100</b>	<b>98</b>	<b>99</b>	<b>98</b>	96	94
<i>No of errors**</i>	0	1	0	2	1	3	3	5
EW/TTB	96	<b>100</b>	93	94	93	95	95	<b>98</b>
<i>No of errors**</i>	1	0	9	7	8	6	4	0
EW-3/TTB	x	x	97	95	98	96	<b>99</b>	96
<i>No of errors**</i>			4	6	3	5	0	2

Notes:

NonCF = non-compensatory functions

CF = compensatory functions

\* Functions contain some ambiguous cases (3 for CF4 and 5 for CF5 -- see Appendix). Thus, even though a model may make no errors, its expected predicted accuracy is less than 100% due to the presence of these ambiguous cases.

\*\* Errors involve misclassifications by the models (from totals of 28 and 120 choices for the 3- and 4-cue cases, respectively).

Table 4 -- Exemplar weights of different weighting functions

4-cue models

	<u><math>\beta_1</math></u>	<u><math>\beta_2</math></u>	<u><math>\beta_3</math></u>	<u><math>\beta_4</math></u>
NonCF	0.53	0.24	0.13	0.10
CF1	0.57	0.19	0.14	0.10
CF2	0.48	0.29	0.14	0.10
CF3	0.48	0.22	0.17	0.13
CF4	0.42	0.31	0.15	0.12
CF5	0.40	0.25	0.20	0.15
CF5	0.29	0.26	0.24	0.21

5-cue models

	<u><math>\beta_1</math></u>	<u><math>\beta_2</math></u>	<u><math>\beta_3</math></u>	<u><math>\beta_4</math></u>	<u><math>\beta_5</math></u>
NonCF	0.55	0.24	0.11	0.07	0.03
CF1	0.53	0.25	0.10	0.07	0.05
CF2	0.56	0.21	0.13	0.09	0.01
CF3	0.55	0.20	0.10	0.09	0.06
CF4	0.52	0.19	0.15	0.08	0.06
CF5	0.55	0.21	0.12	0.10	0.02
CF6	0.49	0.26	0.17	0.05	0.03
CF7	0.48	0.26	0.13	0.08	0.06
CF8	0.48	0.26	0.14	0.08	0.05
CF9	0.50	0.23	0.13	0.08	0.06
CF10	0.47	0.21	0.17	0.08	0.07
CF11	0.49	0.19	0.16	0.10	0.07
CF12	0.46	0.28	0.16	0.06	0.05
CF13	0.46	0.28	0.11	0.09	0.06
CF14	0.45	0.27	0.14	0.08	0.05
CF15	0.44	0.27	0.14	0.09	0.06
CF16	0.44	0.25	0.18	0.09	0.04
CF17	0.42	0.20	0.14	0.13	0.11
CF18	0.42	0.36	0.12	0.07	0.03
CF19	0.35	0.33	0.16	0.10	0.07
CF20	0.34	0.29	0.18	0.09	0.09
CF21	0.36	0.29	0.17	0.10	0.07
CF22	0.32	0.27	0.22	0.17	0.02
CF23	0.36	0.21	0.18	0.15	0.09
CF23	0.24	0.22	0.20	0.18	0.16

Table 5 -- Expected predictive accuracy (%'s) for the 5-cue case

<u>Models</u>	<u>NonCF</u>	<u>CF2</u>	<u>CF4</u>	<u>CF6</u>	<u>CF8</u>	<u>CF10</u>	<u>CF12</u>	<u>CF14</u>	<u>CF16</u>	<u>CF18</u>	<u>CF20</u>	<u>CF22</u>	<u>CF23</u>
DOMRAN	71	71	71	71	71	71	71	71	71	71	71	71	71
EW	78	78	79	78	78	80	79	79	80	81	81	83	85
<i>No of errors*</i>	55	53	45	54	52	44	49	47	39	34	30	14	2
TTB	<b>100</b>	<b>100</b>	<b>98</b>	<b>100</b>	<b>99</b>	<b>98</b>	<b>99</b>	<b>99</b>	<b>97</b>	96	96	93	91
<i>No of errors*</i>	0	2	6	1	3	7	3	5	9	9	11	15	21
EW/TTB	89	89	90	89	90	91	90	90	91	91	92	93	<b>95</b>
<i>No of errors*</i>	55	53	45	54	52	44	49	47	39	34	30	14	2
EW-4/TTB	93	93	94	93	93	94	94	94	95	95	96	<b>96</b>	<b>95</b>
<i>No of errors*</i>	36	34	28	35	33	27	30	28	22	16	12	1	5
EW-3/TTB	97	96	95	97	97	95	98	97	96	<b>98</b>	<b>98</b>	95	93
<i>No of errors*</i>	16	18	22	15	17	21	10	12	16	0	2	6	13
Ambiguous cases**	0	0	6	0	0	6	3	3	9	19	21	38	43

Notes:

NonCF = non-compensatory functions

CF = compensatory functions

\* Errors involve misclassifications by the models (from total of 496 choices).

\*\* Functions contain some ambiguous cases (see Appendix). All models are assumed to be 50% correct on these cases.

Table 6 -- Sensitivity to different cue orderings for  
3-cue case (populations of distinctive pairings)

	<u>Cue orderings</u>			<u>Expected correct - %</u>			
	<u>1st</u>	<u>2nd</u>	<u>3rd</u>	<u>Non-compensatory</u>		<u>Compensatory</u>	
				<u>TTB</u>	<u>EW/TTB</u>	<u>TTB</u>	<u>EW/TTB</u>
1	x1	x2	x3	<b>100</b>	96	96	<b>100</b>
2	x1	x3	x2	<b>93</b>	89	89	<b>93</b>
3	x2	x1	x3	86	<b>89</b>	89	<b>93</b>
4	x2	x3	x1	79	<b>82</b>	82	<b>86</b>
5	x3	x1	x2	79	<b>82</b>	82	<b>86</b>
6	x3	x2	x1	71	<b>75</b>	75	<b>79</b>
			Means	<u>85</u>	<u>86</u>	<u>86</u>	<u>89</u>

Notes:

(1) **Bold** entries indicate best predictions within orderings.

(2) Expected performance (%) of models not affected by ordering:

	<u>NonCF</u>	<u>CF</u>
DOMRAN	84	84
EW	86	89

Table 7 -- Sensitivity to different cue orderings for  
some 4-cue models (populations of distinctive pairings)

					<u>Expected correct -- %</u>					
<u>Cue orderings</u>					<u>Non-Compensatory</u>		<u>Compensatory</u>			
							<u>CF2</u>		<u>CF5</u>	
<u>1st</u>	<u>2nd</u>	<u>3rd</u>	<u>4th</u>		<u>TTB</u>	<u>EW/TTB</u>	<u>TTB</u>	<u>EW/TTB</u>	<u>TTB</u>	<u>EW/TTB</u>
1	x1	x2	x3	x4	<b>100</b>	93	<b>99</b>	93	94	<b>98</b>
2	x1	x2	x4	x3	<b>97</b>	89	<b>96</b>	90	90	<b>95</b>
3	x1	x3	x2	x4	<b>93</b>	89	<b>93</b>	90	90	<b>95</b>
4	x1	x3	x4	x2	<b>90</b>	86	<b>89</b>	87	87	<b>91</b>
5	x1	x4	x2	x3	<b>90</b>	86	<b>89</b>	87	87	<b>91</b>
6	x1	x4	x3	x2	<b>87</b>	83	<b>86</b>	83	84	<b>88</b>
7	x2	x1	x3	x4	87	<b>88</b>	<b>88</b>	<b>88</b>	88	<b>94</b>
8	x2	x1	x4	x3	83	<b>84</b>	84	<b>85</b>	85	<b>90</b>
9	x2	x3	x1	x4	80	<b>84</b>	81	<b>85</b>	83	<b>90</b>
10	x2	x3	x4	x1	77	<b>81</b>	78	<b>82</b>	80	<b>87</b>
11	x2	x4	x1	x3	77	<b>81</b>	78	<b>82</b>	80	<b>87</b>
12	x2	x4	x3	x1	73	<b>78</b>	74	<b>78</b>	76	<b>84</b>
13	x3	x1	x2	x4	80	<b>84</b>	81	<b>85</b>	85	<b>90</b>
14	x3	x1	x4	x2	77	<b>81</b>	78	<b>82</b>	81	<b>87</b>
15	x3	x2	x1	x4	73	<b>81</b>	74	<b>82</b>	80	<b>87</b>
16	x3	x2	x4	x1	70	<b>78</b>	71	<b>78</b>	76	<b>84</b>
17	x3	x4	x1	x2	70	<b>78</b>	71	<b>78</b>	76	<b>84</b>
18	x3	x4	x2	x1	67	<b>74</b>	68	<b>75</b>	73	<b>80</b>
19	x4	x1	x2	x3	77	<b>81</b>	78	<b>82</b>	79	<b>86</b>
20	x4	x1	x3	x2	73	<b>78</b>	74	<b>78</b>	75	<b>83</b>
21	x4	x2	x1	x3	70	<b>78</b>	71	<b>78</b>	75	<b>83</b>
22	x4	x2	x3	x1	67	<b>74</b>	68	<b>75</b>	72	<b>80</b>
23	x4	x3	x1	x2	67	<b>74</b>	68	<b>75</b>	72	<b>80</b>
24	x4	x3	x2	x1	63	<b>71</b>	64	<b>72</b>	69	<b>76</b>
				Means	<u>79</u>	<u>81</u>	<u>79</u>	<u>82</u>	<u>81</u>	<u>87</u>

Notes:

(1) **Bold** entries indicate best predictions within orderings.

(2) Expected performance (%) of models not affected by ordering

	<u>NonCF</u>	<u>CF2</u>	<u>CF5</u>
DOMRAN	77	77	77
EW	81	82	87

Table 8 -- Sensitivity to different cue orderings for some 5-cue models (populations of distinctive pairings)

Cue orderings						Expected correct -- %									
						Non-compensatory		Compensatory							
								CF1		CF9		CF17		CF23	
1st	2nd	3rd	4th	5th	TTB	EW/TTB	TTB	EW/TTB	TTB	EW/TTB	TTB	EW/TTB	TTB	EW/TTB	
1	x1	x2	x3	x4	x5	<b>100</b>	89	<b>99</b>	90	<b>99</b>	90	<b>96</b>	93	91	<b>95</b>
2	x1	x2	x3	x5	x4	<b>98</b>	87	<b>98</b>	88	<b>97</b>	89	<b>94</b>	91	90	<b>94</b>
3	x1	x2	x4	x3	x5	<b>97</b>	87	<b>98</b>	88	<b>97</b>	89	<b>94</b>	91	90	<b>94</b>
4	x1	x2	x4	x5	x3	<b>95</b>	86	<b>96</b>	87	<b>95</b>	87	<b>92</b>	90	88	<b>92</b>
5	x1	x2	x5	x3	x4	<b>95</b>	86	<b>96</b>	87	<b>95</b>	87	<b>92</b>	90	88	<b>92</b>
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
11	x1	x3	x5	x2	x4	<b>89</b>	83	<b>88</b>	84	<b>88</b>	85	<b>89</b>	88	85	<b>90</b>
12	x1	x3	x5	x4	x2	<b>87</b>	82	<b>86</b>	83	<b>87</b>	83	<b>87</b>	86	83	<b>88</b>
13	x1	x4	x2	x3	x5	<b>90</b>	85	<b>91</b>	86	<b>91</b>	86	<b>91</b>	89	87	<b>92</b>
14	x1	x4	x2	x5	x3	<b>89</b>	83	<b>90</b>	84	<b>90</b>	85	<b>90</b>	88	85	<b>90</b>
15	x1	x4	x3	x2	x5	<b>87</b>	83	<b>88</b>	84	<b>88</b>	85	<b>89</b>	88	85	<b>90</b>
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
21	x1	x5	x3	x2	x4	<b>86</b>	82	<b>86</b>	83	<b>87</b>	83	<b>87</b>	86	83	<b>88</b>
22	x1	x5	x3	x4	x2	<b>84</b>	80	<b>85</b>	81	<b>85</b>	82	<b>85</b>	84	81	<b>86</b>
23	x1	x5	x4	x2	x3	<b>84</b>	80	<b>85</b>	81	<b>85</b>	82	<b>85</b>	84	81	<b>86</b>
24	x1	x5	x4	x3	x2	<b>82</b>	78	<b>83</b>	79	<b>83</b>	80	<b>84</b>	82	79	<b>85</b>
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25	x2	x1	x3	x4	x5	<b>87</b>	85	<b>86</b>	<b>86</b>	<b>86</b>	<b>86</b>	85	<b>89</b>	86	<b>92</b>
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
31	x2	x3	x1	x4	x5	81	<b>83</b>	80	<b>83</b>	80	<b>84</b>	79	<b>86</b>	82	<b>91</b>
32	x2	x3	x1	x5	x4	79	<b>81</b>	78	<b>82</b>	78	<b>82</b>	77	<b>85</b>	81	<b>89</b>
33	x2	x3	x4	x1	x5	77	<b>81</b>	77	<b>82</b>	76	<b>82</b>	76	<b>85</b>	80	<b>89</b>
34	x2	x3	x4	x5	x1	76	<b>79</b>	75	<b>80</b>	75	<b>81</b>	74	<b>83</b>	78	<b>87</b>
35	x2	x3	x5	x1	x4	76	<b>79</b>	75	<b>80</b>	75	<b>81</b>	74	<b>83</b>	78	<b>87</b>
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
41	x2	x4	x5	x1	x3	73	<b>78</b>	73	<b>78</b>	73	<b>79</b>	73	<b>82</b>	77	<b>86</b>
42	x2	x4	x5	x3	x1	71	<b>76</b>	72	<b>77</b>	72	<b>77</b>	71	<b>80</b>	75	<b>84</b>
43	x2	x5	x1	x3	x4	76	<b>79</b>	77	<b>80</b>	76	<b>81</b>	76	<b>83</b>	79	<b>87</b>
44	x2	x5	x1	x4	x3	74	<b>78</b>	75	<b>78</b>	75	<b>79</b>	74	<b>82</b>	77	<b>85</b>
45	x2	x5	x3	x1	x4	73	<b>78</b>	73	<b>78</b>	73	<b>79</b>	73	<b>82</b>	76	<b>85</b>
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
61	x3	x4	x1	x2	x5	71	<b>78</b>	70	<b>79</b>	71	<b>80</b>	74	<b>83</b>	77	<b>87</b>
62	x3	x4	x1	x5	x2	69	<b>77</b>	69	<b>78</b>	69	<b>78</b>	72	<b>81</b>	75	<b>85</b>
63	x3	x4	x2	x1	x5	68	<b>77</b>	67	<b>78</b>	68	<b>78</b>	71	<b>81</b>	75	<b>85</b>
64	x3	x4	x2	x5	x1	66	<b>75</b>	65	<b>76</b>	66	<b>77</b>	69	<b>80</b>	73	<b>84</b>
65	x3	x4	x5	x1	x2	66	<b>75</b>	65	<b>76</b>	66	<b>77</b>	69	<b>80</b>	73	<b>84</b>
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
81	x4	x2	x3	x1	x5	68	<b>77</b>	69	<b>78</b>	69	<b>78</b>	71	<b>81</b>	75	<b>85</b>
82	x4	x2	x3	x5	x1	66	<b>75</b>	67	<b>76</b>	68	<b>77</b>	70	<b>80</b>	73	<b>83</b>
83	x4	x2	x5	x1	x3	66	<b>75</b>	67	<b>76</b>	68	<b>77</b>	70	<b>80</b>	73	<b>83</b>
84	x4	x2	x5	x3	x1	65	<b>74</b>	65	<b>74</b>	66	<b>75</b>	68	<b>78</b>	72	<b>82</b>
85	x4	x3	x1	x2	x5	68	<b>77</b>	69	<b>78</b>	69	<b>78</b>	72	<b>81</b>	75	<b>85</b>
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
116	x5	x4	x1	x3	x2	63	<b>72</b>	64	<b>73</b>	64	<b>73</b>	67	<b>76</b>	67	<b>78</b>
117	x5	x4	x2	x1	x3	61	<b>72</b>	62	<b>73</b>	63	<b>73</b>	65	<b>76</b>	67	<b>78</b>
118	x5	x4	x2	x3	x1	60	<b>70</b>	61	<b>71</b>	61	<b>72</b>	63	<b>74</b>	66	<b>77</b>
119	x5	x4	x3	x1	x2	60	<b>70</b>	61	<b>71</b>	61	<b>72</b>	63	<b>74</b>	66	<b>77</b>
120	x5	x4	x3	x2	x1	58	<b>69</b>	59	<b>70</b>	60	<b>70</b>	62	<b>73</b>	64	<b>75</b>
					Means	<u>74</u>	<u>78</u>	<u>74</u>	<u>79</u>	<u>74</u>	<u>79</u>	<u>76</u>	<u>82</u>	<u>77</u>	<u>85</u>

Notes:

(1) Bold entries predict the best predictions within orderings.

(2) Expected performance (%) of models not affected by ordering:

	NonCF	CF1	CF9	CF17	CF23
DOMRAN	71	71	71	71	71
EW	78	79	79	82	85

Table 9-- Some distributions of cue profiles for the 3-cue case

<u>Cue profile type</u>	<u>x1</u>	<u>x2</u>	<u>x3</u>	<u>Distributions</u>		
				(entries: number of profiles)		
				<u>I</u>	<u>II</u>	<u>III</u>
A	1	1	1	1	1	0
B	1	1	0	1	1	2
C	1	0	1	1	1	0
D	1	0	0	1	1	2
E	0	1	1	1	0	6
F	0	1	0	1	1	0
G	0	0	1	1	1	3
H	0	0	0	1	1	3

Characteristics of distributions

a) Total number of binary choices	28	21	120
b) Percentages of choices involving dominance	68	71	51
c) Presence of repeats?	No	No	Yes
d) TTB-error cases DE, as percentage of total	4	0	10
e) Overall characterization	TTB-neutral	TTB-friendly	TTB-unfriendly

Predicted correct (%'s) assuming  
compensatory weighting function:

TTB	96	<b>100</b>	80
EW	89	91	83
EW/TTB	<b>100</b>	<b>100</b>	<b>90</b>
DOMRAN	84	86	75

Note: **Bold** entries indicate best predictions within distributions.

Table 10 -- Characterization of empirical datasets

<u>Datasets</u>	<u>n</u>	<u>Choices (%) involving:</u>		<u>Ratio of expected TTB</u>		<u>Overall</u>	
		<u>Dominant</u>	<u>Repeats</u>	<u>errors:Uniform/actual</u>		<u>characterized as</u>	
		<u>Pairs</u>	<u>Repeats</u>				
<b><u>3-cues:</u></b>						<b><u>CF</u></b>	
Ozone	11	53	18	Infinite			F
Attractiveness of women	30	51	17	Infinite			F
Attractiveness of men	32	57	21	2.67			F
Fish fertility	395	72	21	Infinite			F
<b><u>4-cues:</u></b>				<b><u>CF2</u></b>	<b><u>CF5</u></b>	<b><u>CF2</u></b>	<b><u>CF5</u></b>
Oxidant	17	45	6	Infinite	2.50	F	F
Land rent	58	42	9	0.29	0.39	U	U
<b><u>5-cues:</u></b>				<b><u>CF9</u></b>	<b><u>CF23</u></b>	<b><u>CF9</u></b>	<b><u>CF23</u></b>
Homelessness	50	47	6	Infinite	1.24 to 4.94	F	F
Body fat	218	66	8	7.62	4.88	F	F
City populations	83	67	25	28/1 to 63/1	84/1 to 189/1	F	F
High school dropout rates	57	70	26	4.67	8.40	F	F
Cows' manure	14	13	33	Infinite	Infinite	F	F
Mortality	20	71	8	1.17	1.62	N	F
House prices	22	75	13	1.75	3.00	N	F
Car accidents	37	35	4	Infinite	1.11	F	N
Rainfall	24	49	4	1.75	1.24	N	N
Obesity at 18	46	72	10	0.41 to 1.65	0.55 to 2.21	N	N
Fuel consumption	48	47	6	0.13 to 0.52	0.38 to 1.50	U	N
Professors' salaries	51	48	9	0.47	0.89	U	N
Mammals' sleep	35	49	14	0.35	1.00	U	N
Biodiversity	26	42	7	0.35	0.68	U	U
Means		<u>54</u>	<u>13</u>				

Legend:

F stands for TTB-friendly  
U stands for TTB-unfriendly  
N stands for TTB-neutral

Table 11 -- Theoretical predictions (% correct) for 20 datasets based on  
distributions of cue profiles

Datasets

**3-cue:**

	<u>NonCF</u>		<u>CF</u>		
	<u>DOMRAN</u>	<u>TTB</u>	<u>TTB</u>	<u>EW/TTB</u>	<u>EW</u>
Ozone	76	91	91	91	87
Attractiveness of women	75	90	91	90	80
Attractiveness of men	78	89	89	90	81
Fish fertility	86	89	89	89	86
Means	<u>79</u>	<u>90</u>	<u>90</u>	<u>90</u>	<u>84</u>

**4-cue:**

	<u>NonCF</u>		<u>CF2</u>			<u>CF5</u>		
	<u>DOMRAN</u>	<u>TTB</u>	<u>TTB</u>	<u>EW/TTB</u>	<u>EW</u>	<u>TTB</u>	<u>EW/TTB</u>	<u>EW</u>
Oxidant	72	97	96	91	81	93	95	84
Land rent	71	96	92	82	71	79	91	82
Means	<u>72</u>	<u>97</u>	<u>94</u>	<u>87</u>	<u>76</u>	<u>86</u>	<u>93</u>	<u>83</u>

**5-cue:**

	<u>NonCF</u>		<u>CF9</u>			<u>CF23</u>		
	<u>DOMRAN</u>	<u>TTB</u>	<u>TTB</u>	<u>EW/TTB</u>	<u>EW</u>	<u>TTB</u>	<u>EW/TTB</u>	<u>EW</u>
Homelessness	74	97	97	94	84	94	95	86
Body fat	83	96	96	93	89	94	95	91
City populations	84	87	87	87	84	87	87	84
High school dropout rates	85	87	87	87	86	86	87	86
Cows' manure	57	84	84	82	82	83	83	83
Mortality	85	96	93	86	83	87	92	89
House prices	87	93	92	89	87	89	92	90
Car accidents	68	98	98	91	79	91	94	84
Rainfall	74	98	97	90	82	89	95	88
Obesity at 18	86	95	93	91	86	90	94	90
Fuel consumption	73	97	92	87	79	88	91	83
Professors' salaries	74	96	91	78	71	82	86	80
Mammals' sleep	75	93	90	79	72	89	79	73
Biodiversity	71	97	81	90	75	81	90	84
Means	<u>77</u>	<u>94</u>	<u>91</u>	<u>87</u>	<u>81</u>	<u>88</u>	<u>90</u>	<u>85</u>

Table 12 -- Relative accuracy of cue validities from known populations

		<u>Estimated cue validities</u> <u>match relative sizes</u> <u>of <math>\beta</math>-parameters (%)</u>			<u>Cue associated with largest</u> <u><math>\beta</math>-parameter has largest</u> <u>estimated cue validity (%)</u>			<u>Size of error</u> <u>expressed as <math>(1-R^2)</math></u>		
		<u>Error:</u> <u>None</u>	<u>Medium</u>	<u>Large</u>	<u>None</u>	<u>Medium</u>	<u>Large</u>	<u>None</u>	<u>Medium</u>	<u>Large</u>
<u>Distributions</u>										
<u>(A)</u>	Uniform (n=16)									
	NonCF	32	12	7	100	54	33	0.00	0.41	0.55
	CF5	28	10	7	80	40	30	0.00	0.41	0.61
<u>(B)</u>	TTB-friendly CF5 (n=8)	58	32	31	97	68	55	0.00	0.14	0.33
	TTB-unfriendly CF5 (n=8)	25	18	28	58	46	54	0.00	0.41	0.44
<u>(C)</u>	Oxidant (n=17)									
	CF2	14	7	6	92	45	30	0.00	0.47	0.68
	CF5	15	7	5	75	36	31	0.00	0.50	0.66
<u>(D)</u>	Land rent (n=58)									
	CF2	35	14	6	100	60	35	0.00	0.60	0.80
	CF5	25	11	4	38	44	31	0.00	0.64	0.81

Table 13 -- Correct predictions (%) for different populations and models under varying sets of weighting parameters and error conditions: averages based on 10.000 trials (4-cue case)

(A) <u>Distribution:</u>	<u>Uniform (n = 16)*</u>							
<u>β-parameters:</u>	<u>NonCF</u>				<u>CF5</u>			
	<u>TTB</u>	<u>EW/ TTB</u>	<u>EW</u>	<u>DOMRAN</u>	<u>TTB</u>	<u>EW/ TTB</u>	<u>EW</u>	<u>DOMRAN</u>
<u>Error</u>								
None	<b>91</b>	83	73	70	86	<b>89</b>	81	73
Medium	<b>62</b>	<b>62</b>	60	57	60	<b>61</b>	60	57
Large	49	49	<b>50</b>	47	47	47	<b>48</b>	47
(B) <u>Distribution:</u>	<u>TTB friendly (n = 8)**</u>				<u>TTB unfriendly (n = 8)***</u>			
<u>β-parameters:</u>	<u>CF5</u>				<u>CF5</u>			
	<u>TTB</u>	<u>EW/ TTB</u>	<u>EW</u>	<u>DOMRAN</u>	<u>TTB</u>	<u>EW/ TTB</u>	<u>EW</u>	<u>DOMRAN</u>
<u>Error</u>								
None	<b>97</b>	95	90	78	<b>56</b>	53	54	<b>56</b>
Medium	<b>68</b>	<b>68</b>	66	60	51	<b>52</b>	<b>52</b>	<b>52</b>
Large	50	50	<b>51</b>	48	<b>45</b>	<b>45</b>	<b>45</b>	<b>45</b>
(C) <u>Distribution:</u>	<u>Oxidant (n = 17)</u>							
<u>β-parameters:</u>	<u>CF2</u>				<u>CF5</u>			
	<u>TTB</u>	<u>EW/ TTB</u>	<u>EW</u>	<u>DOMRAN</u>	<u>TTB</u>	<u>EW/ TTB</u>	<u>EW</u>	<u>DOMRAN</u>
<u>Error</u>								
None	<b>81</b>	<b>81</b>	78	70	80	<b>81</b>	77	70
Medium	<b>62</b>	<b>62</b>	61	57	<b>59</b>	<b>59</b>	58	56
Large	49	49	<b>50</b>	48	48	48	<b>49</b>	47
(D) <u>Distribution:</u>	<u>Land rent (n = 58)</u>							
<u>β-parameters:</u>	<u>CF2</u>				<u>CF5</u>			
	<u>TTB</u>	<u>EW/ TTB</u>	<u>EW</u>	<u>DOMRAN</u>	<u>TTB</u>	<u>EW/ TTB</u>	<u>EW</u>	<u>DOMRAN</u>
<u>Error</u>								
None	<b>83</b>	76	68	65	79	<b>84</b>	78	70
Medium	<b>60</b>	59	57	56	58	<b>59</b>	58	56
Large	<b>48</b>	<b>48</b>	<b>48</b>	47	47	47	<b>48</b>	46

Notes:

<u>β-parameters</u>	<u>NonCF</u>	<u>CF2</u>	<u>CF5</u>
β1	0.52	0.48	0.40
β2	0.24	0.28	0.25
β3	0.14	0.14	0.20
β4	0.10	0.10	0.15

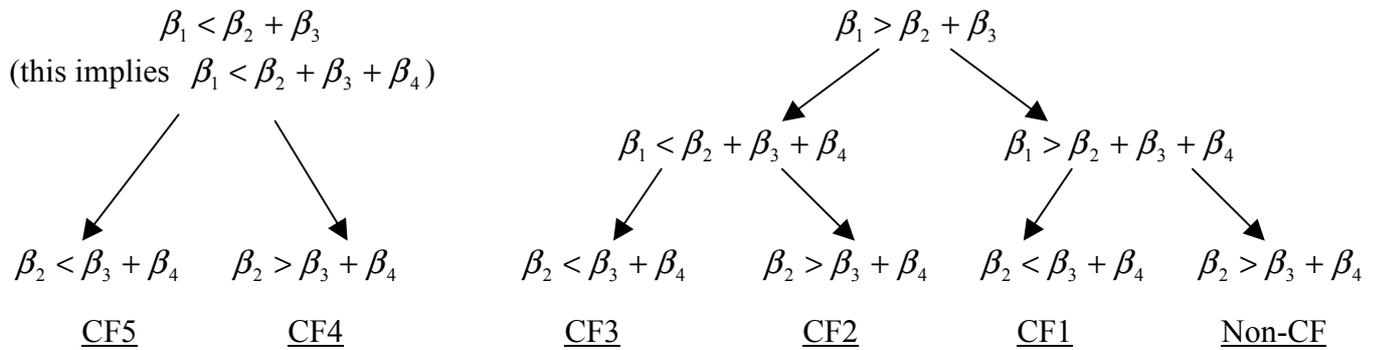
\* Population of distinctive cue profiles profiles involved in TTB errors

\*\* Population of distinctive cue profiles less

\*\*\* Population of distinctive cue profiles less profiles not involved in TTB errors

Errors are normally distributed with mean of 0 and standard deviation of 0,40 (medium) and 1,00 (large).

**Figure 1 -- Classification of compensatory (CF) and non-compensatory functions (non-CF) for the 4-cue case**



In tabular representation:

Functions / Between-weights relations	$\beta_1$ vs. $\beta_2 + \beta_3$	$\beta_1$ vs. $\beta_2 + \beta_3 + \beta_4$	$\beta_2$ vs. $\beta_3 + \beta_4$
Non-CF	$\beta_1 > \beta_2 + \beta_3$	$\beta_1 > \beta_2 + \beta_3 + \beta_4$	$\beta_2 > \beta_3 + \beta_4$
CF1	$\beta_1 > \beta_2 + \beta_3$	$\beta_1 > \beta_2 + \beta_3 + \beta_4$	$\beta_2 < \beta_3 + \beta_4$
CF2	$\beta_1 > \beta_2 + \beta_3$	$\beta_1 < \beta_2 + \beta_3 + \beta_4$	$\beta_2 > \beta_3 + \beta_4$
CF3	$\beta_1 > \beta_2 + \beta_3$	$\beta_1 < \beta_2 + \beta_3 + \beta_4$	$\beta_2 < \beta_3 + \beta_4$
CF4	$\beta_1 < \beta_2 + \beta_3$	$\beta_1 < \beta_2 + \beta_3 + \beta_4$	$\beta_2 > \beta_3 + \beta_4$
CF5	$\beta_1 < \beta_2 + \beta_3$	$\beta_1 < \beta_2 + \beta_3 + \beta_4$	$\beta_2 < \beta_3 + \beta_4$

Figure 2 -- Predictions in holdout samples with 1,000 trials.

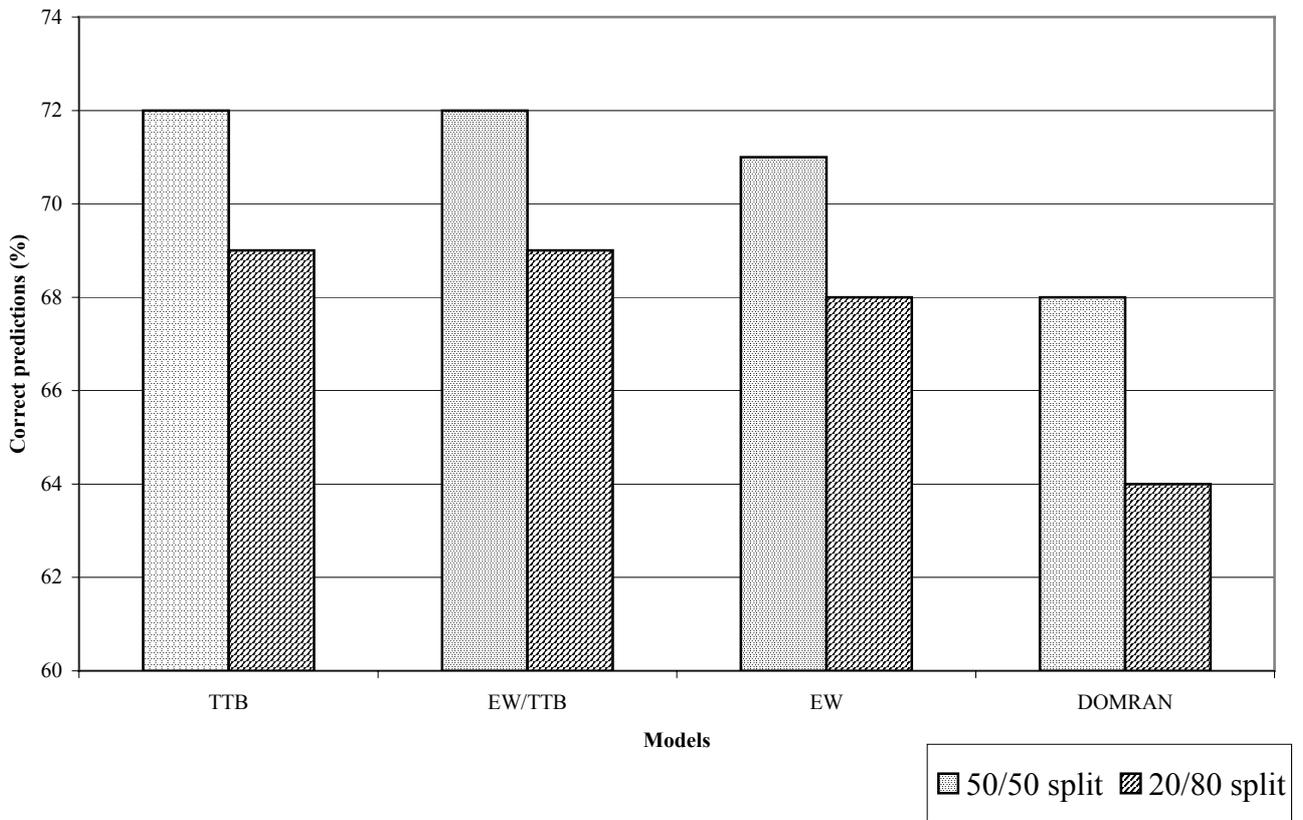
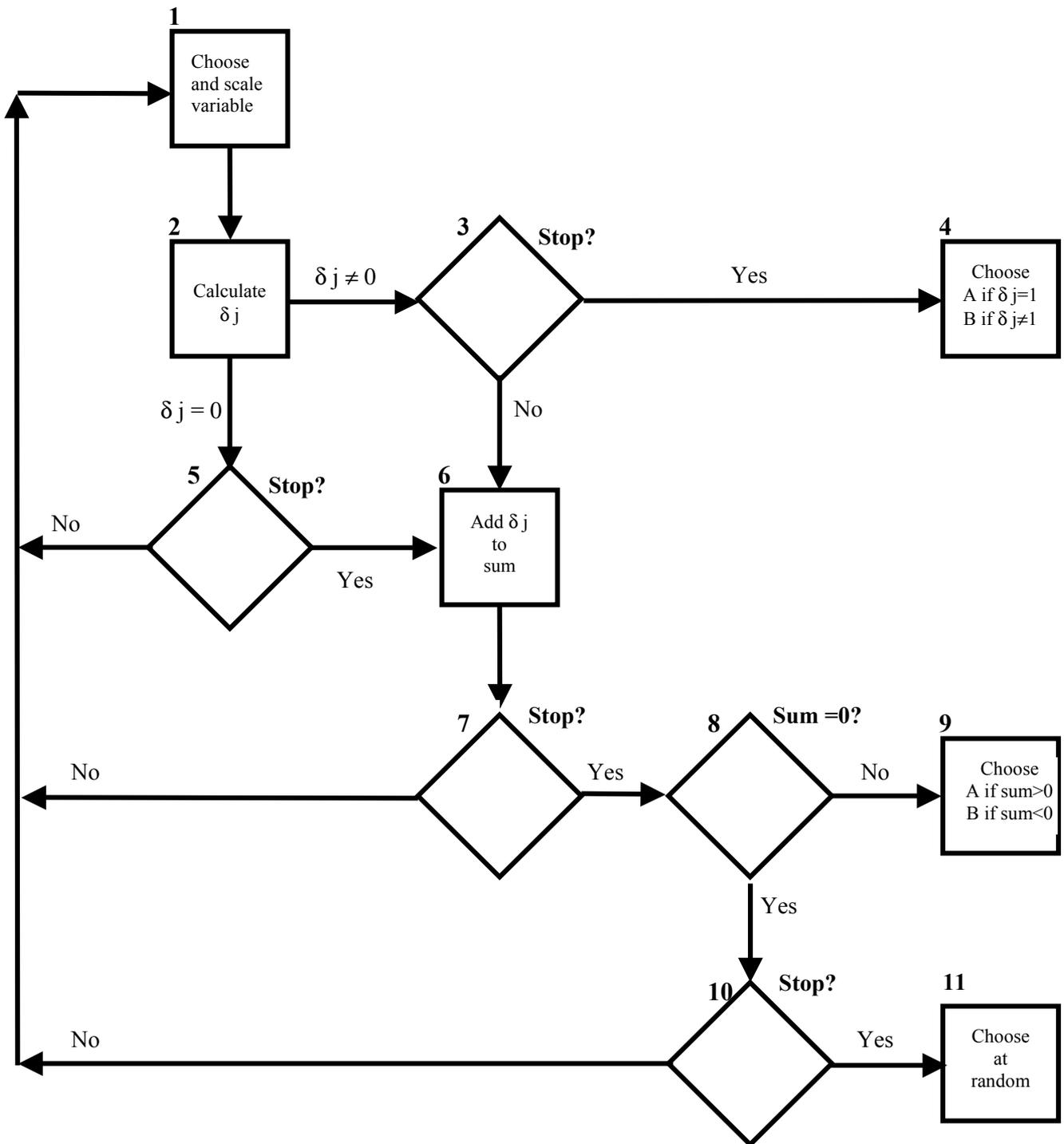


Figure 3 -- Flowchart of simple binary choice models



## Appendix – Specification of “incorrect” and “uncertain” cases for the 3-, 4-, and 5-cue models

As demonstrated by Martignon and Hoffrage (1999; 2002), TTB is optimal within a theoretical world when the weighting function is non-compensatory. Using the same notation as the main text, Martignon and Hoffrage define non-compensatory weighting functions as having the property that, when ordered from largest to smallest, each weight is greater than the sum of all weights to its right-hand side, i.e.

$$\beta_j > \sum_i \beta_i, \text{ for any } i > j, j = 1, \dots, k - 1. \quad (\text{A1})$$

All other weighting functions are said to be compensatory. Thus, all compensatory weighting functions are *broadly* defined by

$$\beta_j \leq \sum_i \beta_i, \text{ for any } i > j, j = 1, \dots, k - 1, \quad (\text{A2})$$

which is the complementary set to that defined by (A1).

To understand why we stress the broadness of this definition, consider the following two weighting functions for the 4-cue case. In the first, all the weights  $\beta_j$  satisfy the compensatory condition (A2). In the second, all the weights  $\beta_j$  satisfy the non-compensatory condition (A1) except for the second weight  $\beta_2$ , where  $\beta_2 < \beta_3 + \beta_4$ . Clearly, by the definition in (A1), both functions are compensatory. However, as this example illustrates, compensatory functions can be distinguished by “degree of compensation” which, in turn, can be defined by the number of restrictions on the relative sizes of the weights that do not satisfy the non-compensatory condition (A1). Thus, whereas all the weights of the first function fail to satisfy condition (A1), there is one weight in the second equation,  $\beta_2$  that does meet the non-compensatory condition (A1). In this sense, therefore, we can say that the first function is “more compensatory” than the second.

It should also be emphasized that the range of compensatory functions includes functions with both weak and strict compensatory relations between cues, i.e. with  $\beta_j \leq \sum_i \beta_i$  and  $\beta_j < \sum_i \beta_i$ , for any  $i > j, j = 1, \dots, k - 1$ . Thus, functions that involve a mix of weak compensatory, strict compensatory and non-compensatory relations are also classified as compensatory due to fact they are not in the set defined by (A1). For example, the 4-cue case function, which is characterized by the following restrictions  $\beta_1 \leq \beta_2 + \beta_3$ ,  $\beta_1 < \beta_2 + \beta_3 + \beta_4$ , and  $\beta_2 > \beta_3 + \beta_4$ , is also considered compensatory by the Martignon and Hoffrage (1999; 2002) definition.

To understand the effectiveness of TTB when dealing with compensatory weighting functions, the “optimal” would be to prove a theorem analogous to the non-compensatory case. Unfortunately, this approach has not yet yielded satisfactory results. On the other hand, by specifying forms of compensatory weighting schemes that span the possible space of functions, we can still gain considerable insight into the properties of TTB. This is the pragmatic approach adopted in this paper and involves the following steps.

First, we examine the nature of the conflicts between the  $n(n-1)/2$  possible distinctive pairings of cue profiles where  $n = 2^k$  and  $k$  is the number of cues in the model being examined. This examination is done using the vector of differences in cue profiles as described in the main text (equation 4). As an example, consider two cue profiles, A and B, and  $k$  cues. This leads to the difference vector

$$A - B = \{x_{a1}-x_{b1}, x_{a2}-x_{b2}, \dots, x_{ak}-x_{bk}\} \quad (A3)$$

where the variables have been ordered, from left to right, in accordance with the relative sizes of cue weights (from largest to smallest) and, by definition, elements can only take the values 1, 0, or -1. (As noted in the main text, application of TTB simply involves reading the difference vector from left to right and choosing according to the

first nonzero element. If this is 1, the choice is A; if it is -1, the choice is B. If all elements are equal to 0, choice is made at random. However, this latter case cannot occur when only distinctive pairings are considered).

Second, having specified the difference vectors, these are separated into four categories: dominance; correct; incorrect; and uncertain.

For cases of dominance, all elements of each difference vectors must be non-negative and at least one must be positive. In these cases, TTB makes the same choices as any weighting strategy, compensatory or not, characterized by  $\beta_j \geq 0$  for all  $j = 1, \dots, k$ .

To classify difference vectors to the other categories, we need to provide more precise specification of compensatory weighting functions, broadly defined by (A2).

For ease of exposition, consider first the 3-cue case. To classify the difference vectors that do not involve dominance, we exploit the fact that  $\beta_j > \beta_{j+1}$  for  $j = 1, \dots, k-1$ . There are 6 distinct difference vector profiles: (1,1,-1), (1, 0,-1), (1,-1,1), (1,-1,0), (0,1,-1), and (1,-1,-1) – see Table 2. (Note, some difference vectors have the same profiles.) For the first five profiles, it must be the case that any rule satisfying  $\beta_j > \beta_{j+1}$  for  $j = 1, \dots, k-1$ , will make the same choices as TTB and these are therefore all classified as “correct.” To see this, note that the effect of any element having the value -1 will be more than compensated for by any element to its left that has the value 1.

The only vector that does not meet this criterion is the last one (1,-1,-1). Here the effect of the second element is clearly more than compensated for by the first; however, this still leaves a -1 as the third element. If the compensatory function

involves strict variants of the restrictions (A2), i.e.,  $\beta_1 < \beta_2 + \beta_3$ ,<sup>15</sup> then the difference vector (1,-1,-1) is classified as “incorrect.” If, however, the compensatory relation (A2) is weak, i.e.,  $\beta_1 \leq \beta_2 + \beta_3$ , then the difference vector (1,-1,-1) may be classified as “incorrect” or “correct”. In the case of equality, i.e.,  $\beta_1 = \beta_2 + \beta_3$ , TTB will be correct, on average, 50% of the time because here we assume that choice under the compensatory strategy is made at random.

As this example illustrates, TTB will perform marginally better under a weak compensatory scheme than under the corresponding strict compensatory scheme. In other words, a compensatory scheme that involves only strict inequalities of the form  $\beta_j < \sum_i \beta_i$ , for any  $i > j, j = 1, \dots, k-1$  provides a lower bound on the performance of TTB.

While in the 3-cue case there are only two different compensatory functions consistent with the definition (A2) (i.e., involving either  $\beta_1 < \beta_2 + \beta_3$  or  $\beta_1 \leq \beta_2 + \beta_3$ ), for the 4- or 5-cue models the classification of compensatory functions is more complex. However, conditional on the classification (see below), examination of the 4- and 5-cue models is carried out in the same manner.

First, we identify all the dominance pairs. Second, using the fact that  $\beta_j > \beta_{j+1}$  for  $j = 1, \dots, k-1$ , we isolate those difference vectors that unambiguously lead to correct predictions for the TTB model. Third, we allocate the remaining difference vectors into “correct”, “incorrect” and “uncertain” categories in accordance with the inequalities that characterize particular compensatory strategies.

---

<sup>15</sup> Note that – using the expressions (1) and (2) – the only relation between weights that needs to be considered to classify a 3-cue weighting function as compensatory or not, is that between  $\beta_1$  and  $\beta_2 + \beta_3$ .

Consider the classification of the space of compensatory and non-compensatory functions for the 4-cue case. Consistent with what is stated above, we simplify our task by only considering strict compensatory schemes. (This reduces the number of possible schemes but, as noted above, provides a more conservative estimate of TTB's performance.) Our classification of the different models is illustrated by the tree diagram in Figure 1 in the main text.

To initiate the classification, relate the first weight,  $\beta_1$ , to the sum of the next two, i.e.,  $\beta_2 + \beta_3$ , and observe that there are two cases,  $\beta_1 < \beta_2 + \beta_3$  and  $\beta_1 > \beta_2 + \beta_3$ . For the former, note that  $\beta_1 < \beta_2 + \beta_3$  implies that  $\beta_1 < \beta_2 + \beta_3 + \beta_4$  such that when this condition is subdivided, we are left with two possible cases denoted CF5 and CF4. Thus, CF4 is characterized by  $\beta_1 < \beta_2 + \beta_3$  and  $\beta_2 < \beta_3 + \beta_4$  and CF5 by  $\beta_1 < \beta_2 + \beta_3$  and  $\beta_2 > \beta_3 + \beta_4$ . (Recall also that  $\beta_j > \beta_{j+1}$  for  $j = 1, \dots, k-1$ .)

The right hand side of Figure 1 deals with  $\beta_1 > \beta_2 + \beta_3$ . This is accompanied by two other conditions ( $\beta_1 < \beta_2 + \beta_3 + \beta_4$  and  $\beta_1 > \beta_2 + \beta_3 + \beta_4$ )<sup>16</sup> which, in turn, can each have two further conditions ( $\beta_2 < \beta_3 + \beta_4$  and  $\beta_2 > \beta_3 + \beta_4$ ), thereby resulting in four additional models. Of these models, three are compensatory (CF3, CF2, and CF1) and one meets the non-compensatory condition specified in (A1) above. The full classification of strict compensatory strategies therefore includes five different sets of inequalities characterizing between-weights relations (CF1-CF5). These are also characterized in tabular form in the lower part of Figure 1.

Now, for each compensatory strategy depicted in Figure 1, we classify the ambiguous difference vectors as “correct”, “incorrect” or “uncertain”, making use of the inequalities that characterize each compensatory strategy. The resulting

---

<sup>16</sup> For more-than-4-cue cases, the tree should be extended at this point by relating  $\beta_1$  and further sums of the next weights, i.e. for the 5-cue case we should also relate  $\beta_1$  and  $\beta_2 + \beta_3 + \beta_4 + \beta_5$ , for the 6-cue case – additionally  $\beta_1$  and  $\beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6$ , and so on.

classifications of the difference profiles for each of the five compensatory strategies are shown in Table A1.

-----  
Insert Table A1 about here  
-----

First, note that there are 10 ambiguous difference-vector profiles, i.e., cases that do not involve dominance and cannot be classified as “correct” or “incorrect” in accordance with the constraint that  $\beta_j > \beta_{j+1}$  for  $j = 1, \dots, k-1$ . Given that, in the 4-cue case, there are 120 possible pairings, this means that, when the population consists of all distinctive pairings, the lower bound of TTB’s performance cannot fall below 92 % (i.e., 110 of 120).

As can be seen from Table A1, the 4-cue case involves profiles that, for the compensatory functions specified, cannot be classified as “correct” or “incorrect.” These are referred to as “uncertain” and arise because we are now dealing with 4- as opposed to 3-element vectors. To illustrate how this occurs, consider, the compensatory function 4 (CF4) and the difference profile (1,-1,-1,1). Under this weighting function,  $\beta_1$  is smaller than  $(\beta_2 + \beta_3)$ . However, it is not clear whether  $(\beta_1 + \beta_4)$  is greater or smaller than  $(\beta_2 + \beta_3)$ .

For the 5-cue model, we have adopted the same line of reasoning as above for specifying different strict compensatory strategies. See Figure A1 and Table A2. For the 5-cue case, the total number of ambiguous profiles to be classified as either “correct”, “incorrect” or “uncertain” is 66 from a total of 496. This means that the lower bound of TTB’s performance is 87% for a population of distinctive pairings. Indeed, under the “most compensatory” function, analogous to CF5 in the 4-cue case, TTB is expected to make 91 % correct predictions (see Table 6).

-----  
Insert Figure A1 and Table A2  
-----

**Table A1 -- Classification of ambiguous difference vector profiles  
under different compensatory functions: 4-cue case**

<u>Difference profiles, classified differently under different functions</u>					<u>Weighting functions</u>					
<u>x1</u>	<u>x2</u>	<u>x3</u>	<u>x4</u>	<u># cases</u>	<u>Non-CF</u>	<u>CF1</u>	<u>CF2</u>	<u>CF3</u>	<u>CF4</u>	<u>CF5</u>
1	-1	-1	0	2	c	c	c	c	w	w
1	-1	-1	-1	1	c	c	w	w	w	w
0	1	-1	-1	2	c	w	c	w	c	w
1	0	-1	-1	2	c	c	c	c	c	u
1	-1	0	-1	2	c	c	c	c	u	u
1	-1	-1	1	1	c	c	c	c	u	u
Total				<u>10</u>						
Total correct					120	118	119	117	114	110
Total incorrect					0	2	1	3	3	5
Total uncertain					0	0	0	0	3	5

Legend:

c = correct

w = incorrect

u = uncertain

**Table A2 -- Classification of ambiguous difference vector profiles under different compensatory functions: 5-cue case**

Difference profiles.

classified differently under different functions

x1	x2	x3	x4	x5	# cases	Non-CF	CF1	CF2	CF3	CF4	CF5	CF6	CF7	CF8	CF9	CF10	CF11	CF12	CF13	CF14	CF15	CF16	CF17	CF18	CF19	CF20	CF21	CF22	CF23	
1	0	-1	-1	-1	2	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	c	c	u	u	u	u	
1	-1	0	-1	-1	2	c	c	c	c	c	c	c	c	c	c	c	c	c	u	c	u	c	u	u	w	u	w	u	w	
1	-1	-1	0	0	4	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	w	w	w	w	w	w	
1	-1	-1	0	-1	2	c	c	c	c	c	c	c	c	c	c	c	c	u	u	u	u	u	u	w	w	w	w	w	w	
1	-1	-1	-1	0	2	c	c	c	c	c	c	c	c	c	c	c	c	w	w	w	w	w	w	w	w	w	w	w	w	
1	-1	-1	-1	-1	1	c	c	c	c	c	c	w	w	w	w	w	w	w	w	w	w	w	w	w	w	w	w	w	w	
0	1	-1	-1	0	4	c	c	c	c	w	w	c	c	c	c	w	w	c	c	c	c	w	w	c	c	c	c	w	w	
0	1	-1	-1	-1	2	c	c	w	w	w	w	c	c	w	w	w	w	c	c	w	w	w	w	c	c	w	w	w	w	
0	0	1	-1	-1	4	c	w	c	w	c	w	c	w	c	w	c	w	c	w	c	w	c	w	c	w	c	w	c	w	
1	1	-1	-1	-1	1	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	c
1	0	0	-1	-1	4	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u
1	0	-1	0	-1	4	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	u
1	0	-1	-1	1	2	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	u
1	0	-1	-1	0	4	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	u
1	-1	1	-1	-1	1	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	c	u	c	u
1	-1	0	0	-1	4	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	u	u	u	u	u
1	-1	0	-1	1	2	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	u	u	u	u	u
1	-1	0	-1	0	4	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	u	u	u	u	u
1	-1	-1	1	1	1	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	c	u	c	u	c
1	-1	-1	1	0	2	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	u	u	u	u	u
1	-1	-1	1	-1	1	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	u	u	u	u	u
1	-1	-1	0	1	2	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	u	u	u	u	u	u
1	-1	-1	-1	1	1	c	c	c	c	c	c	c	c	c	c	c	c	u	u	u	u	u	u	u	u	u	u	u	u	u
0	1	0	-1	-1	4	c	c	c	c	c	u	u	c	c	c	c	u	c	c	c	c	c	u	c	c	c	c	c	u	u
0	1	-1	0	-1	4	c	c	c	c	u	u	c	c	c	c	u	u	c	c	c	c	u	u	c	c	c	c	u	u	
0	1	-1	-1	1	2	c	c	c	c	u	u	c	c	c	c	u	u	c	c	c	c	u	u	c	c	c	c	u	u	
<b>Total</b>					<b>66</b>																									

Total correct	496	492	494	490	484	476	495	491	493	489	483	475	490	484	488	482	478	466	468	464	464	460	443	432
Total incorrect	0	4	2	6	6	10	1	5	3	7	7	11	3	7	5	9	9	13	9	15	11	17	15	21
Total uncertain	0	0	0	0	6	10	0	0	0	0	6	10	3	5	3	5	9	17	19	17	21	19	38	43

Legend:  
c = correct  
w = incorrect  
u = uncertain

**Figure A1 -- Classification of compensatory (CF) and non-compensatory functions (NonCF) for the 5-cue case .**

