Title
A Test of the Decision-Time Predictions of the `Take the Best' Model

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 30(30)

ISSN
1069-7977

Authors
Schultz, Benjamin
Navarro, Daniel J.

Publication Date
2008

Peer reviewed

A Test of the Decision-Time Predictions of the 'Take the Best' Model

Benjamin Schultz (benjamin.schultz@adelaide.edu.au)
School of Psychology, University of Adelaide, SA 5005, Australia

Daniel J. Navarro (daniel.navarro@adelaide.edu.au)
School of Psychology, University of Adelaide, SA 5005, Australia

Abstract

The "take the best" model of decision making proposes that people make decisions by sequentially searching amongst cues for one that best discriminates between the options being assessed. The search process starts with the best cue and proceeds in descending order of cue validity until one is found that differentiates between the options. It follows, therefore, that the more cues a person is required to use, the longer it will take to make a decision. This study explored the relationship between response time and the number of cues needed to answer a binary choice question correctly. Participants were asked a series of questions about mammals and their response times were recorded. Results support the hypothesis that response time increases as the number of cues required increases. This gives further evidence that a sequential search is occurring during binary-choice decision-making.

Keywords: take the best; decision-making; response time

Introduction

Examining human response times (RT) in simple decision tasks is one of the oldest ideas in experimental psychology, dating back to the analysis of individual differences in RT among astronomers in the 18th and 19th centuries (see Boring, 1950), and in the intervening years a very considerable literature has built up around the topic. The central idea in modeling RT data is that the time course of human decisions can be considered as a form of sequential analysis (e.g., Wald 1947), where people sample data from the environment, and decisions are made once a sufficient amount of evidence has accrued favoring one response over another (e.g., Ratcliff, 1978, Ratcliff & Smith 2004, Vickers 1979, Townsend & Ashby 1983, Luce 1986). In recent years, a number of the same basic ideas have been independently proposed in the context of somewhat more complex decisions. For instance, the "fast and frugal" approach to decision making advocated by Gigerenzer and colleagues (Gigerenzer & Goldstein, 1996; Gigerenzer & Todd, 1999) proposes that higher-level decisions operate along similar lines, with people engaging in a strategic, self-terminating search through memory or the environment (see also Stewart, Chater & Brown 2006).

Although response latency in higher-level tasks would naturally be expected to relate to the number of informative "cues" that must be examined to justify a decision (as per the "take the best" (TTB) model, for instance; Gigerenzer & Goldstein, 1996; Gigerenzer & Todd, 1999), this assumption has not been thoroughly investigated. This is not to say that no relevant studies have been carried out: several studies have employed artificial or learned cues to examine the sequential aspects of decision-making strategies (e.g., Bröder & Gaissmaier, 2007; Bergert & Nosofsky, 2007; Nosofsky & Bergert, 2007), finding evidence supporting TTB's assumption of a sequential search. Similarly, other studies have explored semantic and categorical structure using RT (e.g., Hampton, 1979; Nosofsky & Palmeri, 1997), and looked at the relationships between RT and choice behaviour generally (e.g., Ratcliff & Smith, 2004), but the core proposition behind TTB, that people search through cues sequentially in more cognitively-oriented naturalistic decision tasks has not been directly tested by looking at empirical response times.

In this paper, we present a simple experimental test of this proposition. Using Gigerenzer and Goldstein's (1996) "take the best" model as a canonical example of the class of theories under consideration, we look at the relationship between human RT and the number of cues TTB needs to look up, in the context of making simple semantic decisions about familiar animals. Using data collected by Ruts et al. (2004) as a proxy for the cues people might use, we find a highly consistent pattern of weak but significant correlations, suggesting that the basic notion of sequential sampling common in basic psychophysical decision tasks can be generalized successfully to higher level tasks.

Take the Best: A Fast and Frugal Decision Model

TTB is a decision strategy proposed to account for how people make choices about real-world stimuli, making use of the knowledge available to people in an fast and economical manner (Gigerenzer & Goldstein, 1996; Gigerenzer & Todd, 1999). The basic concept is very simple: when choosing between two alternatives, people search for cues sequentially in order of (some measure of) validity, and make a decision based on the *first* informative cue that differentiates the alternatives. When interpreted in terms of standard "sequential sampling models", this is equivalent to solving a sequential analysis problem with a very high tolerance for errors (e.g., Lee & Cummins 2004). The basic process is illustrated in Figure 1.

Though one of the desirable characteristics of TTB described in Gigerenzer and Goldstein's (1996) original work relates to the fact that it makes quick decisions, the paper makes no attempt to record human response times in decision tasks or examine the relationship with the number of cues that TTB needs to examine before the process terminates. However, as commented on by Bergert and Nosofsky (2007), the model makes quite explicit predictions in this regard: to our

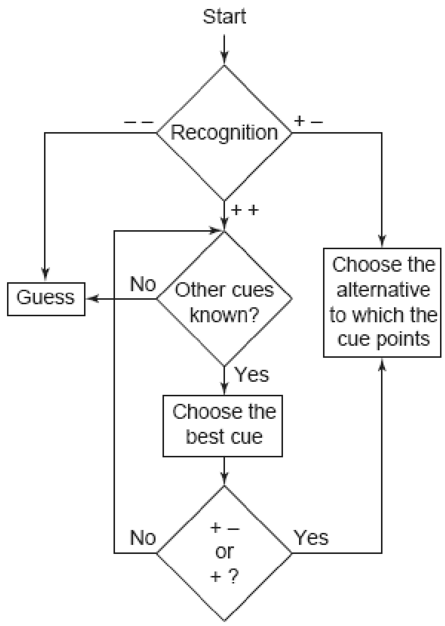

Figure 1: Flow diagram for TTB, adapted from Gigerenzer and Goldstein (1996). If both objects are recognized (always true in our experiment), cues are examined sequentially in order of decreasing validity. The search process stops and a decisions is made as soon as a cue is found upon which the two items differ.

knowledge, the Bergert and Nosofsky (2007) paper is the only one to have looked at RT in this context. However, their work involved artificial stimuli; to fully investigate the relationship between TTB behavior and human RT, it is important to look at human decisions in fairly naturalistic domains, since TTB relies heavily on a theoretical view that holds that human decision making should be evaluated primarily with respect to real world domains.

Experiment

In this experiment, we examine human choice behavior when asked to make decisions regarding the relative heights, weights and speeds of 25 land-based animals. The goal was to determine (a) the extent to which human decision-times correlate with the number of cues that TTB needs to examing to make the corresponding decisions, and (b) the extent to which the TTB model makes the same decisions as human participants.

Method

Participants. Twelve people (8 female, 4 male) participated in the study, all students at the University of Adelaide. Participants ranged in age from 18 to 26 with a mean age of 21.9 years (SD = 2.28), and received either course credit or a $40 book voucher for their participation.

Materials. Data collection proceeded via a computer-based task, in which participants were presented with pairs of animals and asked to decide between them on some attribute (see below). The task was implemented using custom software written in Visual Basic, that recorded participant choices and response times. During the experiment, screen refresh rates were held constant at approximately 16.7ms, and response times were recorded using a high precision timer (though for this experiment the RTs were large enough for this to be a somewhat unnecessary precaution).

Table 1: Animals used as query items in the experiment. The indices for each animal correspond to the column numbers in Figure 2.

1	Monkey	9	Polar Bear	17	Rhinoceros
2	Bison	10	Kangaroo	18	Hippopotamus
3	Camel	11	Cat	19	Elephant
4	Squirrel	12	Cow	20	Horse
5	Donkey	13	Rabbit	21	Tiger
6	Giraffe	14	Llama	22	Pig
7	Deer	15	Lion	23	Fox
8	Dog	16	Mouse	24	Wolf
				25	Zebra

In order to provide a reasonable approximation to the semantic structure involved when making decisions about animals, a cue matrix for the 25 land-dwelling mammals listed in Table 1 was constructed using data taken from a study by Ruts et al. (2004). In that paper, 640 participants were given the name of an object and asked to list 10 features (which could be perceptual, functional, or ad hoc characteristics). Each feature was then assigned a rating between 0 and 3 to indicate the frequency with which it was used to describe an object. Since TTB operates on binary attributes, we treated a rating of 0 in the original study as equivalent to "cue not present" and a rating of 3 as indicating "cue present". For the intermediate values (1-2) we took a pragmatic approach. When the feature was inherently subjective (e.g., "is big") we treated these cases as "feature unknown". However, when these values occurred with respect to easily verified objective characteristics (e.g., "has a tail") the cases were treated as "present" or "absent" depending on the true state of affairs.

The structure of the resulting 181×25 cue matrix is shown in Figure 2. Obviously, given the pragmatic choices made in constructing this cue matrix, it cannot be treated as a literal description of people's mental representations of animals, but it does have the desirable characteristic that it explicitly derived from human judgments, suggesting that some relationship should exist. The cue matrix was augmented with three continuous-valued "target attributes" namely fastest observed speed, largest observed height, and largest observed weight for each of the animals, estimated using online databases.

Procedure. Before the trials began, participants were asked to rate the familiarity of the items on a 10-point scale ranging from "not at all familiar" to "very familiar", since familiarity should also be expected to be mediated by the cue matrix. However, since TTB does not make explicit predictions about familiarity, we omit the analysis of these data in this paper. In any case, this served both as an auxiliary data collection phase, and as a method of familiarizing the participants with the set of stimuli to be used later in the experiment.

After the familiarization phase was complete, participants then proceeded to the experimental trials. On any given trial participants were shown two animals in either picture form or simply presented with the animal names, and asked one of three possible questions:

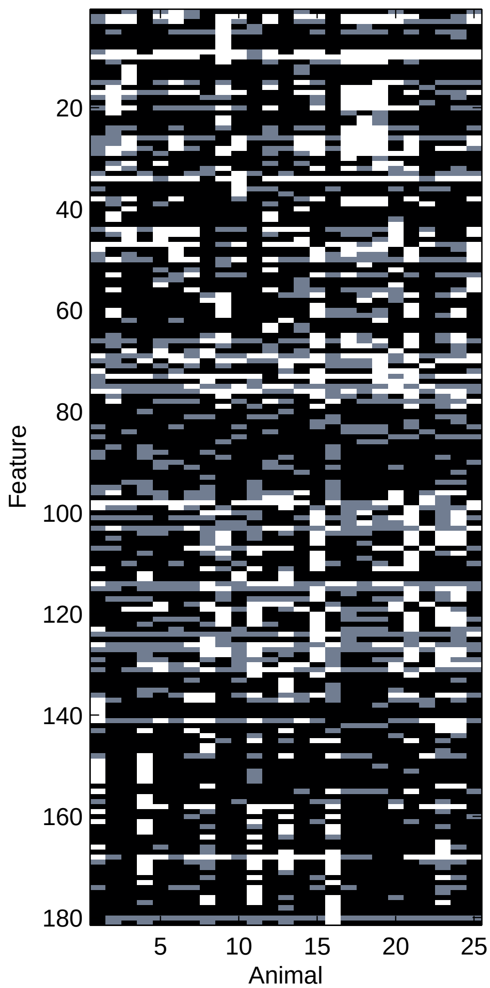

Figure 2: Animal by feature matrix, sorted by validity for the "height" question. Highly valid cues are at the top. White cells indicate feature present (+), black cells indicate feature not present (-), and grey cells indicate that the feature value is unknown.

When travelling at their fastest, which of the following two animals is faster?

At their tallest, which of the following two animals is taller?

At their heaviest, which of the following two animals weighs more?

There was a 3000ms gap between trials in which participants could read the question associated with the stimuli. Once this period elapsed, two animals appeared on screen (in either text or picture form), and the participant was required to select the animal they thought was faster, taller, or heavier "as quickly and as accurately as possible". These instructions were chosen in order to impose an explicit speed-accuracy tradeoff upon the decision maker, typical in studies of response time (see Luce 1986). Participants gave their answers by pressing a key on either the left (using the Z key on a standard Australian keyboard) or right (using the M key) side of the keyboard. This procedure was adopted in order to avoid the concern that mouse-operating skill would contribute substantially to the variability in RT. Moreover, to control for possible differences in speed associated with handedness, the location of each item (i.e., either left or right) was randomized across trials. Finally, the order of questions and of pairs of animals was randomized; however, the formats (i.e., picture or text) occurred in blocks, with the text version always appearing first.

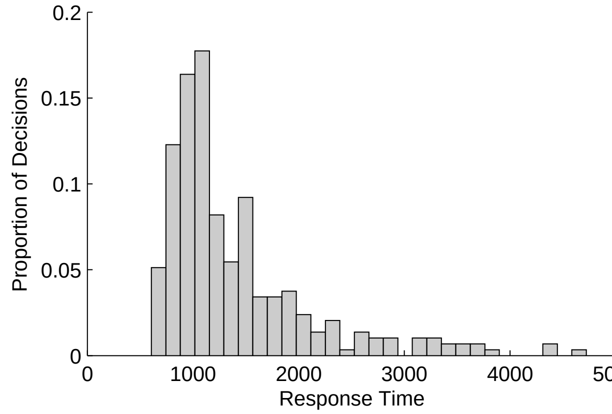

Figure 3: Probability distribution over response times for participant 7 in the context of "height" decisions presented pictorially. As is generally the case for response time distributions, the data are unimodal and positively skewed.

In total, each participant provided 1800 judgments (300 pairs $\times$ 3 question types $\times$ 2 formats), in addition to the familiarization trials. Due to the large number of trials involved, the textual and pictorial tasks were both split into two blocks (450 trials each) with 15 minute breaks between each, in order to reduce fatigue, eye-strain, and boredom. Note that the task involved no supervised learning: participants were not given feedback as to which animals were bigger, faster or heavier (though the anwers were available at the end of the experiment if participants were curious), with questions relying instead on the assumed general knowledge of the participants. By relying on pre-existing knowledge about real world things, the approach is considerably more naturalistic than earlier studies (e.g., Lee and Cummins, 2004; Bergert and Nosofsky, 2007), but as previously noted relies on the assumption that our participants made use of representational structures not too dissimilar from the cue matrix collected by Ruts et al. (2004). This choice was deliberate, since the whole concept of fast and frugal models is built around the "ecological rationality" hypothesis, that people's decision processes should be expected to look sensible only in fairly naturalistic domains.

Results

Descriptive statistics. Of the 12 participants within the study, only 11 completed both the pictorial and textual stimuli form conditions. The other participant completed only the textual stimuli condition. Table 3 shows the means, standard deviations, and skewness for all six conditions. Response

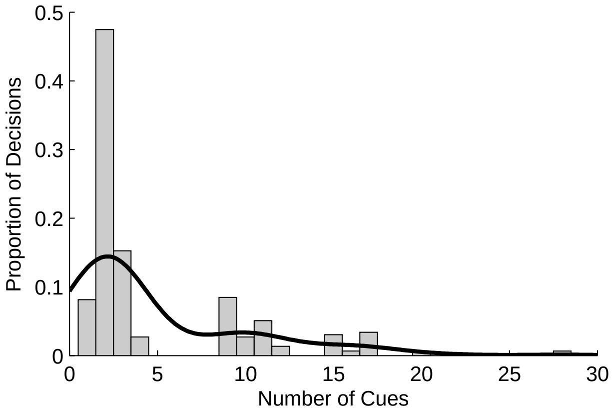

Figure 4: Probability distribution over number of cues looked up by TTB in the context of "height" decisions, and a smoothed estimate of this distribution (estimated by standard nonparametric kernel density estimation methods). The similarity to the empirical RT distribution in Figure 3 is not surprising, but failure to observe this would represent a failure of the TTB model.

times varied considerably, as one might suspect, and tended to be longer than is typically the case for simple perceptual forced choice tasks (e.g., Luce 1986) since some deliberation is required, but were never longer than than 10s. As is often the case, RTs for all conditions were positive skewed, as illustrated in Figure 3. Preceding the correlational analysis, data was screened and it was deemed necessary to treat response times over 5000ms and numbers of cues required over 30, as outliers. Thus these were removed from analysis.

TTB decision processes. Cue validities were calculated from the matrix in Figure 2 using the Bayesian approach discussed by Lee, Chandrasena, & Navarro (2002), based on a simple beta-binomial model. Each cue is able to discriminate between some number of object pairs, but this number is quite variable. Accordingly, it is important to note that there is much more uncertainty about the usefulness of a cue that is able to make only a few decisions, when compared to a cue that is able to make many decisions. If a cue makes g good decisions and b bad decisions if it is relied on exclusively, then the expected validity v according to a beta-binomial model is simply

$$v = \frac{g+1}{g+b+2}.$$

The Bayesian aspect to this approach arises from the $1/2$ term, which results when no data are available for that cue (i.e., $g = b = 0$), and represents a prior belief about the probability that a cue will be useful. The use of this prior imposes a "regression to the mean" effect, which Lee et al. (2002) found to be extremely important in natural domains (in that case, e-mail classification). Obviously, a different set of cue validities was calculated for each of question type, since people would be expected to search through memory in a different fashion when asked about speeds than when asked about heights.

Having found these validities, we then calculated the number of cues TTB needed to look up for all possible decisions about heights, weights and speeds of animals, as well as the decision that TTB predicts in each case. The main variable of interest is the number of cues examined: for height, the number of cues ranged from 1 to 70, with mean 5.71 and standard deviation 8.11. For questions about weight, the average number of cues needed is 6.93 (standard deviation 8.18, ranging from 1 to 61), while questions about speed could require up to 53 cues, with an average of 10.34 (standard deviation 8.42). As one would expect, all three distributions were positively skewed, with height being the most skewed (skew = 4.20), followed by weight (skew = 3.68), then speed (skew = 1.04). The distribution over the number of cues required by TTB when judging heights is shown in Figure 4. The solid black line shows a smoothed version of the distribution. The fact that the distribution is highly skewed is not surprising since it represents the distribution of a minimum statistic (in statistics, extreme-value distributions are always skewed), but it is nevertheless a desirable characteristic since empirical response time distributions are invariably skewed as well (e.g., Luce, 1986).

Agreement with human decisions. Before considering the relationship between human RT and the number of cues TTB examines in any detail, some preliminary checks are in order. In particular, an important check is to see if a TTB procedure based on the cue structure shown in Figure 2 is broadly in agreement with human decisions in this task. To do so, we looked to see if TTB makes errors on the same decisions that human participants do. The probability with which TTB made the same decisions as the human participants across the various conditions is shown in Table 2. Those agreement rates that are significantly higher than would be expected by chance (i.e., if TTB decisions were unrelated to human decisions) are shown in bold.[1] On the whole, TTB choices agree with human choices significantly more frequently than one would expect by chance.

Correlations with human decision times. Since response time and the TTB cue number distributions are both highly non-normal, and that the relationship between them may not be linear, Spearman rank-order correlations were used to measure the strength of the relationship between the two. These correlations are shown in Figure 4: of the 69 correlation coefficients estimated, a total of 40 were significant at the .05 level. This pattern is exceedingly unlikely to represent a chance relationship. Moreover, of the 29 correlations that do not reach significance in their own right, 26 trend in the correct direction (i.e. only 3 of the 69 rank-order correlation coefficients are negative). While none of the effects are particularly large (the interquartile range on ρ runs from .08 to .23), they are nearly always consistent with the TTB model, which is quite remarkable: given that TTB relies on a cue matrix produced by Flemish-speaking participants living in Belgium, the ability to successfully predict the decisions of English-speaking Australian participants is moderately im-

[1]If TTB is correct with probability θ_t and humans are correct with probability θ_h, then expected rate of agreement by chance is simply $\phi = \theta_t\theta_h + (1-\theta_t)(1-\theta_h)$. The tests we conducted looked to see whether the number of agreements between TTB and the human data could be plausibly be treated as a sample from a binomial distribution with rate parameter ϕ. This is somewhat oversimplistic since it ignores uncertainty about the value of ϕ, but with 69 tests at sample size 300 each, this seems a minor issue.

Table 2: Probability of agreement between TTB decisions and human decisions. The values range from 52% agreement and 89% agreement. More importantly, in 53 of the 69 cases (indicated in bold), the probability of agreement is higher than would be expected by chance (i.e., significant at $p < .05$).

ID#	PW	TW	PS	TS	PH	TH
1	**.89**	**.88**	**.65**	**.63**	**.79**	**.80**
2	**.85**	**.82**	.62	.59	**.83**	**.79**
3	**.82**	**.83**	.58	**.63**	**.77**	**.81**
4	**.88**	**.87**	**.65**	**.64**	**.79**	**.82**
5	**.84**	**.84**	.53	.58	**.78**	**.80**
6	.69	.72	.54	.52	.57	**.70**
7	**.89**	**.87**	**.64**	**.65**	**.83**	**.84**
8	**.79**	**.84**	.60	**.63**	**.79**	**.83**
9	**.76**	**.83**	.55	.57	**.77**	**.80**
10	**.77**	**.77**	.58	.61	**.78**	**.76**
11	**.88**	**.86**	.58	**.62**	**.82**	**.81**
12	-	**.88**	-	**.66**	-	**.83**

Table 3: Response time (ms) means, standard deviations, and skewness for pictorial (P) and text-based (T) stimuli, across the three different question types; height (H), weight (W), and speed (S).

Format	Question	Mean	Std Dev.	Skewness
P	W	1085	671	4.64
T	W	1475	834	3.09
P	S	1246	824	3.49
T	S	1698	1015	2.89
P	H	1094	669	4.34
T	H	1480	764	2.77

pressive. To provide a sense of these relationships, Figure 5 plots the number of cues that TTB looks up for each decision against all 1800 response times made by participant 7 (for whom the relationships are strongest), along with the best fitting linear function.

Discussion

The consistent positive correlations between response time and number of cues TTB examines found in this study, combined with the above-chance rates of agreement between TTB choices and human choices, provide support to the notion that a sequential search is occurring in binary choice decision-making. Our approach has a number of advantages over other experimental designs. By examining participants' choices and decision-times across several conditions in a real-world domain (animals), we are able to ensure that the effects are not artifacts of averaging or the use of artificial stimuli. However, there are some equally-important shortcomings that should be noted. Most notably, as a consequence of the choice to use a pre-existing naturalistic domain, the actual cues used by people are unknown. The reliance on the Ruts et al. (2004) data is intuitively plausible, and in some regards the fact that we observed the expected correlations suggests that it is reasonable to do so, but it would be desirable to investigate the latent mental representations used by our participants. Additionally, a more complete analysis would use the cue matrix to make predictions not only about response times, but also a range of other variables that relate to people's beliefs, such as recognition, familiarity and typicality. On a more minor note, extending the work to a wider variety of domains would be useful. Overall, however, the results are fairly encouraging.

Table 4: Spearman rank-order correlations for all participants and for all six conditions. Conditions are listed as pictorial (P) or textual (T), and by the question asked, namely height (H), weight (W), or speed (S). Bold entries indicate significant correlations at the .05 level.

ID#	PW	TW	PS	TS	PH	TH
1	**.21**	**.12**	**.14**	**.12**	**.15**	**.33**
2	**.14**	**.11**	-.05	.09	**.24**	**.20**
3	**.23**	.09	.11	.06	**.13**	.07
4	**.31**	**.16**	**.28**	**.23**	**.42**	**.31**
5	**.22**	**.26**	.08	.10	**.32**	.09
6	.06	**.21**	.07	.06	**.14**	.03
7	**.23**	**.24**	**.13**	**.28**	**.39**	**.24**
8	.10	**.16**	-.00	.09	.10	.11
9	.05	**.16**	.08	-.06	**.15**	.04
10	.04	**.15**	.01	.07	**.14**	.11
11	**.18**	**.23**	**.16**	**.21**	**.35**	**.15**
12	-	.09	-	**.23**	-	**.24**

Viewed from a broader perspective, our suspicion is that the results should be interpreted not so much as strong evidence for TTB, but rather as support for a more general class of sequential sampling models. That is, while our findings are consistent with the TTB model, they are also very likely to be consistent with a wide variety of sequential search procedures. For example, within the approach discussed by Lee and Cummins (2004) which treats TTB as a special case of a sequential sampling model inspired by Ratcliff (1978) and Vickers (1979), search terminates after a limited (but variable) number of informative cues are found instead of just one. While we have not considered richer models such Lee and Cummins' in this paper, a natural extension of this work would be to do exactly this.

In light of this, we suggest that the basic idea that people make high-level decisions by sampling information from the environment or from memory requires a broader examination. In related work, we have considered the idea that people explicitly sample hypotheses in categorization (Navarro 2007; Sanborn, Griffiths & Navarro 2006) and looked at whether sampling from memory mediates basic decision-making biases (Bruza, Welsh & Navarro, submitted), but as the literature stands at present, it is difficult to say with any certainty that the sorts of sequential sampling processes common in simple decisions (Ratcliff 1978, Vickers 1979) are replicated in higher-level cognitive processing. By explicitly correlating decision-times with simple sequential search processes defined as "known" semantic structures, we are able to take some steps in the direction of verifying this hypothesis.

Acknowledgements. We thank Michael Lee for the TTB code; Peter Hughes for the experimental software; Nancy

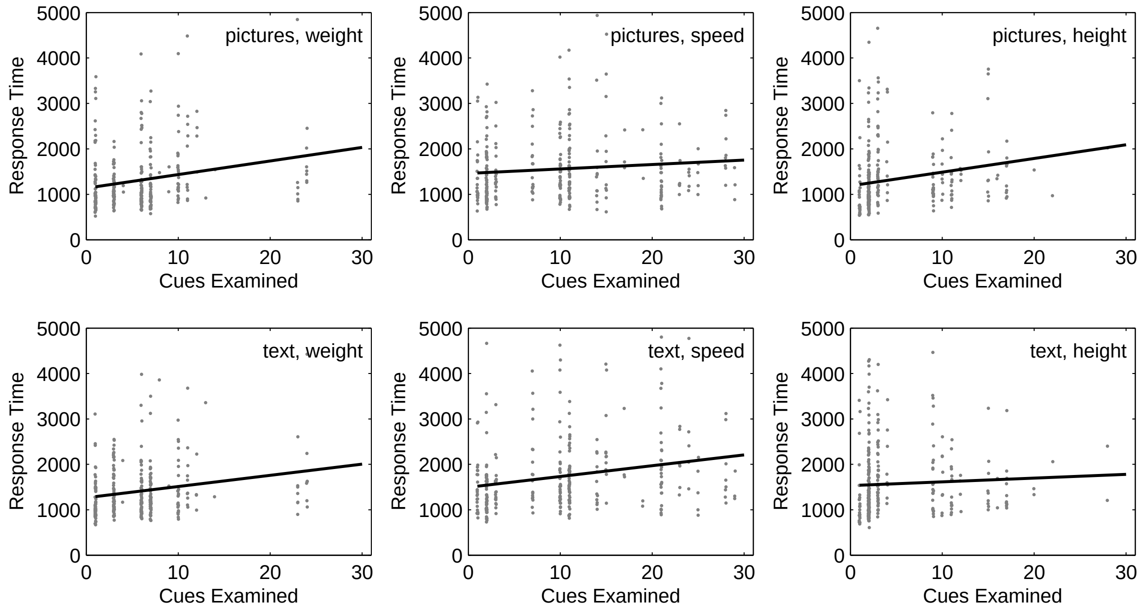

Figure 5: Response times for participant 7 plotted as a function of the number of cues TTB looks up in the Ruts et al. matrix, for all conditions. All six conditions show a weak positive correlation.

Briggs and Matthew Welsh for helpful discussions; and the reviewers for their comments. DJN was supported by an Australian Research Fellowship (ARC grant DP0773794).

References

Bergert, F. B. & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and Take-the-Best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 107-129.

Boring, E. G. (1950). *A History of Experimental Psychology* (2nd ed). New York: Appleton-Century-Crofts

Bröder, A. & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multiattribute decisions. *Psychonomic Bulletin & Review, 14*, 895-900.

Bruza, B. Welsh, M. B. & Navarro, D. J. (in press). Does memory mediate susceptibility to cognitive biases? Implications of the decision-by-sampling theory. In V. Sloutsky, B. Love, & K. McRae (Eds.) *Proceedings of the 30th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650-669.

Gigerenzer, G., & Todd, P. M. (1999). *Simple Heuristics That Make Us Smart.* New York: Oxford University Press.

Goldstein, D. G. & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, & P. M. Todd, (Eds.). *Simple Heuristics That Make Us Smart*. Oxford: Oxford University Press.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning & Verbal Behavior, 18*, 441-461.

Lee, M. D., Chandrasena, L. & Navarro, D. J. (2002). Using cognitive decision models to prioritize E-mails. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society*. (pp. 478-483). Mahwah, NJ: Lawrence Erlbaum.

Lee, M. D. & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the take the best and the rational models. *Psychonomic Bulletin & Review 2004, 11*(2), 343-352.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* New York: Oxford University Press; Oxford: Clarendon Press.

Navarro, D. J. (2007). On the interaction between exemplar-based concepts and a response scaling process. *Journal of Mathematical Psychology, 51*, 85-98

Nosofsky, R. M. & Palmeri, T. J. (1997) An exemplar-based random walk model of speeded classification. *Psychological Review. 104*(2), 266-300.

Nosofsky, R. M., & Bergert, F. B. (2007). Limitations of exemplar models of multi-attribute probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 999-1019.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review 85*, 59-108.

Ratcliff, R. & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*(2), 333-367.

Ruts, W., De Deyne, S. Ameel, E., Vanpaemel, W., Verbeemen, T., & Storms, G. (2004). Flemish norm data for 13 natural concepts and 343 exemplars. *Behavior Research Methods, Instruments, and Computers, 36*, 506-515.

Sanborn, A. N., Griffiths, T. L. & Navarro, D. J. (2006). A more rational model of categorization. In R. Sun & N. Miyake (Eds), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 726-731). Mahwah, NJ: Lawrence Erlbaum.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology, 53*, 1-26.

Townsend, J. T. & Ashby, F. G. (1983). *Stochastic Modeling of Elementary Psychological Processes*. Cambridge, UK: Cambridge University Press.

Vickers, D. (1979). *Decision Processes in Visual Perception*. New York: Academic Press.

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