

Magnitude comparison revisited: An alternative approach to binary choice under uncertainty

Norman R. Brown · Shawn Tan

Published online: 17 February 2011
© Psychonomic Society, Inc. 2011

Abstract Two generations of psychologists have been interested in understanding binary choice under uncertainty. In the 1970s and 1980s, researchers assumed that people rely on a two-stage *magnitude comparison* process to make these decisions (Banks, 1977; Moyer & Dumais, 1978). More recently, the focus has shifted to approaches that rely on probabilistic cues and simple heuristics (Gigerenzer & Goldstein, *Psychological Review* 103, 650–669, 1996). Here, we test competing predictions derived from these two very different approaches and conclude that the magnitude comparison process plays a central role in this task. In support of this conclusion, we present an experiment in which participants were timed as they decided which of two vehicles was more expensive. Pairs composed of one luxury vehicle (e.g., BMW 323i) and one nonluxury vehicle (e.g., Toyota 4Runner) were critical because the magnitude comparison approach correctly predicted that reaction times would decrease with *subjective distance*, whereas the heuristics approach incorrectly predicted that there would be no relation.

Keywords Magnitude comparison · Take the best · Binary choice · Paired comparison · Heuristics · Symbolic distance effect

Which is taller: a fox or an elk?
Which has the larger population: Winnipeg or Vancouver?

Which is more expensive: a BMW 323 sedan or a Toyota RAV4 SUV?

Two generations of researchers have been interested in understanding how people answer questions like these. In the 1970s and 1980s, the *paired comparison task* was studied extensively from an information-processing perspective. This research focused on the origins of several robust reaction time (RT) phenomena (e.g., the *symbolic distance effect*; see below) and their representational implications (Banks, 1977; Brown & Siegler, 1991; Hamilton & Sanford, 1978; Kerst & Howard, 1977; Kosslyn, Murphy, Bemesderfer, & Feinstein, 1977; Moyer & Dumais, 1978; Siegler & Robinson, 1982). More recently, Gigerenzer, his colleagues (Gigerenzer & Goldstein, 1996; Hoffrage, Hertwig, & Gigerenzer, 2000; Martignon & Hoffrage, 2002; Persson & Rieskamp, 2009), and their critics (Bröder, 2000; Bröder & Schiffer, 2003; Dougherty, Franco-Watkins, & Thomas, 2008; Newell & Shanks, 2003; Newell, Weston, & Shanks, 2003) have used the *binary choice task* to investigate one-reason decision making and the psychological reality of the ecological analysis underpinning a family of “fast and frugal” heuristics. Although these two research streams have used the identical task, they have been *completely* isolated from one another—“two solitudes,” as it were. Indeed, we know of no case in which the RT studies of the 1970s and 1980s have been cited in the more recent decision-making literature.

This scholarly gulf is paralleled by a theoretical one. At issue is the fundamental nature of the processes people use when they make binary decisions. The information-processing models of the 1970s and 1980s incorporated two central processes. During the first stage, magnitude values are generated or retrieved for each item in the pair.

N. R. Brown (✉) · S. Tan
Department of Psychology, University of Alberta,
Edmonton, AB T6G 2E9, Canada
e-mail: nrbrown@ualberta.ca

During the second stage, these values are compared; when they match, additional information is sought; when they differ, an appropriate response is registered (Banks, 1977; Moyer & Dumais, 1978).

In contrast to this *magnitude comparison* (MaC) approach, the *probabilistic mental models* (PMMs) approach holds that choice is based solely on the values of probabilistic cues and that magnitude information, per se, plays no role in the decision. This account requires that people have access to a set of task-relevant cues, that these cues are ordered according to their ecological validities, and that this database is interrogated in a serial manner from the most predictive cue to the least. On this view, search is terminated and a decision made when a positive cue value is encountered for one member of the pair, but not for the other. This strategy is often referred to as *take-the-best* (TTB), because each decision is based on the single cue that is most likely to point to a correct answer (Gigerenzer & Goldstein, 1996; Gigerenzer, Todd, & The ABC Research Group, 1999; Goldstein & Gigerenzer, 2002).

The present study was undertaken with two goals in mind. The first was to (re)introduce the idea that people often engage a MaC process when they make binary decisions. The second was to demonstrate the utility of process-based measures for selecting between these two very different approaches.

In the service of these goals, we conducted an experiment in which participants were timed as they decided which of two vehicles was more expensive. There were two reasons for using this domain. First, a recent

study has demonstrated that university students do not store automobile prices in memory but can estimate them with a fair degree of accuracy (Murray & Brown, 2009). This indicates that people have a reasonable amount of useful task-relevant knowledge. Second, across three experiments, price estimates were well fit by additive models that made use of just two cues, *brand status* (luxury vs. nonluxury) and *product class* (SUV vs. sedan), and that weighted the former more heavily than the latter. This indicates that people treat brand status as “the best,” or at least the most important, of the two cues. Implications of these facts are considered next.

Predicting decision times and response patterns

In this experiment, participants were timed as they decided which of two vehicles was more expensive. The 16 vehicles used in this study were drawn from four categories identified by crossing brand status (luxury, nonluxury) with product class (SUV, sedan). During the decision task, these vehicles were exhaustively paired with one another. This procedure produced the 10 *pair types* listed in Table 1. The label assigned to each of these types has four parts; the two properties listed to the left of the double colon define one pair member, and the two listed to the right define the other. For example, when a pair involves, say, a Mercedes C-Class Sedan and Toyota 4Runner SUV, it is classified as a luxury-sedan::nonluxury-SUV trial, because the pair is composed of a luxury sedan (the Mercedes) and a non-luxury SUV (the Toyota).

Table 1 Description of pair types used in the present experiment, along with relevant processing values and reaction time (RT) predictions derived from PMM and MaC accounts

Pair-Type Category	Pair Type	PMM Account		MaC Account			
		# Cue Evaluations	RT Predictions	Subjective Magnitude _A	Subjective Magnitude _B	Subjective Difference	RT Predictions
Same-status:: same-class	luxury-SUV::luxury-SUV	3	slowest	4	4	0	slowest
	luxury-sedan::luxury-sedan	3	slowest	3	3	0	slowest
	nonluxury-SUV::nonluxury-SUV	3	slowest	1	1	0	slowest
	nonluxury-sedan::nonluxury-sedan	3	slowest	0	0	0	slowest
Same-status:: different-class	luxury-SUV::luxury-sedan	2	medium	4	3	1	medium
	nonluxury-SUV::nonluxury-sedan	2	medium	1	0	1	medium
Different-status:: same-class	luxury-SUV::nonluxury-SUV	1	fastest	4	1	3	fast
	luxury-sedan::nonluxury-sedan	1	fastest	3	0	3	fast
Different-status:: different-class	luxury-SUV::nonluxury-sedan	1	fastest	4	0	4	fastest
	luxury-sedan::nonluxury-SUV	1	fastest	3	1	2	medium-fast

It was necessary to adopt this level of description because predictions that follow from the MaC account depend on the specific properties of paired vehicles (see below). Table 1 also includes a more general level of description, the *pair-type category*. This level collapses over the specific values of the defining properties. Instead, it reflects the relationship (same or different) that holds across the paired vehicles for the two properties. Thus, four pair types are classified as same-status::same-class; these include all pairs that involve two luxury SUVs, two luxury sedans, two nonluxury SUVs, or two nonluxury sedans. The same-status::different-class category includes the pairs where the brand status is held constant but one pair member is an SUV and the other a sedan. The opposite situation obtains for the different-status::same-class pairs; here, both pair members are either SUVs or sedans, but only one is produced by a luxury manufacturer. Finally, two pair types are classified as different-status::different-class pairs; one consists of a luxury SUV paired with a nonluxury sedan, and the other of a luxury sedan paired with a nonluxury SUV.

PMM predictions

Predictions generated by the PMM account are presented in Table 1. These predictions assume the following: First, people use the TTB heuristic (Gigerenzer & Goldstein, 1996) to select between vehicles; second, brand status is treated as “the best” cue (see above) and, hence, is evaluated first; and third, the process of retrieving and comparing cue values takes time. It follows that decisions requiring a single cue evaluation should be faster than those that require two and those requiring two evaluations should be faster than those that require three (Bröder & Gaissmaier, 2007). Thus, (1) participants should respond fastest to status-different pairs, because they involve just one cue evaluation. Also, (2) RTs for these pairs should be unaffected by product class information, because search terminates once the best cue, brand status, has been evaluated. In other words, when a pair consists of one luxury vehicle (e.g., a BMW) and one nonluxury vehicle (e.g., a Chevrolet), the PMM account predicts that participants will rapidly select the former, because the decision is made on the basis of the first cue (the brand status cue) that is evaluated.

In addition, on this view, same-status::different-class pairs require two evaluations and same-status::same-class pairs require at least three. Thus, (3) same-status::different-class pairs should be slower than different-status pairs but (4) faster than same-status::same-class pairs. Finally, the PMM approach predicts that (5) participants will always select a luxury vehicle over a nonluxury vehicle when

confronted with a different-status pair and that (6) they will always select an SUV over a sedan when confronted with a same-status::different-class pair.

MaC predictions

Table 1 also presents the RT predictions generated from the MaC account. In making these predictions, it was assumed that RTs display a classic *symbolic distance effect*—that is, that RTs decrease as the difference in subjective magnitudes increases. Subjective magnitudes, presented in Table 1, were computed using the following weighted additive model,¹ where brand status is set to 1 for luxury brands and 0 for nonluxury brands and product class is set to 1 for SUVs and 0 for sedans:

$$\text{subjective magnitude}_x = 3(\text{brand status}_x) + 1(\text{product class}_x). \quad (1)$$

To take a concrete example, consider a pair that consists of a luxury SUV (Lexus RX350) and a nonluxury sedan (Ford Taurus). Because the subjective magnitudes are 4 for the former and 0 for the latter, subjective distance equals 4. Similarly, the subjective distance for a luxury-sedan::nonluxury-SUV (e.g., Cadillac CTS vs Toyota 4Runner) pair is 2, because luxury sedans and nonluxury SUVs are assigned magnitudes of 3 and 1, respectively. Given that the subjective distance is greater on luxury-SUV::nonluxury-sedan trials than on luxury-sedan::nonluxury-SUV trials, the MaC account predicts that participants should respond more quickly to the former than to the latter.

MaC makes the following predictions: (1) Same-status::same-class trials should be the slowest trials; same-status::different-class trials should be (2) faster than same-status::same-class trials but (3) slower than different-status trials; (4) luxury-sedan::nonluxury-SUV trials should be the slowest of the different-status trials; (5) luxury-SUV::nonluxury-sedan trials should be the fastest; and (6) two different-status-same-class trial types should fall in between.

The MaC account also predicts that larger subjective distances should be associated with more consistent choice (Banks, 1977). Thus, for status-different pairs, the probability of selecting the luxury vehicle on different-status

¹ This 3-to-1 weighting of brand status relative to product class closely reflects the ranked price estimates collected in the present study (see below). Although the exact value of this ratio varies somewhat across experiments, materials, and statistical approaches (Murray & Brown, 2009), it is always true that brand status is weighted more heavily than product class.

trials should be highest for luxury-SUV::nonluxury-sedan pairs and lowest for luxury-sedan::nonluxury-SUV pairs. Similarly, participants should be less consistent in their choice of SUVs when responding to same-status::different-class pairs than they are in their choice luxury of vehicles when responding to different-status pairs.

Competing predictions

As is apparent from Table 1, MaC and PMM predictions differ for different-status pairs but not for same-status pairs. Again, for reasons developed above, PMM predicts that RTs should be equally short across different-status trial types and that luxury vehicles should be selected when there is a status difference. In contrast, MaC predicts that luxury-sedan::nonluxury-SUV pairs will be the slowest of the different-status trial types and that luxury-SUV::nonluxury-sedan pairs will be the fastest. It also predicts that participants should be least likely to select a luxury vehicle when responding to luxury-sedan::nonluxury-SUV pairs and most likely to do so when responding to luxury-SUV::nonluxury-sedan pairs.

Method

Participants

Twenty-eight undergraduates took part in this experiment; they received course credit for their cooperation.

Materials

Sixteen vehicle names were used in this experiment. Four referred to luxury SUVs (BMW X3, Cadillac SRX, Lexus RX350, Mercedes B-Class), four to luxury sedans (BMW 323i, Cadillac CTS, Lexus IS250, Mercedes C-Class), four to nonluxury SUVs (Chevrolet Suburban, Ford Expedition MAX, Subaru Tribeca, Toyota 4Runner), and four to nonluxury sedans (Chevrolet Impala, Ford Taurus, Subaru Legacy, Toyota Avalon).

Procedure

Each session consisted of three tasks: a familiarity task, a binary decision task, and an estimation task. During task 1, participants indicated how familiar they were with the target vehicles;² during task 3, they estimated the current base price of each.

² This task was included to ensure that participants had been exposed to each target name prior to the paired-comparison task.

During task 2, participants were presented with one vehicle name, along with its product class (e.g., Lexus RX350 SUV) on the left side of the computer screen and with another on the right. They were instructed to decide which of the two was more expensive (i.e., had a higher base price) and to indicate their belief, as quickly and accurately as possible, by pressing either the “C” key or the “M” key. The 16 target vehicles were exhaustively paired with one another, resulting in 120 trial pairs.

A 20-trial practice block preceded task 2. Participants were presented with either a C or an M and were required to press the corresponding key on their keyboards. No feedback was given during any of the tasks. In all tasks, materials were presented on a video display, and responses were collected from the computer keyboard; presentation order and left/right screen position (task 2) were randomized separately for each participant.

Results

Price estimates

MaC focuses on relative magnitude differences, rather than absolute ones. For this reason, each participant’s price estimates were ranked from smallest to largest and then averaged across the four vehicles in a given category. Figure 1 presents means of these averages. Consistent with prior research (Murray & Brown, 2009), the price estimates reflected an additive combination of brand status, $F(1, 27) = 242.20$, $MSE = 3.90$, $p < .001$, $\eta_p^2 = .90$, and product class, $F(1, 27) = 20.38$, $MSE = 5.17$, $p < .001$, $\eta_p^2 = .43$ [status \times class interaction, $F(1, 27) = 1.85$, $MSE = 2.77$, $p > .10$], and brand status ($\beta = 5.86$) was weighted more heavily than product class ($\beta = 2.06$).

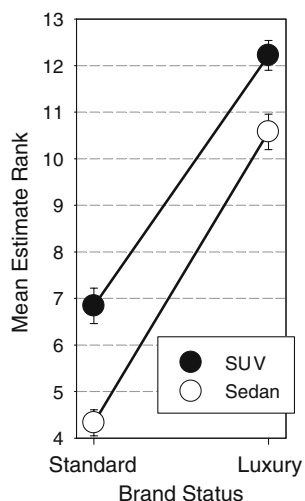
Comparison times

Ten comparison time means, one for each trial type, were computed for each participant. The average of these means is plotted against predicted number of PMM evaluations (Fig. 2a) and against average absolute rank difference³ (Fig. 2b).

Recall that the PMM and MaC accounts make competing predictions only for the different-status pairs. Data relevant to these predictions appear (circled) at the lower right in Fig. 2a and b. As this figure makes clear, there were reliable differences between the different-status trials types, $F(3, 81) = 10.30$, $MSE = 70,907$, $p < .001$, $\eta_p^2 = .28$, and

³ An absolute rank difference value was computed for each of 120 vehicle pairs—that is, rank difference = $|\text{rank}_A - \text{rank}_B|$. Ten trial type means were computed from these values for each participant; the averages of these means are represented along the abscissa in Fig. 2b.

Fig. 1 Mean ranked price estimates (and standard errors) for the four vehicle categories



these differences, which reflected subjective distance between paired vehicles, were consistent with the MaC predictions. Specifically, a series of planned contrasts indicated that luxury-sedan::nonluxury-SUV trials (1.98 s) were reliably slower than luxury-sedan:: nonluxury-sedan trials (1.76 s), $F(1, 27) = 7.09$, $MSE = 178,856$, $p < .05$, $\eta_p^2 = .21$; that luxury-SUV::nonluxury-SUV trials (1.76 s) were reliably slower than luxury-SUV::nonluxury-sedan trials (1.58 s), $F(1, 27) = 8.56$, $MSE = 108,395$, $p < .01$, $\eta_p^2 = .24$; and that luxury-SUV::nonluxury-SUV trials and luxury-sedan::nonluxury-sedan trials were equally fast, $F < 1.0$.

Choice

The MaC account, not the PMM account, also predicted that choice data should reflect subjective distance. In contrast, the PMM account predicted that participants would invariably select the luxury vehicle when presented with different-status pairs and the SUV when presented with same-status::different-class pairs. To assess these predictions, we computed, for the six relevant pair types, the percentage of each participant's responses that were in line with the PMM account.

For the four different-status pairs, the overall choice pattern paralleled the RT pattern, $F(3, 81) = 9.86$, $MSE = 0.05$, $p < .001$, $\eta_p^2 = .27$.⁴ For these pair types, participants were least likely to select the luxury vehicle when selecting between a luxury sedan and a nonluxury SUV ($M = 83\%$, $SE = 4.6$), $F(1, 27) = 12.48$, $MSE = 0.12$, $p < .01$, $\eta_p^2 = .32$, and most likely to select the luxury vehicle when choosing between a luxury SUV and a nonluxury sedan ($M = 96\%$, $SE = 1.6$), $F(1, 27) = 4.59$, $MSE = 0.05$, $p < .05$, $\eta_p^2 = .14$. Participants selected luxury vehicles 94%

($SE = 1.8$) of the time when responding to luxury-SUV:: nonluxury-SUV pairs and 93% ($SE = 2.4$) of the time when responding to luxury-sedan:: nonluxury-sedan pairs.

Overall, participants made fewer responses that were in line with PMM predictions when responding to same-status:: different-class pairs ($M = 76\%$) than when responding to different-status pairs ($M = 92\%$), $F(1, 27) = 16.70$, $MSE = 0.09$, $p < .001$, $\eta_p^2 = .38$. On average, SUVs were selected 76% ($SE = 1.6$) of the time on luxury-SUV::luxury-sedan trials and 75% ($SE = 1.6$) of the time on nonluxury-SUV:: nonluxury-sedan trials.

Strategy differences

When strategies are well defined and their behavioral consequences well understood, it is possible to assign individuals to groups on the basis of their performance and, in so doing, to determine the relative prevalence of the strategies under investigation (Bröder, 2000; Bröder & Schiffer, 2003; Persson & Rieskamp, 2009; Siegler, 1987; Siegler & Robinson, 1982). To this end, we used the choice data to identify individuals who appear to have relied on TTB and then examined RT data from this TTB group, along with RT and choice data provided by the remaining participants.

Again, the TTB approach predicts that participants should consistently choose luxury vehicles when confronted with different-status pairs and SUVs when confronted with same-status::different-class pairs. Only 1 of 28 participants exactly conformed to this pattern. However, 8 participants selected luxury vehicles on all different-status trials. Consistent with the notion that participants in this group might have used the status cue in a noncompensatory manner, different-status RTs did not differ reliably as function of trial type, $F < 1.0$. For this group, mean RTs were 1.72 s ($SE = 0.15$) for luxury-sedan:: nonluxury-SUV trials, 1.55 s ($SE = 0.16$) for luxury-SUV::nonluxury-sedan, 1.57 s ($SE = 0.19$) for the luxury-SUV::nonluxury-SUV trials, and 1.60 s ($SE = 0.12$) for luxury-sedan::nonluxury-sedan trials.

In contrast, RTs of the other 20 participants were strongly affected by trial type, $F(3, 57) = 11.20$, $MSE = 69,140$, $p < .001$, $\eta_p^2 = .37$. Consistent with the MaC predictions, RTs were shortest for luxury-SUV::nonluxury-sedan trials ($M = 1.59$ s, $SE = 0.12$), longest for luxury-sedan::nonluxury-SUV trials ($M = 2.08$ s, $SE = 0.17$), and intermediate when the vehicles were both SUVs ($M = 1.84$ s, $SE = 0.16$) or both sedans ($M = 1.83$ s, $SE = 0.15$). The choice data paralleled this pattern and were also consistent with the MaC account, $F(3, 57) = 9.64$, $MSE = 0.07$, $p < .01$, $\eta_p^2 = .34$. The choice percentages for the four pair types just listed were 95% ($SE = 0.09$), 78% ($SE = 0.26$), 92% ($SE = 0.11$), and 90% ($SE = 0.14$), respectively. In brief, these data indicate that some participants responded to status-different

⁴ Inferential statistics reported in this section were conducted on arcsine-transformed percentages.

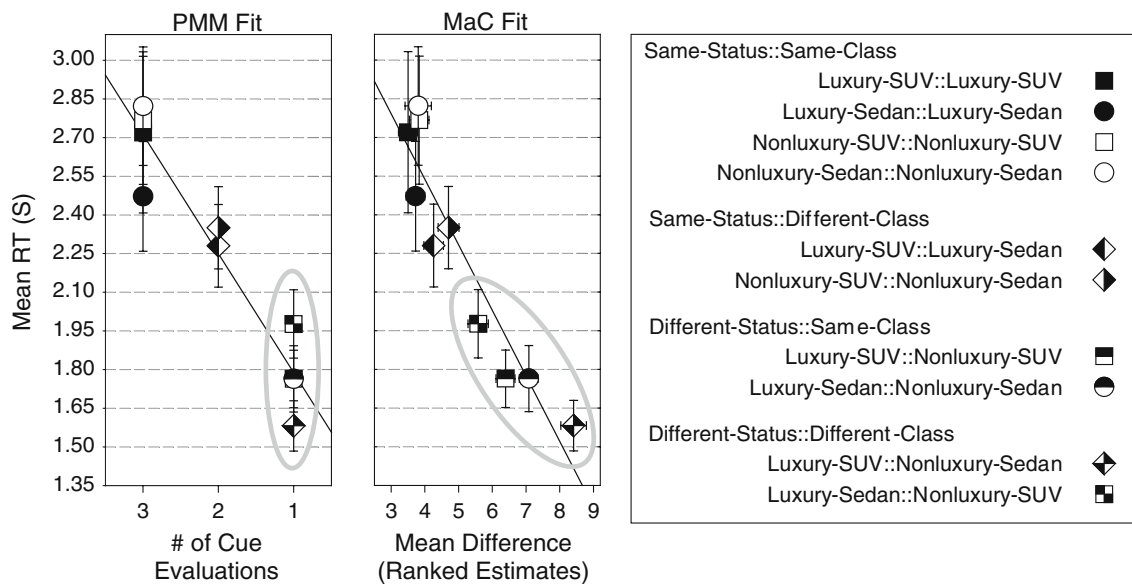


Fig. 2 Mean reaction time (RT; with standard error) for each pair type plotted against number of cue evaluations (**a** $R^2 = .91$) and mean difference in ranked price estimates (**b** $R^2 = .89$). Data for the different-status pairs are highlighted by the gray ovals appearing at the lower right of each figure. The markers in the legend are grouped

according to pair-type category. The first four markers represent same-status::same-class pairs, the next two represent same-status::different-class pairs, the two markers after that represent different-status::same-class pairs, and the final two represent different-status::different-class pairs

trials as if they were using the TTB heuristic, but most appeared to have relied on the MaC process to generate their responses.⁵

Discussion

This study was designed to test predictions derived from two very different approaches to the binary choice task. Both the PMM and MaC accounts correctly predicted longer RTs for same-status::same-class pairs than for same-status::different-class pairs and longer RTs for same-status::different-class pairs than for different-status pairs. Importantly, only the MaC account correctly predicted the observed pattern of RTs for the different-status pairs and the general pattern of choice for the different-status pairs, as well as the same-status::different-class pairs. These results indicate that participants often used the MaC process when selecting the more expensive of two vehicles. More

generally, this study suggests that this process plays an important role when people make binary decisions about real-world quantities (e.g., Brown & Siegler, 1991; Maki, 1981; Schweickart, Brown, & Lee, 2009).

A follow-up analysis indicated that participants could be divided into two groups; it appears that some individuals did rely on TTB but that many more engaged the MaC process to make their decisions. This conclusion is consistent with much prior research. It is well established that TTB and other one-reason decision strategies are not used all the time (e.g., Bröder, 2000; Bröder & Schiffer, 2003; Newell & Shanks, 2003; Newell et al., 2003), and there is abundant evidence that people have different strategy preferences (e.g., Payne, Bettman, & Johnson, 1993; Persson & Rieskamp, 2009; Rieskamp & Otto, 2006). What is new here is the notion that people employ the MaC process when they make binary decisions under uncertainty and the demonstration that they often rely on this process even when the task in question is clearly amenable to a cue evaluation approach.

Because the MaC process is known to play a central role in perceptual and symbolic comparisons (Moyer & Bayer, 1976), it is not surprising, at least not in retrospect, that it also dominated performance in the present study. Indeed, it could be that the MaC approach is the default strategy for making binary decisions under uncertainty. Whether it is or is not is an empirical question. Of course, an answer to this question must await additional research, research that explicitly recognizes the existence of the MaC process and acknowledges its potential relevance.

⁵ A detailed discussion of individual differences is beyond the scope of this report. However, it is worth noting the participants assigned to the TTB group produced price estimates that correlated strongly with the subjective magnitude values derived from Equation 1. For this group, the mean rank order correlation between estimated price and the derived magnitude values was .82 ($SE = .03$), whereas the mean correlation for the MaC group was .61 ($SE = .04$), $t(26) = 3.01$, $p < .01$. Thus, it is not the case that individuals assigned to the MaC group were more knowledgeable of the structure of the domain than those assigned to the TTB group.

Authors' Note This research was funded by the first author's NSERC operating grant. We thank Oliver Schweickart and Peter Lee for their contributions to this project.

References

- Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgments. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 101–159). New York: Academic Press.
- 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*, 1332–1346. doi:10.1037/0278-7393.26.5.1332
- Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multiattribute decisions. *Psychonomic Bulletin & Review*, *14*, 895–900.
- 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*, 277–293. doi:10.1037/0096-3445.132.2.277
- Brown, N. R., & Siegler, R. S. (1991). Subjective organization of U.S. presidents. *American Journal of Psychology*, *104*, 1–33. doi:10.2307/1422849
- Dougherty, M. R., Franco-Watkins, A. M., & Thomas, R. (2008). Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review*, *115*, 199–211. doi:10.1037/0033-295X.115.1.199
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669. doi:10.1037/0033-295X.103.4.650
- Gigerenzer, G., Todd, P. M., & The ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75–90. doi:10.1037/0033-295X.109.1.75
- Hamilton, J. M. E., & Sanford, A. J. (1978). The symbolic distance effect for alphabetic order judgments: A subjective report and reaction time analysis. *Quarterly Journal of Experimental Psychology*, *30*, 33–41. doi:10.1080/14640747808400652
- Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 566–581. doi:10.1037/0278-7393.26.3.566
- Kerst, S. M., & Howard, J. H., Jr. (1977). Mental comparisons for ordered information on abstract and concrete dimensions. *Memory & Cognition*, *5*, 227–234.
- Kosslyn, S. M., Murphy, G. L., Bemesderfer, M. E., & Feinstein, K. J. (1977). Category and continuum in mental comparisons. *Journal of Experimental Psychology: General*, *106*, 341–375. doi:10.1037/0096-3445.106.4.341
- Maki, R. H. (1981). Categorization and distance effects with spatial linear orders. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 15–32. doi:10.1037/0278-7393.7.1.15
- Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, *52*, 29–71. doi:10.1023/A:1015516217425
- Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, *8*, 228–246. doi:10.1016/0010-0285%2876%2990025-6
- Moyer, R. S., & Dumais, S. T. (1978). Mental comparison. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 12, pp. 117–155). New York: Academic Press.
- Murray, K. B., & Brown, N. R. (2009). A feature-based inference model of numerical estimation: The split-seed effect. *Acta Psychologica*, *131*, 221–234. doi:10.1016/j.actpsy.2009.05.007
- 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*, 53–65. doi:10.1037/0278-7393.29.1.53
- 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. doi:10.1016/S0749-5978%2802%2900525-3
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.
- Persson, M., & Rieskamp, J. (2009). Inferences from memory: Strategy- and exemplar-based judgment models compared. *Acta Psychologica*, *130*, 25–37. doi:10.1016/j.actpsy.2008.09.010
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207–236. doi:10.1037/0096-3445.135.2.207
- Schweickart, O., Brown, N. R., & Lee, P. J. (2009, November). *On the role of recognition and magnitude-comparison in binary decision tasks*. Paper presented at the annual meeting of the Society for Judgment and Decision Making, Boston, MA.
- Siegler, R. S. (1987). The perils of averaging data over strategies: An example children's addition. *Journal of Experimental Psychology: General*, *116*, 250–264. doi:10.1037/0096-3445.116.3.250
- Siegler, R. S., & Robinson, M. (1982). The development of numerical understanding. In H. Reese & L. Lipsitt (Eds.), *Advances in child development and behavior* (pp. 242–312). New York: Academic Press.