Psychological Review
2008, Vol. 115, No. 1, 263–273

0033-295X/08/$12.00 DOI: 10.1037/0033-295X.115.1.263

Process Models Deserve Process Data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006)

Eric J. Johnson
Columbia University

Michael Schulte-Mecklenbeck
University of Bergen

Martijn C. Willemsen
Eindhoven University of Technology

Resolution of debates in cognition usually comes from the introduction of constraints in the form of new data about either the process or representation. Decision research, in contrast, has relied predominantly on testing models by examining their fit to choices. The authors examine a recently proposed choice strategy, the priority heuristic, which provides a novel account of how people make risky choices. The authors identify a number of properties that the priority heuristic should have as a process model and illustrate how they may be tested. The results, along with prior research, suggest that although the priority heuristic captures some variability in the attention paid to outcomes, it fails to account for major characteristics of the data, particularly the frequent transitions between outcomes and their probabilities. The article concludes with a discussion of the properties that should be captured by process models of risky choice and the role of process data in theory development.

Keywords: risky choice, decision making, cognitive processes, process tracing

Decision research has largely progressed through the use of models that account solely for observed choices and that say little of the underlying cognitive processes. Great progress has been made in developing descriptive models of human choice behavior, in part through the design of clever experiments. However, there are several cases in which quite different processes are proposed to account for the same outcome data. One example is widely studied in decision making: choice between gambles. A stark contrast exists between *integration models*, such as prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) and others, all of which are modifications of the expected utility model that integrate probabilities and payoffs, and other *heuristic models* such as the priority heuristic (Brandstätter, Gigerenzer, & Hertwig, 2006) and others (see Payne, Bettman, & Johnson, 1993, for a review), which do not integrate these kinds of information, yet are also thought to describe choice data. In this comment, we argue that decision research will progress more quickly by doing more than simply assessing the predictions of as-if models (termed "paramorphic" by Hoffman, Slovic, & Rorer, 1968) and instead focusing on approaches that provide richer descriptions of processes and representations. Such models can also be tested on a functional (process) level rather than just on their outcome predictions, leading to the quicker modifications of models and the development of better models through the incorporation of added constraints. Although there have been historical calls for such a shift (Einhorn, Kleinmuntz, & Kleinmuntz, 1979), there has been little integration of these two traditions.

In this article, we illustrate this point by developing a set of process predictions implied by the priority heuristic (PH), a model posited by Brandstätter et al. (2006) that demonstrates impressive predictive strength in the aggregate and provides accounts for several decision-making phenomena such as Allais' paradox, risk-seeking and risk-aversion behavior elicited by different levels of probability and outcome, the certainty effect, and intransitive preferences. Within the limitations on the applicability of the heuristic,[1] Brandstätter et al. argued that it provides a better account of choices normally explained by alternative theories. The PH is an ideal example because it is

Eric J. Johnson, Columbia Business School, Columbia University; Michael Schulte-Mecklenbeck, Faculty of Psychology, University of Bergen, Bergen, Norway; Martijn C. Willemsen, Department of Technology Management, Eindhoven University of Technology, Eindhoven, the Netherlands.

This research was supported by National Science Foundation Grant SES-0352062 awarded to Eric J. Johnson and by Swiss National Science Foundation Grant PBFR1-110407 awarded to Michael Schulte-Mecklenbeck. Eric J. Johnson and Michael Schulte-Mecklenbeck contributed equally to this research. We thank participants at the Decision Workshop at Universitat Pompeu Fabra and the Workshop on The Application of Conversational Pragmatics to Understanding Reasoning and Decision Making, Toulouse, France as well as James Bettman, Arndt Bröder, Ben Newell, John Payne, J. Edward Russo, and George Wu for their helpful comments.

Correspondence concerning this article should be addressed to Eric J. Johnson, Columbia Business School, Columbia University, Uris Hall, 3022 Broadway, New York, NY 10027. E-mail: ejj3@columbia.edu

[1] Limitations suggested by Brandstätter et al. (2006) included the following: The heuristic does not apply for gambles where one option dominates the other and where expected values are "strikingly" different. Individual differences may account for different orders and different aspiration levels. Low stakes can result in the same reversals as individual differences. Discrepant EVs result in the PH making incorrect predictions (e.g., with the Erev, Roth, Slonim, & Barron, 2002, data). Differences in choices manipulated by problem representation (combination from three to two options in a gamble) as shown by Birnbaum (2004) are not predicted by PH.

ambitious and goes beyond simply predicting choices; it represents a process model for making these choices, predicting, for example, what information will receive attention and what will be ignored. Thus, Brandstätter et al. argued, the PH should not only provide superior predictions of the output of a choice process but also account for information acquisition as well. Brandstätter et al. did report reaction time data consistent with the PH, but the purpose of this comment was to illustrate the theoretical value of combining stronger process predictions and more informative data for this and other proposed models of choice and inference.

Recent Developments in Process Tracing

Monitoring the information used in the course of making a decision provides one common source of data to describe decision processes. Early applications used manual retrieval during decisions (Bettman & Jacoby, 1976; Payne, 1976) and eye movement recording (Russo & Rosen, 1974) followed by the analysis of click streams in computer-based environments. Recently, these computer-based approaches have been adopted in economics as well (Costa-Gomes & Crawford, 2006; Costa-Gomes, Crawford, & Broseta, 2001; Gabaix, Laibson, Moloche, & Weinberg, 2006; Johnson, Camerer, Sen, & Rymon, 2002). For example, Costa-Gomes et al. (2001) used information acquisition data to significantly improve the prediction of choices and the classification of people into strategic types.

This increased adoption has been due to the development of better theory linking information acquisition to the underlying choice process. Costa-Gomes et al. (2001) suggested two properties that strengthen the mapping of acquisition to process. The first, *occurrence*, appeals to a simple assumption: Information not acquired by the decision maker cannot be used by a hypothesized strategy. The second, *adjacency*, suggests that information used in temporal proximity by a proposed strategy should be acquired in close proximity. For adjacency to hold, one must assume that the information needed by the strategy is easier to acquire than to memorize, a fact that can be empirically tested. Thus, if one sees little or no acquisition of outcomes with small probabilities, one would conclude by occurrence that these probabilities do not affect choices. Similarly, if we see repeated acquisitions of outcomes and their probabilities, we might assume, by adjacency, that the evaluation of the outcomes in some way depend on probabilities.

Process Predictions for the PH

At an abstract level, there are three components of the PH. For two-outcome gambles only containing gains these components are as follows:

1. Priority rule.—Consider reasons in the following order: minimum gain, probability of minimum gain, maximum gain.

2. Stopping rule.—Stop examination if the minimum gains differ by 1/10 (or more) of the maximum gain (aspiration level); otherwise, stop examination if probabilities differ by 1/10 (or more) of the probability scale.

3. Decision rule.—Choose the gamble with the more attractive gain (probability).

For gambles with more than two outcomes, the priority rule adds a fourth reason in the first component: the probability of maximum gain.

These processes provide additional constraints that can be used to assess the PH. Brandstätter et al. (2006, p. 424) suggested one example of such a constraint using the order of information acquisition suggesting that information should be considered in the order minimum gain, probability of minimum gain, and maximum gain.

It is possible to develop a more precise set of predictions by using occurrence and adjacency. First, we translate the rules formulated in the PH into a production systemlike set of process steps following Newell and Simon (1972) and Johnson and Payne (1985). Consider the choice between Gamble A, which pays \$3,000 with a probability of 0.75, or \$4,000 with a probability of 0.25, and Gamble B with a \$5,000 payoff and a probability 0.20, and a \$2,800 payoff with a probability of 0.80. We labeled the payoffs, or amounts to *W*in, $W_a^1 \ldots W_b^n$. The subscripts identify the gamble, *a* or *b*, and the superscripts identify the position of the payoff, using reading order (left to right, top to bottom) in the display. Each outcome has an associated probability that is labeled $P_a^1 \ldots P_b^n$. Thus, the second outcome of Gamble A, \$4,000, is labeled W_a^2, and its probability, 0.25, is P_a^2. To identify the rank order of the outcomes from minimum to maximum, we exchange the superscript denoting its position in the display with *min* and *max*. For example, the \$4,000 outcome for Gamble A, which is its maximum payoff, is now W_a^{max}, with P_a^{max} as its probability. Figure 1 presents the display we used to present this gamble to participants.

Using this notation, the three stages can be represented by a series of process steps.

Step 1. READING: The first step follows from Brandstätter et al. (2006, p. 424), who posited that the PH, like all heuristics, contains an exploratory state in which every piece of information is read in order to identify the relevant items. Thus, because a decision maker may not know the location of relevant information in a gamble display, they must first scan the gambles to locate it. Once this initial stage is finished, the choice phase begins entailing the following steps:

Step 2. CALCULATE $.1 \times W_a^{max}$ (aspiration level).

Step 3. ESTIMATE DIFFERENCE W_a^{min}, W_b^{min}.

Step 4. IF Step 3 $\geq$ Step 2, THEN stop ELSE Step 5 (one-reason stopping rule).

Step 5. ESTIMATE DIFFERENCE P_a^{min}, P_b^{min}.

Step 6. IF Step 5 $\geq$ .10, THEN stop ELSE Step 7 (two-reason stopping rule).

Step 7. CHOOSE the alternative with the greater W^{max} (three-reason stopping rule).

On the basis of these steps, we developed a set of hypothesized properties for the PH for both the attention (frequency and duration), given each of the gamble elements, and the transitions between the elements. We label these measures *f*, *d*, and *t*, respectively.

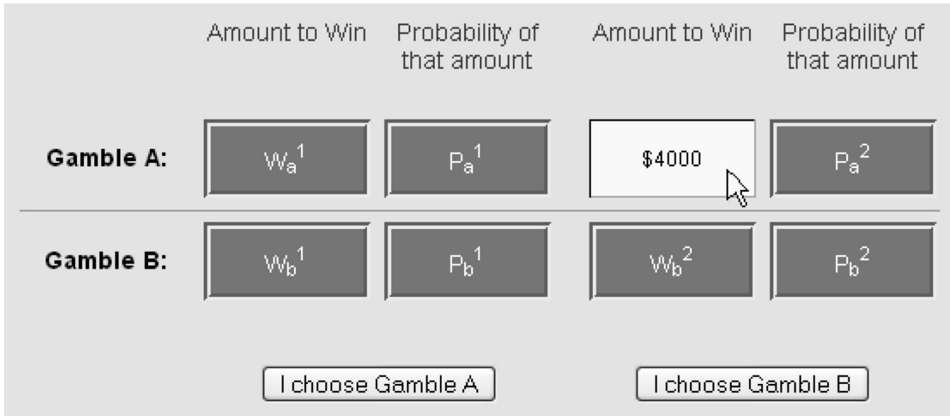

Figure 1. Screenshot of a MouselabWEB gamble with the Mouse opening box W_a^2, that is, \$4,000.

To organize these predictions, Figure 2 presents an icon graph portraying attention and transition predictions in a compact form (Johnson et al., 2002). The height of each rectangle in an icon graph represents f, the mean acquisition *f*requency, and the width represents d, the mean total (i.e., over all acquisitions) *d*uration of attention to that cell. Thus, bigger boxes mean more attention. The location of the boxes corresponds to the display in Figure 1. The arrows and their length portray t, the mean number of *t*ransitions between items in the display. Starting with the substantively more interesting choice phase, we argue that the PH suggests three families of hypotheses.

Hypothesis 1: The choice hypothesis: In the choice phase, comparisons *between* similar elements of Gambles A and B are used in the PH (e.g., W_a^{max} and W_b^{max}), and not, as in integration models, transitions within a gamble (e.g., P_a^{max} and the W_a^{max}). An additional property is that the probabilities of the maximum outcome are irrelevant to the decision maker. However, the processes embodied by the PH are contingent upon the choices themselves because of the following:

Hypothesis 1a: One-reason choices stop at Step 4 above. Thus, transitions between the two minimum payoffs should be made (W_a^{min}, W_b^{min}) in order to execute Step 3, but there should be few transitions between the minimum probabilities and maximum payoffs because Steps 5–7 are not executed. Similarly, attention measures in choice should reflect this pattern.

Hypothesis 1b: Three-reason choices require the execution of Steps 5–7. This means that, in addition to the transitions above, we should see attention to all outcomes (W_a^{max}, W_b^{max}, W_a^{min}, W_b^{min}), and the probabilities of the minimum outcomes (P_a^{min}, P_b^{min}), and transitions between those elements of the gamble, should be greater for three-reason gambles than for one-reason gambles.

Hypothesis 2: The probability-payoff hypothesis: Because the PH explicitly suggests that probabilities and payoffs are not integrated, transitions between an outcome and its probability should be relatively infrequent. We test this by comparing the number of transitions between *W*s and *P*s with the number of all other possible transitions. If $t(x,y)$ represents the number of transitions between two cells x and y, independent of direction, then we expect $t(P_a^{max}, W_a^{max}) + t(P_a^{min}, W_a^{min}) + t(P_b^{max}, W_b^{max}) + t(P_b^{min}, W_b^{min}) < t(W_a^{max}, W_b^{max}) + t(W_a^{min}, W_b^{min}) + t(P_a^{min}, P_b^{min})$. In Figure 2, this corresponds to a surplus of vertical arrows and a deficit of horizontal ones.

Hypothesis 3:The reading hypothesis: In the reading phase, comparisons of outcomes within each gamble are used in PH to find its smallest and largest payoffs (turning W_a^1 and W_a^2 into W_a^{min} and W_a^{max} etc.). Additionally, we should observe more attention (frequency and duration) for all *W*s than for all *P*s (because the *P*s are irrelevant to finding the largest payoffs) as well as a larger number of transitions within all *W*s than within all *P*s (see Table 2 for details of the test). Identifying the *max* and *min* outcome for each gamble is sufficient for reading, and because each probability is uniquely associated with a payoff, identifying the *max* and *min* outcomes also identifies what probability is associated with that outcome, and, strictly speaking, acquiring probabilities is not necessary.

Figure 2 summarizes predictions for the reading (left) and choice (right) phases and one-reason (top) and three-reason (bottom) choices. As displayed in the figure, we expect the reading phase (left half of the figure) to have large boxes for the *W*s, with many transitions among them, and small boxes and few transitions for the *P*s. Because the details of the reading phase are not as well specified, or as substantively important as the details of the choice phase, they receive less attention in our analyses. We also note that this set of hypotheses is an illustrative subset of implications of the PH, focusing on the properties uniquely associated with the heuristic.

Empirical Example

We used 16 of the 40 original gambles from the response time experiment of Brandstätter et al. (2006), using their categorization of choices according to number of *outcomes* and necessary *reasons* to make a choice. We used four choices of each resulting type

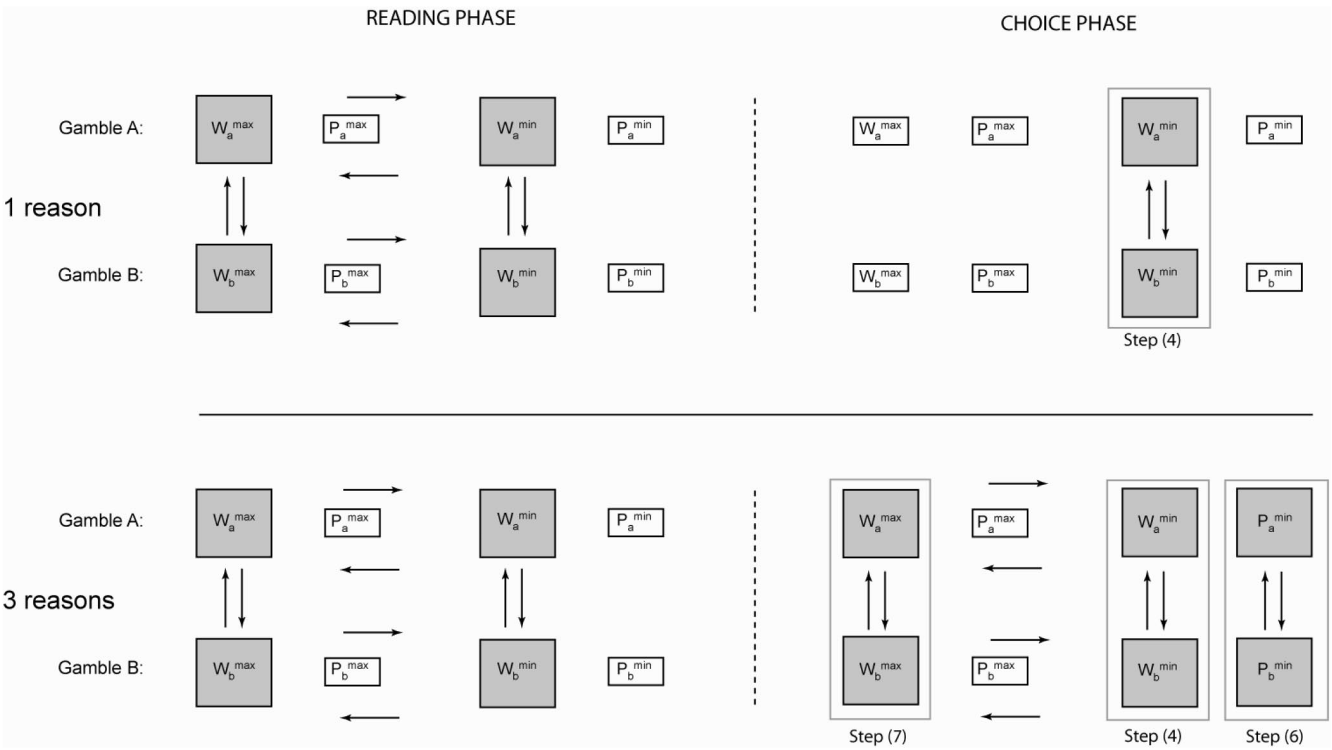

Figure 2. Icon graphs presenting predictions of the priority heuristic for reading phase (left column) and choice phase (right column), separately for one (top row) and three-reason (bottom row) gambles. Boxes with white backgrounds receive minimal attention. Within a graph, each rectangle corresponds to one of the cells in Figure 1.

(two- or five-outcome choices that required either one or three reasons). The position of the gambles in the display, order, etc., was counterbalanced to control for reading order. Respondents made choices in a Web browser running MouselabWEB (Willemsen & Johnson, 2006), which captured acquisition times and search patterns. For methodological details, see Appendix A.

To examine the empirical predictions of the PH, we distinguish between the reading and choice phases. The *reading phase* was identified as all acquisitions made before all outcomes have been examined at least once. The *choice phase* consisted of all subsequent acquisitions. (For a similar rule, see Klayman, 1983.)

Adjacency does seem to hold for these data. Although we removed very brief acquisitions, there are many acquisitions and reacquisitions of the information: For the two outcome gambles, the eight cells were acquired a total of 26.7 times (an average 3.3 times per cell). Overall decisions took, on average, 21.0 s.

We first examine each of our hypotheses using a graphic display and then report the results of more formal statistical tests. Figure 3 presents an icon graph of the observed data, for two-outcome gambles, similar to the prediction graph in Figure 2. The lower left-hand corner presents a scale for each measure. Thus, in the reading phase, the maximum amount to win for Gamble A, W_a^{max}, is examined, on average, for 912 ms, acquired 1.15 times, and transitions between this box and its probability P_a^{max} occur .63 times.

The choice hypothesis (Hypothesis 1a, see also Figure 2) predicts that for one-reason choices, one pair of outcomes, W_a^{min} and W_b^{min}, receives the most attention and dominates the transitions. Figure 3 shows, in the top right, that this did not occur in our data: These boxes do not receive more attention, nor are there substantially more transitions between the two minimum gain outcomes. Attention, instead, is more evenly distributed across payoffs and probabilities.

For three-reason choices (Hypothesis 1b, Figure 3 bottom right), the choice hypothesis predicts that additional time will be spent comparing the probabilities of the minimum outcomes, $P_a^{min}$and P_b^{min}, and the maximum outcomes, W_a^{max} and W_b^{max}. However, although the picture looks very similar to that for one-reason choices, there are differences in the attention given to outcomes and probabilities when we compare one- and three-reason choices. Although there is a slight increase in the number of between-gamble transitions, consistent with the PH, the predicted additional attention to the probabilities of the minimum outcomes (P_a^{min}, P_b^{min}) is not apparent, and there is significant attention given to the probabilities of the maximum outcomes, which, according to the PH, should be ignored in both one-reason and three-reason choices.

The probability-payoff hypothesis (Hypothesis 2) suggests that transitions between outcomes and their corresponding probabilities should be rare. In fact, this is the most common transition, and this pattern appears in both reading and choice phases and for one- and three-reason choices. Note that the transition between a payoff and an unassociated probability (e.g., P_a^{min} and W_a^{max}) serves as one control for accidental transitions. Although they occur, they are

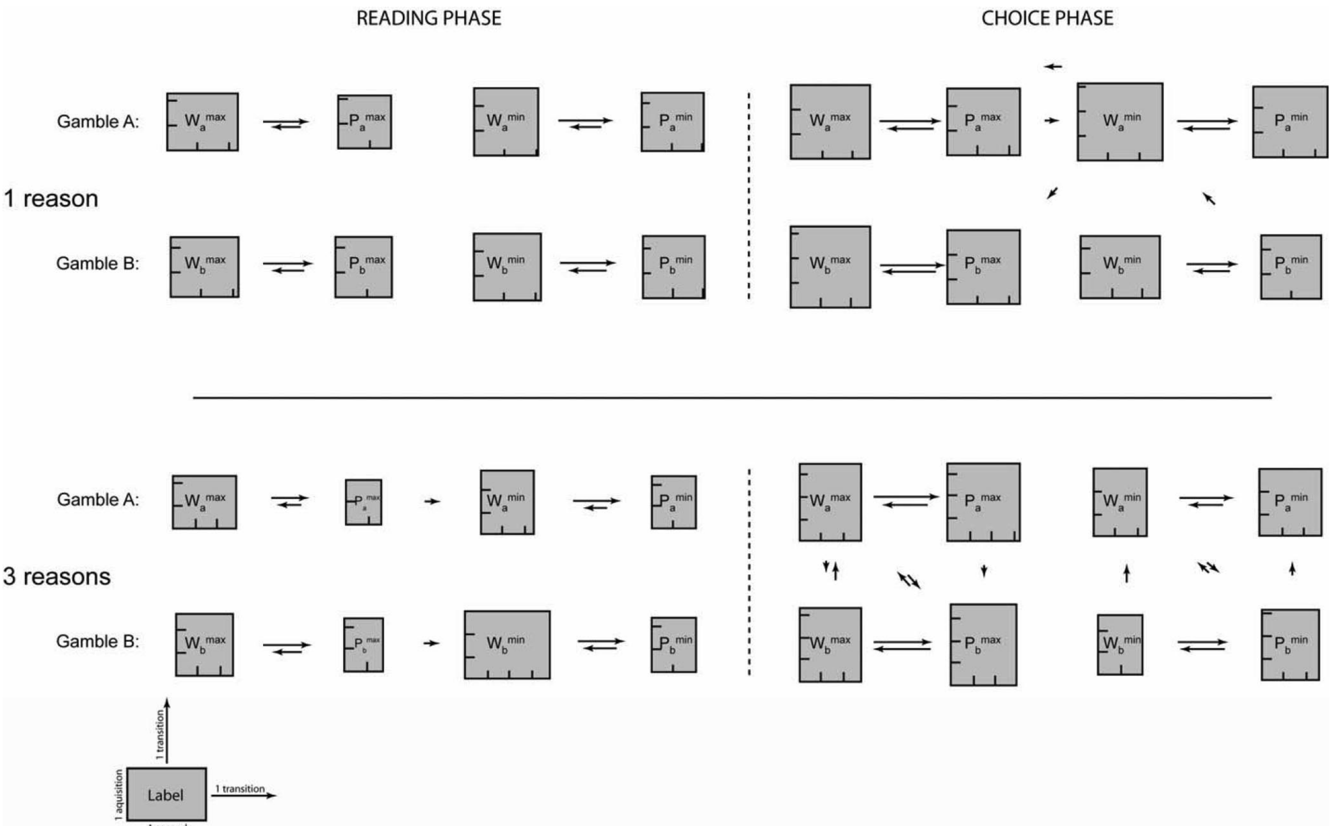

Figure 3. Icon graph of the observed data by phase and gamble type, corresponding to the predictions in Figure 2. For each, rectangle height represents the frequency of acquisition, and width represents the total duration. The length of arrows corresponds to the frequency of transitions. For clarity, transitions occurring fewer than an average of .15 per trial are not displayed. Ticks represent 400 ms and .5 acquisitions.

much less frequent than those between that probability and its outcome.

Finally, the reading hypothesis (Hypothesis 3) predicts that in the reading phase, we should primarily see transitions between the outcomes within a single gamble (e.g., W_a^{max}, W_a^{min}, and W_b^{max}, W_b^{min}, etc.), as seen in the left half of Figure 2. The results in the left half of Figure 3 are very different from the predictions and show roughly the same amount of information acquired for outcomes and probabilities. Figure 3 also reveals that there are very few of the transitions predicted by the PH that compare the payoffs to determine which are greater and which are smaller.

To provide more formal tests of the observations contained in Figure 3, we used a random coefficients model (see Costa-Gomes et al., 2001; Willemsen, Böckenholt, & Johnson, 2006, for a similar approach) that allows intercepts to vary across respondents as well as heterogeneity in how much processing occurs in each phase (see Appendix B for details). Table 1 reports the means and resulting contrasts that represent these hypotheses for the number of acquisitions *f*, the duration of looking time *d*, and number of transitions *t*. The last column summarizes the result for each hypothesis. Note that stronger tests of the PH are possible. For example, whereas a strong test would suggest that no acquisitions of probability information should occur in the reading phase, or that such acquisitions should occur at a chance rate, we actually conducted a less stringent test that requires simply that there are significantly more acquisitions of outcomes than probabilities.

For the choice hypotheses (Hypothesis 1a), describing one-reason choices, the data failed to show significant differences in the expected direction for frequency, duration, and transitions. For the three-reason choices (Hypothesis 1b), attention did shift in the predicted directions, but this difference was most significant for the minimum probabilities. This shift was accompanied by a small but significant increase in the appropriate transitions. The most striking result contradicts the probability-payoff hypothesis (Hypothesis 2) that predicts that between-gamble transitions will be more frequent than payoff-probability transitions. Confirming the visual impression from Figure 3, this result is significant, but in the opposite direction: These transitions are, in fact, much less frequent than those between probabilities and payoffs. Consistent with the reading hypothesis (Hypothesis 3), frequencies and duration show significantly more attention paid to payoffs during reading. However, the pattern of transitions was, in fact, significant in the opposite direction than the PH would predict.

We conducted a similar analysis for five-outcome gambles. Because of space constraints, we only report that the results are similar to those in Figure 3. Table 2 presents the contrasts for the five-outcome gambles following the same structure as the analysis for the two-outcome gambles. Consistent with Hypothesis 3, pay-

Table 1
Priority Heuristic (PH) Predictions and Tests for Two-Outcome Gambles for Frequency, f, Duration, d, in Milliseconds, and the Number of Transitions, t

Hypothesis (H)	Means	Test statistic	Supports PH
H1a: Choice, one reason			
$W_a^{min} + W_b^{min} > W_a^{max} + W_b^{max}$	*f*: 0.70 > 0.68	*f*: $t(3{,}682) = 0.34$, *ns*	No support
	d: 441.85 > 450.68	*d*: $t(3{,}682) = -0.21$, *ns*	
$t(W_a^{min}, W_b^{min}) > t(W_a^{max}, W_b^{max})$	*t*: 0.19 > 0.11	*t*: $t(13{,}855) = 1.85$, *ns*	
H1b: Choice, three reasons			
$W_a^{max} + W_b^{max}$ > for 3- than for 1-reason choices	*f*: 0.88 > 0.68	*f*: $t(1{,}308) = 2.71^{**}$	Mostly supported
	d: 340.25 > 316.26	*d*: $t(4{,}060) = 0.64$, *ns*	
$W_a^{min} + W_b^{min}$ > for 3- than for 1-reason choices	*f*: 0.78 > 0.70	*f*: $t(1{,}308) = 1.13$, *ns*	
	d: 378.72 > 272.13	*d*: $t(4{,}060) = 2.87^{**}$	
$P_a^{min} + P_b^{min}$ > for 3- than for 1-reason choices	*f*: 0.95 > 0.69	*f*: $t(1{,}308) = 3.37^{**}$	
	d: 396.13 > 253.31	*d*: $t(4{,}060) = 3.85^{**}$	
$t(W_a^{min}, W_b^{min}, W_a^{max}, W_b^{max}) + t(P_a^{min}, P_b^{min})$ > for 3- than for 1-reason choices	*t*: 0.18 > 0.13	*t*: $t(13{,}837) = 2.84^{**}$	
H2: Probability-payoff			
$t(P_a^{max}, W_a^{max}) + t(P_a^{min}, W_a^{min}) + t(P_b^{max}, W_b^{max}) + t(P_b^{min}, W_b^{min})$ < all other transitions	*t*: 0.76 < 0.11	*t*: $t(13{,}692) = -62.06^{**}$	Significant in opposite direction
H3: Reading			
$(W_a{}^1 + W_a{}^2 + W_b{}^1 + W_b{}^2) > (P_a{}^1 + P_a{}^2 + P_b{}^1 + P_b{}^2)$	*f*: 1.20 > 0.87	*f*: $t(3{,}682) = 8.81^{**}$	Mixed, as predicted for attention, but opposite direction for transitions.
	d: 805.22 > 553.47	*d*: $t(3{,}682) = 9.78^{**}$	
$t(W_a^{max}, W_b^{max}) + t(W_a^{min}, W_b^{min})$ > all other transitions	*t*: 0.11 > 0.17	*t*: $t(13{,}841) = -2.62^{**}$	

** $p < .01$.

offs receive more attention and result in more transitions during reading. Another prediction made by the PH for the five-outcome data is that the intermediate outcomes, those that are not the minimum or the maximum payoffs of the gamble, should not affect choice or receive attention. In fact, the three intermediate outcomes received substantial attention in our data: On average, these "irrelevant" outcomes and probabilities got more attention (examined for 9.8 s during 13.6 acquisitions) than the maximum and minimum outcomes (examined for 9.5 s during 13.2 acquisitions).

Ordinal Tests

Brandstätter et al. (2006) suggested that information should be considered in a particular order, with the minimum gain first, followed by probability of minimum gain and then maximum gain. We calculated the mean rank of each acquisition separately for the reading and choice phases, as a function of choice, and tested these ranks for differences. The mean rank acquisition, in the order hypothesized by the PH, was 9.75 for W^{min}, 9.71 for P^{min}, and 10.13 for W^{max}. There was no significant difference in the order in which each reason was accessed, either as a main effect, as hypothesized by Brandstätter et al., or within phase or choice type.

Similarly, we examined the final comparison for each choice, based on the notion that this should differ for one- and three-reason choices. For one-reason choices, the final transition should be between the minimum outcomes $t(W_a^{min}, W_b^{min})$ dictated by Step 3, and for three-reason choices, it should be a transition $t(W_a^{max}, W_b^{max})$ suggested by Step 7. We found that a small proportion of all choices ended with the predicted transitions. Transitions between W_a^{min} and W_b^{min} were the last transition for only 2.3% of the one-reason choices and 0.8% for three-reasons choices, and $t(W_a^{max}, W_b^{max})$ occurred 2.3% and 4.5% of the time for one- and three-reason choices, respectively. These frequencies did not differ significantly, $\chi^2(1, N = 256) = 0.98$, *ns*. The most frequent terminal transition was between probabilities and the adjacent payoff: 67.4% for one-reason choices and 52.8% for three-reason choices.

In summary, we observed a very consistent picture. Participants showed interest in both probabilities and payoffs, mainly navigated through the available information within one gamble, and strongly favored transitions between outcomes and their corresponding probabilities. These patterns occurred across one- and three-reason choices and for the reading and choice phases. Consistent with the PH, we did see differences in the attention devoted to one- versus three-reason choices. However, these were not accompanied by an equivalent change in transitions.

Process Data and Choice

As a final check, we attempted to predict the choices made by respondents using the process tracing data. First, if we looked at choices alone, then one might conclude that the PH provides good predictions for choices, identifying the correct option more often than prospect theory for both the two- and five-outcome choices. This illustrates that predicting outcomes alone is not particularly helpful in assessing and building better choice models.

To demonstrate this, we constructed a model for the two-outcome choices, based on the number of transitions made by respondents within each of the two gambles during the choice phase. This statistical model is statistically significant and predicts which gamble will be chosen: The more transitions within a gamble, the more likely it is to be chosen. This suggests that the

Table 2
Priority Heuristic Predictions and Tests for the Five-Outcome Gambles for Frequency, f, Duration, d, in milliseconds, and the number of Transitions, t

Hypothesis (H)	Means	Test statistic	Supports PH
H1a: Choice, one reason			
$W_a^{min} + W_b^{min} > W_a^{max} + W_b^{max}$	*f:* 0.187 > 0.133	$t(13{,}626) = 1.87$, *ns*	Marginal support
	d: 337.54 > 308.26	$t(13{,}632) = 1.94$, *ns*	
$t(W_a^{min}, W_b^{min}) > t(W_a^{max}, W_b^{max})$	*t:* 0.07 > 0.03	$t(196{,}164) = 1.90$, *ns*	
H1b: Choice, three reasons			
$W_a^{max} + W_b^{max} >$ for 3- than for 1-reason choices	*f:* 0.17 > 0.13	$t(13{,}627) = 1.42$, *ns*	Mostly not supported
	d: 252.90 > 252.94	$t(13{,}640) = -0.001$, *ns*	
$W_a^{min} + W_b^{min} >$ for 3- than for 1-reason choices	*f:* 0.187 > 0.184	$t(13{,}627) = 0.11$, *ns*	
	d: 261.44 > 310.51	$t(13{,}640) = -1.65$, *ns*	
$P_a^{min} + P_b^{min} >$ for 3- than for 1-reason choices	*f:* 0.31 > 0.24	$t(13{,}627) = 2.16^{*}$	
	d: 384.91 > 369.73	$t(13{,}640) = 0.46$, *ns*	
$t(W_a^{min}, W_b^{min}, W_a^{max}, W_b^{max}) + (P_a^{min}, P_b^{min}) >$ for 3- than for 1-reason choices	*t:* 0.023 > 0.017	$t(132{,}038) = 1.13$, *ns*	
H2: Probability-payoff			
$t(P_a^{max}, W_a^{max}) + t(P_a^{min}, W_a^{min}) + t(P_b^{max}, W_b^{max}) + t(P_b^{min}, W_b^{min}) <$ all other transitions	*t:* 0.27 < 0.02	$t(254{,}536) = -156.2^{**}$	Significant in the opposite direction
H3: Reading			
$(W_a^1 + W_a^2 + W_a^3 + W_a^4 + W_a^5 + W_b^1 + W_b^2 + W_b^3 + W_b^4 + W_b^5) > (P_a^1 + P_a^2 + P_a^3 + P_a^4 + P_a^5 + P_b^1 + P_b^2 + P_b^3 + P_b^4 + P_b^5)$	*f:* 1.059 > 0.782	$t(13{,}626) = 13.43^{**}$	Supported
	d: 664.83 > 620.21	$t(13{,}632) = 9.24^{**}$	
$t(W_a^1, W_a^2, W_a^3, W_a^4, W_a^5, W_b^1, W_b^2, W_b^3, W_b^4, W_b^5) >$ all other transitions	*t:* 0.06 > 0.03	$t(230{,}421) = 29.31^{**}$	

* $p < .05$. ** $p < .01$.

underlying process driving choice is very different from that proposed by the PH and that increased attention to a gamble, and not certain outcomes, is associated with choice, a frequent result in riskless choice (Payne, 1976). This result also shows that the process data presented here are not epiphenomenal to choice, but are related to outcomes.

Prior Research

Obviously, our results depended, to some extent, on the particular display, gambles, population, and other characteristics unique to our experiment. However, we can examine prior research to see how robust some characteristics of this research might be because other researchers have used different displays, gambles, and information acquisition technologies. Although most of this research does not report as fine-grained analysis as the present article, one measure is included in many articles: a statistic reporting the relative frequency of transitions between two elements of a gamble, relative to transitions within a gamble. This statistic is approximately equivalent to a test of our Hypothesis 2. The PH would predict fewer transitions between probabilities and payoffs (which are consistent with the integration of these elements) than between the two gambles' payoffs and probabilities (which are consistent with the comparison processes posited by the PH). We calculated the ratio of these two types of transitions, as reported in several studies, using very different methods (trackball, mouse and eye tracking; Payne & Braunstein, 1978; Rosen & Rosenkoetter, 1976; Russo & Dosher, 1983; Schkade & Johnson, 1989). Although the PH would suggest that all of these indices should be less than 1, in reality, they all exceed this number. Averaging across studies, there are 70% more transitions within a probability and its payoff than those contrasting the two gambles' probabilities or payoffs.

By using a random coefficient model, we can examine individual differences in transitions across respondents, examining variations from the aggregate across our respondents. We focus on Hypothesis 2, estimating a random effect that allows the number of transitions to vary across people. This random effect was quite significant, so we constructed an index for each participant in our experiment. This index had a mean of 4.47, and none of the individual-level means were significantly less than 1. Thus, all respondents seemed most consistent with a strategy that uses probabilities in the evaluation of payoffs, but to varying degrees.

Conclusion

Our observations raise an important question: What are people doing when they make choices between gambles? It is beyond the scope of the present article to provide an answer. However, our results do suggest some stylized facts for future models:

> The probability-payoff transition is common, and the comparisons of outcomes across gambles are rare. However, the transition between a probability, say 0.3, and its payoff, say \$11, may not represent an explicit multiplication but some other process that weights a payoff by its likelihood. Instead, it may indicate how attention should be given to an outcome.

> Heuristics differ across individuals: Our analysis of the transition data reveals sizable differences in how people approach these choice problems. Future models may abandon the idea of a single underlying heuristic for choice.
>
> Heuristics also seem to differ across gamble types. One of the appealing features of the PH is that it proposes processes that are contingent upon the kind of gamble. Although our results are not particularly supportive of the PH in the way it combines information, we do find support for the idea that attention differs across gamble types. This, of course, may also be consistent with rank-dependent choice theories, including cumulative prospect theory.

These facts suggest solutions to a remaining puzzle: How can a model that seems to predict choices well fail to capture the process? We suspect that the PH captures, in the aggregate, shifts in choice processes that occur across both problems and individuals. Confirmation of this speculation is left for further research.

Our hope in this comment has been to encourage the use of process-oriented models and data in decision research. This represents a challenge because the data that arise from the analysis of these relatively simple gambles provide a richer description of cognition than simply modeling choice. At the same time, the clues and constraints present in process data should allow the rapid development of models that are more faithful to the processes made to make decisions. We optimistically agree with Brandstätter et al. (2006) that "process models of heuristics are key to opening this black box." (p. 427) and add that progress would be made more quickly with the use of process data, both in the study of choice (Payne et al., 1993) and inference.

References

Bettman, J. R., & Jacoby, J. (1976). Patterns of processing in consumer information acquisition. In B. B. Anderson (Ed.), *Advances in consumer research* (Vol. 3, pp. 315–320). Chicago: Association for Consumer Research.

Birnbaum, M. (2004). Causes of Allais common consequence paradoxes: An experimental dissection. *Journal of Mathematical Psychology, 48,* 87–106.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113,* 409–432.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459.

Costa-Gomez, M., & Crawford, V. P. (2006). Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review, 96,* 1737–1768.

Costa-Gomes, M., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica, 69*(5), 1193–1235.

Einhorn, J. H., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Regression models and process tracing analysis. *Psychological Review, 86,* 465–485.

Erev, I., Roth, A. E., Slonim, R. L., & Barron, G. (2002). *Combining a theoretical prediction with experimental evidence to yield a new prediction: An experimental design with a random sample of tasks*. Unpublished manuscript, Columbia University and Faculty of Industrial Engineering and Management, Techion, Haifa, Israel.

Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review, 96,*1043–1068.

Hoffman, P. J., Slovic, P., & Rorer, L. G. (1968). An analysis-of-variance model for the assessment of configural cue utilization in clinical judgment. *Psychological Bulletin, 69,* 338–349.

Johnson, E. J., Camerer, C. F., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory, 104*(1), 16–47.

Johnson, E. J., & Payne, J. W. (1985). Effort and accuracy in choice. *Management Science, 31,* 395–414.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica, 47,* 263–291.

Klayman, J. (1983). Analysis of predecisional search patterns. In P. Humphreys, O. Svenson, & A. Vari (Eds.), *Analyzing and aiding decision processes* (pp. 400–414). Amsterdam: North Holland.

Lohse, G., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes, 68,* 28–43.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Engelwood, CA: Prentice Hall.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Decision Processes, 16,* 366–387.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 534–552.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.

Payne, J. W., & Braunstein, M. L. (1978). Risky choice: An examination of information acquisition behavior. *Memory and Cognition, 6,* 554–561.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). London: Sage.

Rosen, L. D., & Rosenkoetter, P. (1976). An eye fixation analysis of choice of judgment with multiattribute stimuli. *Memory and Cognition, 4,* 747–752.

Russo, J. E., & Dosher, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 676–696.

Russo, J. E., & Rosen, L. D. (1974). An eye fixation analysis of multialternative choice. *Memory and Cognition, 3,* 267–276.

Schkade, D., & Johnson, E. (1989). Cognitive processes in preference reversals. *Organizational Behavior and Human Decision Processes, 44,* 203–231.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5,* 297–323.

Willemsen, M. C., Böckenholt, U., & Johnson, E. J. (2006). *How losses affect choice: Loss encoding, gain enhancement and a comparison of three theories*. Unpublished manuscript, Eindhoven University of Technology, Eindhoven, the Netherlands.

Willemsen, M. C., & Johnson, E. J. (2006). MouselabWEB: Monitoring information acquisition processes on the Web. Retrieved August 15, 2006, from http://www.mouselabweb.org/

Appendix A

Experimental Method

A total of 77 participants (51 women, 26 men) were recruited through the virtual lab at Columbia University and participated online in the experiment. The mean age was 42.8 years ($SD = 12.7$ years). Participants were compensated with \$ 5.

The gambles[1] were presented with a recording tool called MouseLabWEB (http://www.mouselabweb.org), which enables the researcher to run experiments that collect process data with high accuracy. In this methodology, a computer mouse is used by the participant to uncover hidden information from a matrix. As soon as the mouse is moved over an information cell, the content is displayed; after moving the mouse out of the box, the content is hidden again. Frequency and length of viewing are recorded as well as the sequence of acquisition and the final decision the participant makes after the information search.

The two option gambles (Brandstätter et al., 2006) were presented in two basic setups in a between-subjects design. A *horizontal setup* (see Figure 1) with one gamble (two options) per row and a *vertical setup* with one gamble (two options) per column was administered. These setups were used to minimize effects that reading order, for example, could have on the search process. Furthermore for each (win–probability) pair, the position (left, right, top, bottom) was counterbalanced using a Latin square with the restriction that corresponding options had to remain connected (within a row or column). The five-option gambles were only presented in a vertical setup with one gamble (five options) per column. The horizontal layout is not ideal to implement because the width of most computer screens (even with high resolution) would have not been sufficient to display all options at once.

After a detailed stepwise description of the components of a gamble, a warm-up task followed to train participants to use MouseLabWEB. Furthermore, it was ensured that participants understood concepts like "What is the largest amount to win?" or "What is the smallest probability?" through questions targeted at these values.

Of course, the mapping between a hypothesized cognitive process and information acquisition is not perfect. Although the PH does not have an error theory, it is possible that information search contains extraneous acquisitions. To minimize these errors of commission, we eliminated all acquisitions of less than 100 ms as standard practice with this kind of data (Payne, Bettman, & Johnson, 1988).

We performed a number of robustness checks to check the sensitivity of our results to how we preprocessed the data, divided the decision into two stages, etc. For preprocessing, changing the threshold to 200 ms did not change our results. Similarly, dividing each decision into two halves instead of using a data-driven specification of reading versus choice phases did not appreciably change the outcome of our analysis. This division did have one advantage over the data-driven analysis: Some participants in our analysis did acquire all of the information and therefore did not have a choice phase when we used the data-driven definition. Our visual analysis eliminates those observations, but we have performed the statistical analysis both with those values set to zero (reported in the tables) and set to missing, with no substantive change in results.

Appendix B

Model and Estimation

The analysis contained factors from the experimental design, including a fixed between-respondent effect of orientation and fixed within-respondent effects that generate the eight cells seen in Figure 1: gamble (*a* or *b*), cell type (probability or outcome), size (maximum or minimum), as well as a random effect representing the effect of phase (reading or choice). Thus, each choice had 16 observations, defined by the eight cells and two phases. Using the notation of Raudenbush and Bryk (2002), the model to be estimated for the attention measures is specified as:

Level 1:

$$\log(Y_{ij}) = \beta_{\text{int}\,j} + \beta_{phase\,j} \times phase_{ij} + \beta_{nreasons\,j} \times nreasons_{ij} + \sum_{1}^{n} \beta_z \cdot X_{ijz} + r_{ij}, \text{ where } r_{ij} \sim N(0,\sigma^2).$$

Level 2:

$$\beta_{\text{int}\,j} = \gamma_{00} + u_{0j}, \quad \text{where } E[u_{0j}] = 0 \text{ and } \operatorname{Var}[u_{0j}] = \tau_{00}.$$

$$\beta_{phase\,j} = \gamma_{10} + u_{1j}, \quad \text{where } E[u_{1j}] = 0 \text{ and } \operatorname{Var}[u_{1j}] = \tau_{11}.$$

$$\beta_{nreasons\,j} = \gamma_{20} + u_{2j}, \quad \text{where } E[u_{2j}] = 0 \text{ and } \operatorname{Var}[u_{1j}] = \tau_{22}.$$

In this model, Y_{ij} indicates the *i*th observation within subject *j*. X_{ijz} is the matrix of fixed effects as described above. The Level 1 model has three coefficients with random variation over participants, denoted by the random effect u_{xj}. Specifically, we allow the intercept β_{int} (participants differ in the mean number of acquisitions/time), the phase coefficient β_{phase} (participants differ in how they differ between phases within a decision), and the coefficient for number of reasons $\beta_{nreasons}$ (participants differ in how they change strategies for more complex problems) to be random. The

[1] In reading order, we used Gambles 1, 3, 7, 9, 12, 15, 17, and 19 from Brandstätter et al.'s (2006) appendix for both two- and five-outcome choice pairs.

(*Appendix continues*)

random effects are assumed to be normally distributed, with mean vector 0 and variance terms τ_{xy}.

Because the dependent measures, time and frequency, are skewed, they were log-transformed to make them more symmetrically distributed. We note that because of the discrete nature of the frequency data, a linear regression might not be an appropriate model. We therefore also estimated a nonlinear Poisson regression model. However, because using a Poisson specification yielded similar results to the ones obtained under the linear model of the log-transformed data, we do not discuss the results of this analysis in detail.

For transitions, we modeled the relevant transitions between the eight boxes using a similar model, which has 56 observations per gamble per participant due to two factors: a transition identifier (28 levels) and phase. Again, we used an intercept, phase and number of reasons as random factors. In addition, as described in the text, we estimated a model that allowed the impact of kind of transition to vary across respondents, which resulted in improved fit and substantively interesting interpretations. We used two-tailed tests throughout, both because this analysis has substantial power and because prior data suggests results that conflict with the predictions of the priority heuristic.

Received August 16, 2006
Revision received January 10, 2007
Accepted January 11, 2007 ■

Postscript: Rejoinder to Brandstätter, Gigerenzer, and Hertwig (2008)

Eric J. Johnson
Columbia University

Michael Schulte-Mecklenbeck
University of Bergen

Martijn C. Willemsen
Eindhoven University of Technology

We appreciate that Brandstätter, Gigerenzer, and Hertwig (2008) agree that process models are indeed useful for advancing researchers' understanding of choice processes. Their statement that choices represent adaptive process is also welcome (see Payne, Bettman, & Johnson, 1993), as is the emphasis on multiple measures. However, we do disagree on two empirical matters:

Describing our research, Brandstätter et al. (2008) posit that most tests are either null or supportive of the priority heuristic (PH) and that only three of the tests were significant in the opposite direction. This "scorekeeping heuristic" implies that all tests are equally weighted in theory testing. However, we believe that, one, the presence of mostly probability-payoff transitions in the data is a critical test for any falsification of the PH. Although Brandstätter et al. consider the existing evidence on this point mixed, we do not. Table 1 contains the results from a series of studies, almost all conducted before Brandstätter et al. started their research and many of which are well-known in the literature. As can be seen from Table 1, the majority of observations show a predominance of probability-payoff transitions.

Brandstätter et al.'s (2008) effort to provide quantitative predictions is laudatory. However, these predictions are *very* sensitive to the assumptions that are made in the first, reading, phase postulated by Brandstätter et al. They originally suggested (Brandstätter, Gigerenzer, & Hertwig, 2006, p. 424) that all choice heuristics have a reading phase in which one looks for relevant information and a choice phase in which the relevant information is used. This means that the choice phase is useful for understanding heuristics, whereas the reading phase is more epiphenomenal. The purpose of their original reading phase was to find the larger payoffs. On the basis of adjacency (the idea that participants would inspect information when it is needed), we argued that this would naturally lead to comparisons between payoffs (see Figure 2 of Johnson, Schulte-Mecklenbeck, & Willemsen, 2008). Brandstätter et al. now assume that all the information is read first for Gamble A, then for Gamble B, replacing the vertical arrows in our Figure 2 in Johnson et al. (2008) with horizontal arrows. This new position represents a quite testable hypothesis because it suggests large changes in search patterns by phase and orientation.

The bigger point is that Brandstätter et al. (2008), in their new analysis, also ignore empirically their own distinction between reading and choice phase. By stating that their reading phase now consists mainly of probability-payoff transitions, and then combining the two phases, they use those transitions to compensate for their absence in the choice phase, making the PH look artificially good, and remove one of its clearest predictions.

In essence, we agree with the goals of the approach taken by Brandstätter et al. (2008) but argue that if process-tracing data is to inform the development of choice models, then it is important to listen what the data are saying.

Table 1
Ratio of Transitions Types: Between Probabilities and Payoffs/ Within Probabilities and Payoffs, Prior Studies Using Gambles Compatible With the Priority Heuristic

Study	Condition/ experiment	Method	Ratio
Brandstätter et al. (2008)		Mouselab	2.34
Johnson et al. (2008)	2 outcome	Mouselab	2.13
	5 outcome		0.56
Payne & Braunstein (1978)	2 outcome	Trackball	1.94
Rosen & Rosenkoetter (1978)		Eye tracking	1.63
Russo & Dosher (1983)	Experiment 3	Eye tracking	2.20[a]

[a] Weighted by number of fixations.