Psychonomic Bulletin & Review
2010, 17 (6), 923-930
doi:10.3758/PBR.17.6.923

Notes and Comment

Reconsidering "evidence" for fast-and-frugal heuristics

Benjamin E. Hilbig
University of Mannheim, Mannheim, Germany
and Max Planck Institute for Research on Collective Goods, Bonn, Germany

In several recent reviews, authors have argued for the pervasive use of fast-and-frugal heuristics in human judgment. They have provided an overview of heuristics and have reiterated findings corroborating that such heuristics can be very valid strategies leading to high accuracy. They also have reviewed previous work that implies that simple heuristics are actually used by decision makers. Unfortunately, concerning the latter point, these reviews appear to be somewhat incomplete. More important, previous conclusions have been derived from investigations that bear some noteworthy methodological limitations. I demonstrate these by proposing a new heuristic and provide some novel critical findings. Also, I review some of the relevant literature often not—or only partially—considered. Overall, although some fast-and-frugal heuristics indeed seem to predict behavior at times, there is little to no evidence for others. More generally, the empirical evidence available does not warrant the conclusion that heuristics are pervasively used.

Recently, several theoretical contributions and reviews have been published that deal with the fast-and-frugal heuristics research program, advocated by Gigerenzer and colleagues (Gigerenzer, 2008; Gigerenzer & Brighton, 2009; Marewski, Gaissmaier, & Gigerenzer, 2010). In summarizing almost 2 decades of research, they have concluded that successful judgment and decision making stems from reliance on an adaptive toolbox of heuristics. First, they have provided well-argued illustrations of the reasons why simple, often noncompensatory judgment strategies can outperform allegedly more complex ones (cf. Hogarth & Karelaia, 2007). Few, I believe, would actually doubt that, in some situations, more complex strategies could actually lead us astray more severely than simple heuristics.

Second, these reviews have outlined exemplary heuristics that "according to empirical evidence, are likely to be in the adaptive toolbox of humans" (Gigerenzer, 2008, p. 23). The claim thus is that "people's behavior is often better explained by models of heuristics" (Goldstein & Gigerenzer, 2009, p. 762). Likewise, Marewksi, Gaissmaier, and Gigerenzer (2010) summarized previous theoretical and empirical work on "[w]hat heuristics . . . organisms *use* to make decisions" (p. 107, emphasis added). Among these are the recognition heuristic (Goldstein & Gigerenzer, 2002), the fluency heuristic (Schooler & Hertwig, 2005), the take-the-best heuristic (Gigerenzer & Goldstein, 1996), and others.

So, beyond the undisputed normative issue of showing that heuristics can produce good outcomes, the claim made by these authors is that decision makers indeed rely on such heuristics. Stated differently, "the crucial psychological assumption says that decision makers actually *do* use such heuristics" (Fiedler, 2010, p. 21, emphasis in original). However, the empirical evidence is not as unequivocal as their reviews imply. Most problematically, some of those studies seemingly bearing positive evidence for the use of heuristics are marred by methodological caveats. I will demonstrate these by proposing a new fast-and-frugal heuristic—purely for illustration, that is. This example will show that "evidence" for simple heuristics can easily be produced by inappropriate measures, thus leading to the premature conclusion that such heuristics are used. Indeed, most of the evidence available so far for the fluency heuristic is of this nature. I will therefore point to some potential methodological remedies and will both sketch previous findings and report novel results that imply more careful conclusions concerning the use of heuristics, especially the fluency heuristic.

In addition, the view that decision makers pervasively use heuristics has been based on a rather selective set of investigations, whereas other studies—admittedly less supportive of fast-and-frugal heuristics—have typically not been considered. For example, Gigerenzer and Brighton (2009) mentioned only a single critical investigation of the recognition heuristic (which they dismissed), whereas Marewski, Gaissmaier, and Gigerenzer (2010) did not explicitly spell out a single one. Also, both cited previous work as confirming predictions of the recognition heuristic (e.g., Newell & Fernandez, 2006; Pohl, 2006), although the original authors came to less favorable conclusions. A brief overview of the more critical studies on the recognition heuristic not mentioned by Marewski, Gaissmaier, and Gigerenzer or Gigerenzer and Brighton will therefore be provided. Finally, I will point the reader to some of the literature concerning other heuristics, which, to put it carefully, is not entirely favorable for the fast-and-frugal heuristics approach. Note that these overviews will focus explicitly on critical work.

A New Fast-and-Frugal Heuristic

For illustration of what caveats might pertain to investigations of simple strategies, let us consider a new fast-and-frugal heuristic: the *alphabet heuristic* (AH), which

B. E. Hilbig, hilbig@psychologie.uni-mannheim.de

operates in the domain of city-size comparisons between large world cities. It is, literally speaking, a lexicographic judgment strategy, since it comprises the following rules:

Search rule. Search through object names—letter by letter in reading direction—and assess each letter's position in the alphabet.

Stopping rule. Stop search if the letters at the currently searched serial position differ between the cities.

Decision rule. Infer that the city yielding the letter appearing later in the alphabet has more inhabitants.

For example, consider the comparative city-size judgment between Durban (South Africa) and Seoul (South Korea). According to the search rule of the AH, one would initially compare the first letters in each of the city names—that is, D and S. Since they have different positions in the alphabet, the stopping rule applies, and the decision rule implies the choice of Seoul, because S appears later in the alphabet than D. As this also reveals, the AH is a single-cue noncompensatory rule like many other fast-and-frugal heuristics, thus yielding comparable simplicity. Indeed, the AH uses a cue that is even earlier "on the mental stage" (Pachur & Hertwig, 2006) and, thus, even more easily applicable than recognition.

At the same time, the AH does not share many of the theoretical limitations of other fast-and-frugal heuristics (e.g., Dougherty, Franco-Watkins, & Thomas, 2008); for example, unlike take-the-best, it does not require learning or knowing the optimal cue order to make good judgments (but see Gigerenzer, Hoffrage, & Goldstein, 2008). Also, unlike the recognition heuristic, it does not necessitate simplifications such as treating recognition as binary (Erdfelder, Küpper-Tetzel, & Mattern, in press; Newell & Fernandez, 2006). Finally, the AH can be used for every possible pair of objects (see below). It thus avoids what has been called the strategy selection problem (e.g., Glöckner & Betsch, 2010; Newell, 2005)—that is, the unsolved riddle of *how* decision makers will know *when* to switch to *which* heuristic.

Is it often possible and ecologically rational to rely on the AH? In principle, two criteria are relevant for the usefulness of a heuristic in a given domain. One is how often the cue on which this heuristic relies discriminates between pairs of objects, or, stated differently, the proportion of comparative judgments in which the heuristic can be applied. This is known as the *discrimination rate* (e.g., Gigerenzer & Goldstein, 1996). Second, considering only the cases in which the heuristic can be used, how often does it predict a correct judgment with respect to the judgment criterion? In other words, this is the proportion of correct judgments achievable by following the heuristic whenever possible. This criterion is denoted *validity*. Finally, by multiplying discrimination rate and validity, one assesses a heuristic's *success rate* (Newell, Rakow, Weston, & Shanks, 2004).

To compare the discrimination rate, validity, and success of the AH with those of the recognition and the fluency heuristics, data on the set of the 61 largest world cities (Wikipedia, n.d.) were collected. A total of 29 participants (otherwise participating in unrelated studies) were presented with each of these 61 cities (separately and in random order) and were asked to indicate, as speedily as possible, whether they recognized a city or not by pressing one of two keys. Participants' recognition judgments and the corresponding latencies were used to compute the discrimination rate, validity, and success rate of the recognition heuristic and fluency heuristic, respectively (Hertwig, Herzog, Schooler, & Reimer, 2008). The results and comparison with the AH can be found in Table 1.

Table 1
Discrimination Rate (DR), Validity, and Success Rate for Each of the Heuristics Separately As Well As for the Recognition and Fluency Heuristics Taken Together

	DR	Validity	Success
Alphabet heuristic	1.00	.57	.57
Recognition heuristic	.45	.64	.29
Fluency heuristic	.29	.53	.16
Recognition and fluency heuristics	.74	.60	.44

Note—The values for the recognition and fluency heuristics are means across the 29 participants.

As can be seen, the AH had the largest discrimination rate. That is, it made predictions for all cases and, thereby, more than the recognition and fluency heuristics together. At the same time, its validity was slightly below the corresponding value for the recognition heuristic and marginally above that for the fluency heuristic. Taken together, the latter two outperformed the AH by a validity of .03. However, in terms of success, the AH was clearly superior, outperforming both the recognition and fluency heuristics, even when the latter two were considered together. In sum, these analyses show that the AH is at least as useful as the recognition and fluency heuristics.

Is the AH used? As was shown previously, the AH qualifies as a fast-and-frugal heuristic from a normative perspective. The second question then concerns its descriptive adequacy—that is, whether decision makers *use* this heuristic. To this end, I reanalyzed the data of Hilbig and Pohl's (2009) Experiment 2, in which 74 participants performed the city-size task on a random selection of 17 large world cities. In this set, the AH actually had a validity of .68, which is relatively large, although once more below the mean recognition validity (.78). To assess the heuristic's use, the proportion of choices in line with its predictions was computed—that is, the adherence or accordance rate (e.g., Hertwig et al., 2008; Pachur & Hertwig, 2006). Across participants, the AH explained $M = 60\%$ ($SE = 1\%$) of choices. This is highly comparable to what has been reported for the fluency heuristic (Hertwig et al., 2008). The recognition heuristic, by comparison, showed a clearly superior adherence rate of $M = 86\%$ ($SE = 2\%$) in those cases for which it made a prediction. However, this is close to comparing apples and oranges, since the recognition heuristic often makes no predictions. Across *all* choices, the recognition heuristic explained no more than $M = 42\%$ ($SE = 1\%$), which shows that, when it comes to explaining as much behavior as possible, the AH is actually superior, and significantly so [$t(73) = 11.5, p < .001$, Cohen's $d = 1.33$]. From the point of view of Occam's razor, these findings actually imply removing the recognition and fluency heuristics from the adaptive toolbox and replacing them by the AH, which affords fewer theoretical assumptions/simplifications, is at least as well specified, and predicts "more" behavior correctly.

Pitfalls and caveats in the analysis of choice data. Should we now celebrate a new fast-and-frugal heuristic? As the reader will have noticed, there is a hitch. Indeed, there are several. First and foremost, the reported results suffer from a simple confound: In the AH, we have, by definition, a cue that is correlated with the criterion. In other words, correct judgments are more likely in line with the cue's prediction than not. At the same time, we have decision makers who achieve above-chance-level performance in their judgments. Consequently, their choice data will inevitably reflect accordance to our cue. However, they may have achieved their performance by considering a completely different piece of information than the AH proposes—for example, which cities have international airports. Thus, the very precondition for a good heuristic (that its central cue be valid) can easily produce high accordance even if no single participant ever actually considered the cue. Clearly, in the data set analyzed above, this is most likely the case, since I seriously doubt that our participants actually compared cities by considering their alphabetical rank.

So, what do the well-known figures showing high adherence rates to fast-and-frugal heuristics for a majority of decision makers tell us (e.g., Gigerenzer & Brighton, 2009, Figure 7; Goldstein & Gigerenzer, 2002, Figure 4; Hertwig et al., 2008, Figure 4; Marewski, Gaissmaier, Schooler, Goldstein, & Gigerenzer, 2009, Figure 3, and 2010, Figure 3; Pachur, Bröder, & Marewski, 2008, Figure 5; Reimer & Katsikopoulos, 2004, Figure 3)? As the example of the AH has illustrated, very little can and should be concluded concerning the actual use of heuristics on the basis of such measures. Adherence rates can be high even though a cue was *never* even taken into account by participants. Reasonably, it should be noted that proponents of fast-and-frugal heuristics rarely explicitly claim that adherence rates are sufficient to conclude that a specific cue was considered.[1] As such, I do not intend to suggest that they would commit this formal fallacy of affirming the consequent. However, to the extent that we are in agreement on this limitation of adherence rates, more conclusive findings that corroborate *pervasive* reliance on certain cues (such as recognition or retrieval fluency) are still pending.

Worse yet, even if a cue in question was actually considered by decision makers, does following a cue's prediction provide evidence for noncompensatory reliance on this cue alone? Once more, this cannot be deduced so long as the cue in question and a varyingly large number of potential additional cues make the same prediction: In the example above, a participant selecting Seoul as the more populous city (as compared with Durban) could actually have relied on the alphabetical rank cue. However, he or she might have *additionally* considered the capital city cue (Seoul being the capital of South Korea, whereas Durban is not the capital of South Africa) or the international airport cue (Seoul has two international airports, whereas Durban has only one). This list could be continued with many cues. As the example shows, one would have to ensure that all other cues point in a different direction than the AH (for examples, see Bröder & Eichler, 2006; Hilbig, 2008b); only then would high accordance with the AH's predictions become informative. Fortunately for the AH and unfortunately for us, such control over all other cues is impossible if objects (e.g., cities) are known to participants from their real-world experiences (as advocated by Pachur et al., 2008, for studying the recognition heuristic).

Remedies and Some Empirical Findings

A proxy for use of heuristics. Is a cue in question relied on in isolation, and for what proportion of decision makers may this be the case? A crude test of whether a cue may have been considered in a noncompensatory fashion is the discrimination index (DI; Hilbig & Pohl, 2008). Although the DI also has several limitations (Hilbig, 2010), it is the simplest proxy available so far that can be applied to any heuristic so long as inferences can be classified as correct versus false. For a similar approach, see the d' measure proposed by Pachur and Hertwig (2006). The DI is defined as the difference in adherence rates between cases in which a heuristic implies a correct versus a false inference. If decision makers rely on a single cue in isolation, they cannot discriminate such cases; that is, the difference in adherence rates (the DI) must be close to zero—a *necessary* condition. Vice versa, a DI reliably different from zero is *sufficient* for nonusers of the heuristic in question (cf. Hilbig, 2008a). So, to assess whether our participants relied on the alphabetical rank cue in isolation, the data above were analyzed using the DI.

As one might expect, participants often adhered to the AH whenever it made a correct prediction ($M = 72\%$, $SE = 1\%$) but refrained from doing so whenever the AH implied a false inference ($M = 35\%$, $SE = 1\%$), thus resulting in DI = .37, on average. Clearly, the participants must have considered some other cue or information beyond objects' alphabetical rank to show this pattern. It can thereby be ruled out that decision makers used the AH. As an additional advantage, the DI can be computed for each participant—satisfying the call for data analyses at the individual level, often reiterated by proponents of fast-and-frugal heuristics (e.g., Gigerenzer & Brighton, 2009). In the case of the data above on the AH, only 2 participants had a DI within the 95% confidence interval of zero. Only these 4% of the sample may have used the AH. All others (96%), by contrast, can conclusively be considered nonusers of the AH.

As these results reveal, the AH's assumption of noncompensatory decision making must be rejected. However, unmentioned by Marewski, Gaissmaier, and Gigerenzer (2010) or Gigerenzer and Brighton (2009), similar, although less extreme, findings have been reported for the recognition heuristic: Across several experiments, a minority of participants potentially used the recognition heuristic, showing DI scores close to zero. All others, by contrast, can conclusively be considered recognition heuristic nonusers (Hilbig & Pohl, 2008; Hilbig, Pohl, & Bröder, 2009). It should be noted, however, that the proportion of participants who potentially used the recognition heuristic was always substantial (~40%). As an aside, and related to Marewski, Gaissmaier, and Gigerenzer's (2010) general claim that "good judgments do not require complex cognition" (p. 103), participants conclusively identified as nonusers of the recognition heuristic achieved *more* correct judgments than did potential recognition heuristic users (Hilbig & Pohl, 2008). Stated differently, if some participants actually did apply the recognition heuristic, they were outperformed by those who considered some information beyond recognition.

New findings on the fluency heuristic. Unfortunately, no comparable analyses have yet been reported for the fluency heuristic. As such, although some findings clearly point to an influence of fluency on judgments (Hertwig et al., 2008), I am not aware of any study testing the one-reason decision-making aspect of the fluency heuristic. Therefore, I will briefly report reanalyses of an experiment from Hilbig and Pohl (2009, Experiment 3) that closely resembled those previously conducted by Hertwig et al. to study the fluency heuristic. It comprised 68 participants performing the city-size task on the set of the 14 largest cities in Switzerland. On average, fluency (as approximated by recognition latencies for objects' names) was a very valid cue ($M = .72$, $SE = .02$), and participants adhered to the fluency cue with $M = .68$ ($SE = .02$), which strongly resembles previous findings (Hertwig et al., 2008). However, computing the DI yielded rather different results: Whereas decision makers clearly adhered to the fluency heuristic's prediction whenever it implied a correct inference ($M = .78$, $SE = .02$), they refrained from doing so whenever it would have led to a false judgment ($M = .42$, $SE = .03$). Thus, mirroring the results for the AH sketched above, the average DI ($M = .36$, $SE = .03$) was large and significantly different from zero [$t(64) = 10.4$, $p < .001$, Cohen's $d = 1.3$].

Once more focusing on the individual-level, analyses revealed that no more than 9 out of 65 (13.8%) participants had a DI score within the 95% confidence interval of zero. This minority may have used the fluency heuristic. Vice versa, 86% of the sample can conclusively be considered nonusers of the fluency heuristic. Finally, nonusers of the fluency heuristic were substantially more likely to make correct judgments in cases in which the fluency heuristic was applicable ($r = .81$, $p < .001$, between individual DI scores and the proportion of correct inferences), which held when controlling for individual fluency validity ($r_p = .82$, $p < .001$). In sum, only a few participants qualified as potential fluency heuristic users, and they were outperformed by nonusers concerning judgmental accuracy.

Review of the "Other" Literature

The analyses previously reported are by no means the only instances that cast a skeptical light on (some) fast-and-frugal heuristics. On the contrary, there are several well-designed investigations—comprising much cleverer approaches than the DI—which have tested the recognition heuristic, the take-the-best heuristic, and others. Unfortunately, Gigerenzer and Brighton (2009) and Marewski, Gaissmaier, and Gigerenzer (2010) did not consider them in their reviews or focused on those results compatible with the fast-and-frugal heuristics approach.[2] For the sake of brevity, I will merely sketch critical findings pertaining to the recognition heuristic and point only to critical investigations of other heuristics. Especially the latter set will thus necessarily remain incomplete. Note that the following studies have not been considered equally adequate tests of the recognition heuristic (Pachur et al., 2008); however, at the very least, they all contribute to understanding its bounding conditions and thus deserve to be considered. Moreover, several present noteworthy counterevidence against the assumption of noncompensatory processing and one-reason decision making.

Overview of critical findings on the recognition heuristic. With respect to choices, decision makers do not follow the recognition cue if they have criterion knowledge arguing against it or if their recognition of objects stems from a source unrelated to the criterion dimension (Oppenheimer, 2003). As such, there are cases in which recognition information is overruled. When provided with additional cues beyond recognition, decision makers take these further cues into consideration, especially when recognition is not a highly valid cue (Newell & Shanks, 2004). In inferences from memory, the pattern of cues, in addition to recognition, determines decision makers' choices even though all further cues should be irrelevant (Bröder & Eichler, 2006). Enhanced replications in Goldstein and Gigerenzer (2002, Experiment 2) show that an additional successful cue can overrule the recognition cue (Newell & Fernandez, 2006, Experiment 1; Richter & Späth, 2006, Experiment 3). As all these examples show, the recognition cue is not generally considered in isolation; rather, "people appear to use more than one reason" (Hardman, 2009, p. 15).

Further research has tested the claim of noncompensatory reliance on recognition without providing participants with additional cues (in line with the arguments put forward by Pachur et al., 2008). However, the findings remain challenging for the recognition heuristic: As was hinted at previously, a majority of decision makers follow the recognition cue selectively if it leads to a correct inference (Hilbig & Pohl, 2008; Pohl, 2006). Again, this result holds when criterion knowledge is controlled for (Hilbig et al., 2009). In fact, proponents of fast-and-frugal heuristics have reported similar findings (see the Appendix in Pachur & Hertwig, 2006). Clearly, the recognition heuristic fails to explain how decision makers can discriminate between cases in which the recognition cue implies a correct decision and those cases in which it does not—given that the recognition cue is supposed to be relied on in isolation. Similarly, the recognition heuristic cannot explain why switching the direction of the cue–criterion relation (by asking half of the participants which object in each pair scores *lower* on the criterion dimension) leads to choices less in line with its predictions (McCloy, Beaman, Frosch, & Goddard, 2010).

A somewhat different line of research demonstrated that the binary quality of recognition as implemented in the recognition heuristic theory is an oversimplification. As such, adherence to the recognition heuristic depends on whether there is further knowledge available for the recognized object (Pohl, 2006; Snook & Cullen, 2006) and on how speedily it is recognized (Hertwig et al., 2008; Newell & Fernandez, 2006, Experiment 2). Note that the former result holds when participants' criterion knowledge is controlled for (Hilbig et al., 2009). Of course, these findings do not necessarily conflict with the idea of noncompensatory reliance on recognition; however, to be compatible with this notion, additional assumptions about memory processes must be made (Erdfelder et al., in press), thus inflating theoretical complexity.

Concerning the process level, several investigations have analyzed response time data. As an early result, it has been shown that decisions need more time when recognition and further knowledge conflict (Richter & Späth, 2006), which points toward information integration. Indeed, response time patterns across three experiments were explained better by alternative compensatory models—even when controlling for recognition times (Hilbig & Pohl, 2009). Especially, results mostly conflicted with the serial process predictions of the recognition heuristic. These findings were recently replicated and extended to show that task stimuli congruity predicts response time data, whereas the recognition heuristic cannot (Frosch, McCloy, Beaman, & Goddard, 2010). Other researchers have also found that response time data is poorly explained by the stepwise serial process assumptions of the recognition heuristic (Schweickart, Brown, & Lee, 2009).

Finally, some critical formal modeling approaches and model comparisons should be mentioned. When recognition heuristic use is estimated through the *r* model, an unbiased measurement model from the class of multinomial processing tree models (Erdfelder et al., 2009), the recognition heuristic cannot adequately describe empirical choice data (from 10 data sets), not even when implemented in a probabilistic fashion (Hilbig, Erdfelder, & Pohl, 2010). Model comparisons using the multiple-measure maximum-likelihood strategy classification method (Glöckner, 2009; Jekel, Nicklisch, & Glöckner, 2010) show that automatic compensatory processes outperform the recognition heuristic in explaining choices, response latencies, and confidence judgments (Glöckner & Bröder, in press). Recent evidence additionally suggests that two-alternative forced choice judgments are likely to stem from the same underlying processes as continuous judgments for which multiple cues (rather than recognition alone) are necessary (Zhao & Oppenheimer, 2010).

Importantly, none of the authors cited above claimed that the recognition cue is never considered. Nor did they reject the possibility that decisions may more or less often reflect noncompensatory reliance on recognition. Finally, few would have questioned that the recognition heuristic can adequately describe the behavior of some decision makers at times. I am merely suggesting bearing in mind and appropriately acknowledging the evidence sketched above, which, interpreted leniently, can be summarized as follows: No matter how ecologically rational the recognition heuristic often is, pervasive reliance on this strategy is observable only under a very limited set of circumstances. Moreover, whereas the recognition heuristic is indubitably a satisfactory *as-if* model (often achieving high predictive accuracy), most tests of the underlying *processes* have refuted its predictions, so far.

Other critical investigations. It is beyond the scope of this note to provide an overview of the many investigations pertaining to fast-and-frugal heuristics. However, readers should at least be pointed to several critical and often methodologically sound pieces of work. At least concerning the take-the-best heuristic, "well-designed experiments have shown that it is far from a universal heuristic" (Dougherty et al., 2008, p. 209), and several interesting pieces of work are worthy of mention (e.g., Ayal & Hochman, 2009; Bröder, 2000; Bröder & Newell, 2008; Glöckner & Betsch, 2008c, 2010; Glöckner, Betsch, & Schindler, 2010; Glöckner & Moritz, 2009; Hausmann & Läge, 2008; Lee & Cummins, 2004; Newell, Collins, & Lee, 2007; Newell & Lee, in press; Newell & Shanks, 2003; Newell, Weston, & Shanks, 2003). On the basis of this literature, I am skeptical of Marewski, Gaissmaier, and Gigerenzer's (2010) assertion that "*in general*, people tend to make inferences consistent with take-the-best" (p. 112; emphasis added). This is a rather surprising conclusion, given that, across several experiments—including conditions strongly fostering take-the-best use—Bröder and Gaissmaier (2007) found that fewer than 40% of their almost 500 participants made choices most likely produced by the take-the-best heuristic. At the same time, the authors did report a higher proportion of take-the-best users, and it would thus appear more appropriate to conclude that under certain conditions that hamper automatic processing, impose severe information acquisition costs, and/or necessitate effortful retrieval of information from memory, choices tend to resemble use of take-the-best (see also Bröder & Newell, 2008; Glöckner & Betsch, 2008c).

Of course, there are (theoretically numerous) other heuristics, thereby rendering counterevidence to single models less problematic for the entire fast-and-frugal heuristics approach.[3] However, I am aware of few critical and methodologically sound tests that provide evidence for pervasive reliance on strategies such as the fluency heuristic. On the other hand, counterevidence sometimes appears to be quite overwhelming, as the example of the priority heuristic (Brandstätter, Gigerenzer, & Hertwig, 2006) shows (Ayal & Hochman, 2009; Birnbaum, 2008a, 2008b; Birnbaum & LaCroix, 2008; Fiedler, 2010; Glöckner & Betsch, 2008a; Glöckner & Herbold, in press; Hilbig, 2008b; Johnson, Schulte-Mecklenbeck, & Willemsen, 2008; Rieskamp, 2008). A similar point can be made for the quick estimation heuristic (Hausmann, Läge, Pohl, & Bröder, 2007). So, overall "it seems fair to conclude that strict empirical tests have resulted in a more critical picture of the . . . postulated heuristics" (Fiedler, 2010, p. 22).

Conclusions

The criticism sketched above notwithstanding, I concur with Marewski, Gaissmaier, and Gigerenzer (2010) and Gigerenzer and Brighton (2009) in several ways. Above all, few would deny that there exist judgment cues, such as recognition, which are often highly valid and very useful. It is also undeniable that simple strategies need not always yield outcomes inferior to those of more complex algorithms (Hogarth & Karelaia, 2007). Moreover, observable choices can indeed *resemble* noncompensatory reliance on single, highly valid cues (as demonstrated above). However, this cannot and should not be considered evidence for the use of single-cue heuristics. By contrast, a whole host of research has convincingly shown that evidence against the pervasive application of such heuristics is robust, replicable, and observable, using many different experimental designs, measures, and statistical analyses.

In an attempt to reach a balanced conclusion, one might say that *some* fast-and-frugal heuristics indeed provide adequate descriptions of decision makers' behavior at times. As outlined above, there are conditions under which the take-the-best heuristic predicts behavior well (Bröder & Newell, 2008). Indeed, for some decision makers, it has even received confirming evidence on the process level (Bröder & Gaissmaier, 2007). Likewise, across several investigations, about 40% of participants have been identified as potential users of the recognition heuristic (e.g., Hilbig & Pohl, 2008; Pachur et al., 2008). As such, there is always a substantial proportion of choices resembling noncompensatory reliance on recognition as proposed by the recognition heuristic (Hilbig, Erdfelder, & Pohl, 2010). However, on the process level, alternative compensatory models seem to provide a better explanation of how these choices come about (Glöckner & Bröder, in press; Hilbig & Pohl, 2009). In sum, evidence for the take-the-best and the recognition heuristics is mixed, and this should be openly acknowledged in reviews pertaining to fast-and-frugal heuristics.

On the other hand, there is virtually no evidence for models such as the fluency heuristic, the priority heuristic, or the quick estimation heuristic. Especially in the case of the fluency heuristic, this is problematic, since it has been counted among the "well-studied heuristics for which there is *evidence* that they are in the adaptive toolbox of humans" (Gigerenzer & Brighton, 2009, p. 130, emphasis added). As the results reported herein indicate, use of this heuristic could be conclusively ruled out for more than 85% of decision makers. Of course, this does not challenge the fact that retrieval fluency can, at times, be a valid cue. It simply suggests not considering the fluency heuristic a strategy actually used by decision makers, at least not on the basis of above-chance-level validity and adherence rates.

One more general point should be made: Like several others, I partially agree with Marewski, Gaissmaier, and Gigerenzer (2010) that "good judgments do not require complex cognition"; however, this need not be due to pervasive reliance on simple heuristics. Their conclusion critically hinges on the assumption that more complex strategies actually necessitate more effort (Newell & Bröder, 2008). The issue of effort, in turn, applies only if strategies are used deliberatively (Hilbig, Scholl, & Pohl, 2010). However, there is a large and diverse body of literature arguing for automatic processes in judgment and decision making (e.g., Busemeyer & Johnson, 2004; Busemeyer & Townsend, 1993; Dougherty, Gettys, & Odgen, 1999; Glöckner, 2008; Glöckner & Betsch, 2008b; Glöckner & Herbold, in press; Holyoak & Simon, 1999; Juslin & Olsson, 2004; Juslin & Persson, 2002; for a recent review, see also Glöckner & Witteman, 2010), which allow for information integration without imposing severe information processing costs. In this very vein, Newell and Bröder have concluded that "complex process[es] do not necessarily imply the consumption of conscious resources or much processing time and viewed from this perspective, 'simple' heuristics are probably not much simpler, subjectively than complex ones" (2008, p. 201). If decision makers can do better by relying on knowledge beyond single cues (see, e.g., Hilbig & Pohl, 2008, as well as the analyses of the fluency heuristic reported above) and given that this does not necessarily cost them more cognitive effort, it may be time to look beyond simple heuristics to see how smart we really are.

AUTHOR NOTE

I thank Rüdiger F. Pohl, Morten Moshagen, and Andreas Glöckner for important discussions and helpful comments on earlier versions of the manuscript. Also, thanks are due Klaus Fiedler, Philip Beaman (and colleagues), and Jiaying Zhao (and Daniel Oppenheimer) for sharing their work in progress. Correspondence concerning this article should be addressed to B. E. Hilbig, Psychology III, University of Mannheim, Schloss Ehrenhof Ost, 68131 Mannheim, Germany (e-mail: hilbig@psychologie.uni-mannheim.de).

Note—Accepted by Cathleen M. Moore's editorial team.

REFERENCES

Ayal, S., & Hochman, G. (2009). Ignorance or integration: The cognitive processes underlying choice behavior. *Journal of Behavioral Decision Making*, **22**, 455-474.

Birnbaum, M. H. (2008a). Evaluation of the priority heuristic as a descriptive model of risky decision making: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, **115**, 253-260.

Birnbaum, M. H. (2008b). New tests of cumulative prospect theory and the priority heuristic: Probability-outcome tradeoff with branch splitting. *Judgment & Decision Making*, **3**, 304-316.

Birnbaum, M. H., & LaCroix, A. R. (2008). Dimension integration: Testing models without trade-offs. *Organizational Behavior & Human Decision Processes*, **105**, 122-133.

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

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, & Cognition*, **26**, 1332-1346.

Bröder, A., & Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica*, **121**, 275-284.

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

Bröder, A., & Newell, B. R. (2008). Challenging some common beliefs: Empirical work within the adaptive toolbox metaphor. *Judgment & Decision Making*, **3**, 205-214.

Busemeyer, J. R., & Johnson, J. G. (2004). Computational models of decision making. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 133-154). Malden, MA: Blackwell.

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.

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**, 199-213.

DOUGHERTY, M. R., GETTYS, C. F., & ODGEN, E. E. (1999). MINERVA-DM: A memory process model for judgments of likelihood. *Psychological Review*, **106**, 180-209.

ERDFELDER, E., AUER, T.-S., HILBIG, B. E., ASSFALG, A., MOSHAGEN, M., & NADAREVIC, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie*, **217**, 108-124.

ERDFELDER, E., KÜPPER-TETZEL, C. E., & MATTERN, S. (in press). Threshold models of recognition and the recognition heuristic. *Judgment & Decision Making*.

FIEDLER, K. (2010). How to study cognitive decision algorithms: The case of the priority heuristic. *Judgment & Decision Making*, **5**, 21-32.

FROSCH, C. A., MCCLOY, R., BEAMAN, C. P., & GODDARD, K. (2010). *Time to decide: Frugality vs. congruity in comparative judgment*. Manuscript submitted for publication.

GIGERENZER, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, **3**, 20-29.

GIGERENZER, G., & BRIGHTON, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, **1**, 107-143.

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

GIGERENZER, G., HOFFRAGE, U., & GOLDSTEIN, D. G. (2008). Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty, Franco-Watkins, and Thomas (2008). *Psychological Review*, **115**, 230-237.

GLÖCKNER, A. (2008). How evolution outwits bounded rationality: The efficient interaction of automatic and deliberate processes in decision making and implications for institutions. In C. Engel & W. Singer (Eds.), *Better than conscious? Implications for performance and institutional analysis* (pp. 259-284). Cambridge, MA: MIT Press.

GLÖCKNER, A. (2009). Investigating intuitive and deliberate processes statistically: The multiple-measure maximum likelihood strategy classification method. *Judgment & Decision Making*, **4**, 186-199.

GLÖCKNER, A., & BETSCH, T. (2008a). Do people make decisions under risk based on ignorance? An empirical test of the priority heuristic against cumulative prospect theory. *Organizational Behavior & Human Decision Processes*, **107**, 75-95.

GLÖCKNER, A., & BETSCH, T. (2008b). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment & Decision Making*, **3**, 215-228.

GLÖCKNER, A., & BETSCH, T. (2008c). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **34**, 1055-1075.

GLÖCKNER, A., & BETSCH, T. (2010). Accounting for critical evidence while being precise and avoiding the strategy selection problem in a parallel constraint satisfaction approach: A reply to Marewski (2010). *Journal of Behavioral Decision Making*, **23**, 468-472.

GLÖCKNER, A., BETSCH, T., & SCHINDLER, N. (2010). Coherence shifts in probabilistic inference tasks. *Journal of Behavioral Decision Making*, **23**, 439-462.

GLÖCKNER, A., & BRÖDER, A. (in press). Processing of recognition information and additional cues: A model-based analysis of choice, confidence, and response time. *Judgment & Decision Making*.

GLÖCKNER, A., & HERBOLD, A.-K. (in press). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*.

GLÖCKNER, A., & MORITZ, S. (2009). A fine-grained analysis of the jumping-to-conclusions bias in schizophrenia: Data-gathering, response confidence, and information integration. *Judgment & Decision Making*, **4**, 587-600.

GLÖCKNER, A., & WITTEMAN, C. (2010). Beyond dual-process models: A categorization of processes underlying intuitive judgment and decision making. *Thinking & Reasoning*, **16**, 1-25.

GOLDSTEIN, D. G., & GIGERENZER, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 37-58). New York: Oxford University Press.

GOLDSTEIN, D. G., & GIGERENZER, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, **109**, 75-90.

GOLDSTEIN, D. G., & GIGERENZER, G. (2009). Fast and frugal forecasting. *International Journal of Forecasting*, **25**, 760-772.

HARDMAN, D. (2009). *Judgment and decision making*. Malden, MA: Blackwell.

HAUSMANN, D., & LÄGE, D. (2008). Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgment & Decision Making*, **3**, 229-243.

HAUSMANN, D., LÄGE, D., POHL, R. F., & BRÖDER, A. (2007). Testing quickEst: No evidence for the quick-estimation heuristic. *European Journal of Cognitive Psychology*, **19**, 446-456.

HERTWIG, R., HERZOG, S. M., SCHOOLER, L. J., & REIMER, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **34**, 1191-1206.

HILBIG, B. E. (2008a). Individual differences in fast-and-frugal decision making: Neuroticism and the recognition heuristic. *Journal of Research in Personality*, **42**, 1641-1645.

HILBIG, B. E. (2008b). One-reason decision making in risky choice? A closer look at the priority heuristic. *Judgment & Decision Making*, **3**, 457-462.

HILBIG, B. E. (2010). Precise models deserve precise measures: A methodological dissection. *Judgment & Decision Making*, **5**, 272-284.

HILBIG, B. E., ERDFELDER, E., & POHL, R. F. (2010). One-reason decision-making unveiled: A measurement model of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **36**, 123-134.

HILBIG, B. E., & POHL, R. F. (2008). Recognizing users of the recognition heuristic. *Experimental Psychology*, **55**, 394-401.

HILBIG, B. E., & POHL, R. F. (2009). Ignorance- versus evidence-based decision making: A decision time analysis of the recognition heuristic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **35**, 1296-1305.

HILBIG, B. E., POHL, R. F., & BRÖDER, A. (2009). Criterion knowledge: A moderator of using the recognition heuristic? *Journal of Behavioral Decision Making*, **22**, 510-522.

HILBIG, B. E., SCHOLL, S. G., & POHL, R. F. (2010). Think or blink—Is the recognition heuristic an "intuitive" strategy? *Judgment & Decision Making*, **5**, 300-309.

HOGARTH, R. M., & KARELAIA, N. (2007). Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review*, **114**, 733-758.

HOLYOAK, K. J., & SIMON, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General*, **128**, 3-31.

JEKEL, M., NICKLISCH, A., & GLÖCKNER, A. (2010). Implementation of the multiple-measure maximum likelihood strategy classification method in R: Addendum to Glöckner (2009) and practical guide for application. *Judgment & Decision Making*, **5**, 54-63.

JOHNSON, E. J., SCHULTE-MECKLENBECK, M., & WILLEMSEN, M. C. (2008). Process models deserve process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, **115**, 263-272.

JUSLIN, P., & OLSSON, H. (2004). Note on the rationality of rule-based versus exemplar-based processing in human judgment. *Scandinavian Journal of Psychology*, **45**, 37-47.

JUSLIN, P., & PERSSON, M. (2002). PROBabilities from EXemplars (PROBEX): A "lazy" algorithm for probabilistic inference from generic knowledge. *Cognitive Science*, **26**, 563-607.

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

MAREWSKI, J. N., GAISSMAIER, W., & GIGERENZER, G. (2010). Good judgments do not require complex cognition. *Cognitive Processing*, **11**, 103-121.

MAREWSKI, J. N., GAISSMAIER, W., SCHOOLER, L. J., GOLDSTEIN, D. G., & GIGERENZER, G. (2009). Do voters use episodic knowledge to

rely on recognition? In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2232-2237). Austin, TX: Cognitive Science Society.
Marewski, J. N., Gaissmaier, W., Schooler, L. J., Goldstein, D. G., & Gigerenzer, G. (2010). From recognition to decisions: Extending and testing recognition-based models for multi-alternative inference. *Psychonomic Bulletin & Review*, **17**, 287-309.
McCloy, R., Beaman, C. P., Frosch, C. A., & Goddard, K. (2010). Fast and frugal framing effects? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **36**, 1043-1052.
Newell, B. R. (2005). Re-visions of rationality? *Trends in Cognitive Sciences*, **9**, 11-15.
Newell, B. R., & Bröder, A. (2008). Cognitive processes, models and metaphors in decision research. *Judgment & Decision Making*, **3**, 195-204.
Newell, B. R., Collins, P., & Lee, M. D. (2007). Adjusting the spanner: Testing an evidence accumulation model of decision making. In D. McNamara & G. Trafton (Eds.), *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 533-538). Austin, TX: Cognitive Science Society.
Newell, B. R., & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentially of further knowledge: Two critical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, **19**, 333-346.
Newell, B. R., & Lee, M. D. (in press). The right tool for the job? Comparing an evidence accumulation and a naive strategy selection model of decision making. *Journal of Behavioral Decision Making*.
Newell, B. R., Rakow, T., Weston, N. J., & Shanks, D. R. (2004). Search strategies in decision making: The success of "success." *Journal of Behavioral Decision Making*, **17**, 117-137.
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, & Cognition*, **29**, 53-65.
Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision making. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 923-935.
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 & Human Decision Processes*, **91**, 82-96.
Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition*, **90**, B1-B9.
Pachur, T., Bröder, A., & Marewski, J. (2008). The recognition heuristic in memory-based inference: Is recognition a non-compensatory cue? *Journal of Behavioral Decision Making*, **21**, 183-210.
Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **32**, 983-1002.
Pohl, R. F. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, **19**, 251-271.
Reimer, T., & Katsikopoulos, K. V. (2004). The use of recognition in group decision-making. *Cognitive Science*, **28**, 1009-1029.
Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **32**, 150-162.
Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **34**, 1446-1465.
Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, **112**, 610-628.
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.
Snook, B., & Cullen, R. M. (2006). Recognizing National Hockey League greatness with an ignorance-based heuristic. *Canadian Journal of Experimental Psychology*, **60**, 33-43.
Wikipedia (n.d.). List of cities proper by population. Retrieved October 2009 from http://en.wikipedia.org/wiki/List_of_cities_proper_by_population.
Zhao, J., & Oppenheimer, D. M. (2010). *Beyond binary: Limitations of the binary choice paradigm for studying judgment heuristics*. Manuscript submitted for publication.

NOTES

1. Nonetheless, adherence rates have been presented in sections headed "Do People *Use* the Recognition Heuristic?" (Goldstein & Gigerenzer, 1999, p. 50, emphasis added). Likewise, one can come across the figure caption "Systematic individual differences exist in the *use* of heuristics" (Gigerenzer & Brighton, 2009, p. 135, emphasis added) placed under a graph showing adherence rates. Also, Pachur and Hertwig (2006) concluded from (cross-study) differences in adherence rates that "time pressure . . . fostered the *use* of the recognition heuristic" (p. 994, emphasis added). At least to readers less concerned with methodological issues in judgment and decision-making research, these examples would most probably signify that high adherence implies use of a heuristic.

2. For example, Pohl (2006) not only showed that adherence to the recognition heuristic depends on its ecological validity—which, as an aside, would have been predicted by many alternative models, too—but also found that the less-is-more effect predicted by the recognition heuristic is absent, and that participants consider further information to discriminate whether the recognition heuristic implies a correct inference (see above). The latter findings conflict with the recognition heuristic, but only the former—which is compatible with the fast-and-frugal heuristics approach—is typically reiterated (e.g., Gigerenzer, 2008; Gigerenzer & Brighton, 2009; Marewski, Gaissmaier, & Gigerenzer, 2010).

3. This fact has previously attracted criticism (Dougherty et al., 2008), and Fiedler (2010) states that the fast-and-frugal heuristics program "has to go beyond the truism that many heuristics can explain many behaviors" (p. 27).

(Manuscript received January 18, 2010;
revision accepted for publication March 12, 2010.)