Fast and Frugal Heuristics Are Plausible Models of Cognition: Reply to Dougherty, Franco-Watkins, and Thomas (2008)

Gerd Gigerenzer
Max Planck Institute for Human Development

Ulrich Hoffrage University of Lausanne

Daniel G. Goldstein London Business School

M. R. Dougherty, A. M. Franco-Watkins, and R. Thomas (2008) conjectured that fast and frugal heuristics need an automatic frequency counter for ordering cues. In fact, only a few heuristics order cues, and these orderings can arise from evolutionary, social, or individual learning, none of which requires automatic frequency counting. The idea that cue validities cannot be computed because memory does not encode missing information is misinformed; it implies that measures of co-occurrence are incomputable and would invalidate most theories of cue learning. They also questioned the recognition heuristic's psychological plausibility on the basis of their belief that it has not been implemented in a memory model, although it actually has been implemented in ACT-R (L. J. Schooler & R. Hertwig, 2005). On the positive side, M. R. Dougherty et al. discovered a new mechanism for a less-is-more effect. The authors of the present article specify minimal criteria for psychological plausibility, describe some genuine challenges in the study of heuristics, and conclude that fast and frugal heuristics are psychologically plausible: They use limited search and are tractable and robust.

Keywords: adaptive toolbox, cue validity, heuristics, take the best, recognition heuristic

The program of fast and frugal heuristics centers on three questions. The first concerns the adaptive toolbox: What heuristics do organisms use? Answering this involves identifying heuristics, their building blocks, and the evolved capacities that these exploit. The second concerns ecological rationality: What are the environmental structures in which a given heuristic works well or poorly, and how do people adapt heuristics to these structures? The third question concerns applications: How can the study of ecological rationality inform the design of heuristics and environments to improve decision making? For example, ecologically rational design has been used to improve coronary care unit allocations (Green & Mehr, 1997), first-line antibiotic prescription in children (Fischer et al., 2002), and risk

Gerd Gigerenzer, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany; Ulrich Hoffrage, Faculty of Economics and Business Administration, University of Lausanne, Dorigny, Switzerland; Daniel G. Goldstein, Centre for Marketing, London Business School, London, England.

The order of second and third authors was determined by the closing value of the DAX on April 27, 2007. We thank Wolfgang Gaissmaier, Ralph Hertwig, Julian Marewski, Henrik Olsson, Magnus Persson, and Lael Schooler for helpful comments on this article, as well as Michael Dougherty and Rick Thomas for providing their simulation code.

Correspondence concerning this article should be addressed to Gerd Gigerenzer, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, Berlin, Germany 14195. E-mail: sekgigerenzer@mpib-berlin.mpg.de

communication in medicine and law (Gigerenzer, 2002; Gigerenzer & Edwards, 2003; Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000). The results have been summarized in four books (Gigerenzer, 2007; Gigerenzer & Engel, 2006; Gigerenzer & Selten, 2001; Gigerenzer, Todd, & the ABC Research Group, 1999).

Dougherty, Franco-Watkins, and Thomas (2008) commented on our initial work, the 1991 and 1996 *Psychological Review* articles (Gigerenzer & Goldstein, 1996; Gigerenzer, Hoffrage, & Kleinbölting, 1991), but apart from the 2002 *Psychological Review* article on the recognition heuristic (Goldstein & Gigerenzer, 2002), they did not deal with our subsequent research. Their critique concerns the search rule of Take The Best (TTB)—its stopping rule and decision rule are not discussed—and the recognition heuristic, but they inappropriately generalized to all fast and frugal heuristics. To the extent that other heuristics are mentioned at all, the authors passed over or dismissed them. Although Dougherty et al. acknowledged the concept of an adaptive toolbox, they also presented our program

¹ For instance, Dougherty et al. (2008) wrote that the Minimalist heuristic, which does not order cues by validity but searches in random order, is "particularly problematic, in the sense that randomness cannot be empirically validated" (p. xx). This is a surprising statement. One certainly can test randomness as well as models of cognition that make explicit assumptions about stochastic generation of cues or instances (e.g., Bergert & Nosofsky, 2007; Tversky's 1972 elimination by aspects [EBA]).

Table 1
Some Misconceptions, Insights, and Research Questions

Misconception	Clarification	
Cue validity is the same as ecological validity.	Ecological validities are properties of the environment, whereas cue validities are properties of a probabilistic mental model.	
Take The Best needs to compute precise ecological validities to order cues.	Take The Best only requires ordering cues, not computing quantitative validities. Depending on how cue orders are learned, there will be individual differences that depart from the ecological validity order. If the distribution of ecological cue validitie is flat, then different individual orders have little effect on accuracy; the more it is skewed, the larger the effect but the smaller the likelihood of large individual differences (see Martignon & Hoffrage, 2002).	
A memory representation cannot register the absence of information, that is, negative cue values, and therefore cue validity cannot be computed.	Absence of information is not the same as a negative cue value. Negative cue values indicate lower criterion values and do not necessarily correspond to the absence of a cue. Negative cue values, like positive ones, are either encoded or, if they have not been encoded, then they are inferred or not known (i.e., a missing value in a probabilistic mental model, not a negative value). Cue validity can be computed independently of whether negative cue values have been encoded.	

Insight

- 1. Forgetting enables the recognition heuristic and the fluency heuristic to make better inferences (Schooler & Hertwig, 2005).
- Less information, time, and computation can improve cognitive and motor performance in a number of situations (Gigerenzer, 2007; Hertwig & Todd, 2003).
- 3. Ordering cues in a simple, unconditional way can improve judgments as compared with rational conditional ordering or weighing of cues (Brighton, 2006; Czerlinski, Gigerenzer, & Goldstein, 1999). Thus, the cognitive inability to monitor dependencies between cues can actually enhance the accuracy of inference.

Research question

- 1. How do people select between heuristics?
- 2. How do evolved capacities, including cognitive limitations, support the efficiency of heuristics?
- 3. What are the relevant structures of physical and social environments that specific heuristics can exploit?

as if it were about one heuristic only (as it was in 1991). They argued that Bröder (2000) and Newell "have shown that it [TTB] is far from a universal heuristic" (Dougherty et al., 2008, p. 209). Bröder (2000) indeed initially invoked the question of whether TTB is a "universal" heuristic used by everyone in every situation. However, he soon began to ask in which situations people rely on which heuristic (Bröder, 2003; Bröder & Schiffer, 2003). In contrast, Dougherty et al. still promoted the concept of an all-purpose heuristic and did not investigate each heuristic's ecological rationality. We now remedy some misunderstandings (see Table 1).

Fast and Frugal Heuristics Do Not Need an Automatic Frequency Counter

Dougherty et al.'s (2008) first conjecture is that "PMM and the fast and frugal algorithms assume that cue validities are based on a frequency-counter process, such as that proposed by Hasher and Zacks (1979)" (p. 200). They then declared Hasher and Zacks's (1979) frequency counter—and, by association, all the fast and frugal heuristics—to be "implausible." This conclusion, in particular its generalization to all heuristics, is astonishing. Only one of the five models of heuristics in Gigerenzer and Goldstein's (1996) article orders cues by validities (TTB); the others rely on equal weights, recency, or random search. Moreover, heuristics, such as satisficing, tit for tat, and imitate the successful, do not rely on cue order but on principles such as aspiration levels, reciprocity, and behavior copying. What about TTB, however? Does it need an automatic fre-

quency counter to order cues? The answer is no. In Gigerenzer et al.'s (1991) article, we explicitly remained open as to how frequency encoding is carried out, which Dougherty et al. cited but nevertheless disregarded in asserting that an automatic frequency encoding "is necessary (though not sufficient) for the functioning of...the TTB algorithm" (p. 201). This view is incorrect. In Gigerenzer et al.'s (1991) article, we acknowledged that "memory is often (but not always) excellent in storing frequency information" (p. 510). No perfect automatic frequency encoding is assumed.² In Gigerenzer and Goldstein's (1996) article, we made it clear that the "Take The Best algo-

² The assumption of automatic frequency counting is also not part of the Brunswikian psychology on which probabilistic mental model theory is based. Following Helmholtz, Brunswik emphasized the role of frequency learning against the Gestalt theorists' emphasis on innate perceptual structures. However, for Brunswik, the question of to what extent frequency judgments are accurate or depend on other variables was an empirical one, not an a priori assumption. In his multidimensional psychophysics, Brunswik (1937) argued against the one-dimensional psychophysics from Fechner to Stevens, which treated perceived frequency, loudness, or area as independent from context variables. Brunswik was probably the first to study how estimates of the number of coins depend on their size and value and vice versa. In his words, judgments are perceptual compromises, not accurate and automatic counts of frequencies independent of context. Similarly, probabilistic mental model theory emphasizes that the accuracy of frequency judgments depends on the reference class activated (Gigerenzer et al., 1991) and on the size of the reference class (Hoffrage & Hertwig, 2006). Thus, there is no automatic frequency counter in the Brunswikian framework either.

rithm assumes a subjective rank order of cues" (p. 653), that is, an ordering by subjective, not ecological, validities (see Table 1). Furthermore, TTB only needs to order cues, not to compute quantitative values for cue validities (see Table 1). There are three ways cues can be ordered.

Evolutionary Learning

As exemplified by the classic work of Tinbergen (1958), animal biologists have studied rules of thumb (what psychologists call "heuristics"). Natural selection can produce good cue orders. For instance, female sage grouses first screen males on the basis of their songs and then visit only those who pass this test for a closer visual inspection of display rate (Gibson, 1996). Such a sequential use of cues seems to be extremely widespread in sexual selection as well as in animal navigation and food choice (for an overview, see Hutchinson & Gigerenzer, 2005). Similarly, evolutionary psychologists have argued that partner choice in humans involves gender-specific cues that evolved through sexual and natural selection, modified by social and individual learning. Without the benefit of evolutionary learning, the mind would be a "blank slate" (Pinker, 2002).

Social Learning

Cultural evolution is guided by three transmission factors: teaching, imitation, and language (Boyd & Richerson, 2001, 2005). Each of these factors also enables learning of cue orders. For instance, medical students and physicians are taught which diagnostic cues to check in which order, such as when a patient is suspected of acute ischemic heart disease. Learning cue orders by encoding co-occurrence frequencies can be a slow and dangerous process in medicine and beyond. In humans, social learning is probably the most widespread method of learning the identity and order of cues. Examples range from employers learning what to look for in potential employees to mountaineers learning how to predict avalanches (McCammon & Hägeli, 2007). Similarly, psychologists often use social learning when testing models of cognitive heuristics, for instance, when instructing participants about cue orders and other properties of the task (e.g., Rieskamp & Hoffrage, in press; Rieskamp & Otto, 2006).

Individual Learning

Learning by individual experience is a third way to order cues and is the only one Dougherty et al. (2008) considered. It is slow compared with social learning and evolutionary learning (from the perspective of the individual, not evolution). Apart from that, it can be too dangerous (think of learning by feedback about which mushrooms are poisonous) or practically impossible when the events are rare or feedback absent or unreliable. Experiments indicate that with a sufficient number of learning trials, participants can reliably learn to order cues by validity or success, without the aid of social learning (Garcia-Retamero, Hoffrage, & Dieckmann, 2007; Newell, Rakow, Weston, & Shanks, 2004; Rakow, Hinvest, Jackson, & Palmer, 2004; Rakow, Newell, Fayers, & Hersby, 2005).

To summarize, (a) most heuristics do not involve ordering cues by validity, (b) TTB can rely on three ways for ordering cues, and (c) even individual learning does not postulate perfection or independence from attention. No automatic frequency counter is necessary for establishing the cue order in a probabilistic mental model, TTB, and fast and frugal heuristics in general.

Minds Can Infer Negative Cue Values

Dougherty et al.'s (2008) second conjecture is that the definition of cue validities "requires complementary knowledge of events that are present and events that are absent from the environment" (p. 202), and they "see no way that a memory representation can register the absence of information" (p. 203). From this they concluded that cue validity cannot be computed and that its definition is "fundamentally flawed." We, in turn, conclude that their conclusion is fundamentally misinformed.

First, binary cue values (+ and -) do not necessarily stand for presence and absence of a cue. Being female, for instance, is not the absence of being male. In contrast, Dougherty et al. (2008) stated that "...cue values with minus signs correspond to the absence of the cue" (p. 202; also see their Figure 2). Yet one can easily register whether a person is male or female, or young or old. All the information necessary for computing validity is in memory. However, let us focus on the subset of cues that can be characterized by presence and absence.

One example they give is that while reading their article, a person will not encode that the words "xylophone" and "pepper" are not present in one of the sentences. However, Dougherty et al. (2008) confused what we encode with what we can know by inference, either implicitly or explicitly. Although memory might have encoded the fact that Dallas has an NFL football team, according to Dougherty et al., it is impossible to know that Honolulu does not. However, many people do. Doctors who test hypotheses about the condition of a patient may check for the presence of fever and, as a consequence, know if it is absent. There are even cases in which people pay specific attention to the absence of a cue, such as whether a friend avoids eye contact, a child forgets to say "thank you," and a colleague does not cite them. Cognition and memory involve more than encoding and retrieving facts; they also infer cue values and even update missing cue values (which is the essence of our hindsight bias model; Hoffrage, Hertwig, & Gigerenzer, 2000). There are further treatments of how people gain knowledge without having experienced it (e.g., Landauer & Dumais, 1997).

What, then, about the "xylophone" and "pepper" case? The point is that in the context of Dougherty et al.'s (2008) article, neither pepper nor xylophones play a role. If they had instead written a recipe for chili, one would have registered the absence of pepper, although still not that of xylophones. The same holds for cue-based inferences, where the absence of highly associated information would be registered. The absence of irrelevant information is not the same as a negative value of a relevant cue.

If Dougherty et al.'s (2008) conclusion were true, then it would imply that most cognitive theories based on validity or other measures of co-occurrence are psychologically implausible. To begin with, the definition of cue validity is identical, up to a linear transformation, to the Goodman–Kruskal rank correlation, which is close to Spearman's rho (Gigerenzer, 1981). Next, if the discrimination rates of binary cues are the same, then ordering cues by validity yields the same ordering as success (Martignon & Hoffrage, 1999). The ranking of cues by success is identical to that

by the Bayesian expected change in belief (Klayman & Ha, 1987), also known as the expected information gain (Oaksford & Chater, 1998), and to Shannon's information measure (see Rakow et al., 2004). In addition, the Pearson correlation between a cue and the criterion (Alternative A is greater vs. B is greater) can also be expressed as a function of validity (Rakow et al., 2005). This list can be continued. Thus, if cue validity could not be computed (because memory representations cannot register missing information), then no other theory that postulates ordering or weighing cues by success, rank correlations, Pearson correlations, Bayesian expected information gain, or similar weighting functions would be psychologically plausible. That would invalidate many cognitive theories and pose the question of how humans and other animals are able to learn correlations in the wild (say, between disease and symptom or cause and effect) where no-symptom and no-effect cases represent the absence of information.

To summarize, Dougherty et al. (2008) overlooked that binary cues do not generally imply absence or presence and that positive and negative cue values can be known by inference even without previous encoding. In contrast to their claim, cues can be ordered according to validity or success. The empirical evidence shows the same picture.

Inference From Memory

If TTB's memory requirements were implausible, as Dougherty et al. (2008) suggested, then there should be no evidence for this heuristic when people make inferences from memory, only when they make inferences from givens, that is, when cue values are displayed externally, as on a computer screen. Bröder and Schiffer (2003) tested and rejected this conjecture. When participants had to retrieve cue values from memory, a Bayesian model selection criterion classified 64% of 50 participants as using TTB, whereas two competing linear models accounted for only 12% of the participants each. Similarly, when Bröder and Schiffer (2003) directly compared inferences from memory with inferences from givens, the proportion of participants classified as using TTB was twice as large for inferences from memory. Hoffrage, Hertwig, and Gigerenzer (2000) implemented TTB in a model of hindsight bias and successfully tested its predictions in a task in which participants had acquired cue values in several preceding learning trials. Furthermore, Bröder and Gaissmaier (in press) analyzed response times from five experiments and reported that for participants classified as using TTB, response times increased monotonically with the number of cues that had to be retrieved from memory, as this heuristic would predict. With inferences from givens, Bergert and Nosofsky (2007) tested the reaction time predictions of TTB against those of a weighted additive model and concluded that the vast majority of participants were consistent with the heuristic's process predictions. In sum, when people had previously encoded positive and negative cue values in memory, their inferences were more often consistent with TTB as compared with external displays of cue values.

Some Genuine Challenges in the Investigation of Cue Orderings

Learning cue orders appears irrelevant if one assumes that all cues or features contribute equally to a decision, as in Dougherty, Gettys, and Ogden's (1999) Minerva-DM model. However, the assumption that cues are always equally weighted is unrealistic. In contrast, we assume that the adaptive toolbox contains both heuristics that order cues and heuristics that treat them equally, allowing organisms to select heuristics according to the problem at hand. For instance, Martignon and Hoffrage (1999) and Hogarth and Karelaia (2007) specified conditions under which TTB with its cue ordering is superior to weighing cues equally, such as high dispersion of cue validities, whereas Rieskamp and Otto (2006) showed that people intuitively adapt their heuristics to environmental conditions.

Having clarified the misunderstandings in Dougherty et al.'s (2008) conjectures, we now describe some of what we see as the real issues relating to cue-ordering heuristics, such as TTB, elimination by aspects, and fast and frugal trees (Gigerenzer, 2004).

How Do Social and Individual Learning Combine When Ordering Cues?

Gigerenzer and Goldstein (1996) did not deal with the question of how cue orders are learned. If one had asked us at that point in time, we would have conjectured that it occurs through feedback learning. When Todd and Dieckmann (2005) put this idea to test, it turned out that simulated agents can learn cue orders that way but that convergence toward the ecological cue order is slow if learning occurs while inferences are being made with TTB. To avoid slow convergence, a person who has little knowledge of cues can start with a trial-and-error phase for exploring cues and only later switch to an application phase in which the heuristic is used. However, there is a more interesting social solution to the problem of slow individual learning. Individual learners might exchange the information they have after a number of individual trials, continue with individual learning, converse again, and so on. We tested several social exchange rules: the average rule, the Borda rule, the Condorcet rule, and imitate the successful. All social rules, except the Borda rule, sped up individual learning. The simplest rule—to imitate the cue order of the most successful member-led to the best results (Garcia-Retamero, Takezawa, & Gigerenzer, 2006). Psychologists still know little about how social learning supports individual learning in judgment, classification, and decision making, which mirrors the unfortunate truth that many cognitive theories ignore the crucial role of social input.

How Do Minds Create Robust Rather Than Optimal Cue Orders?

Does it pay to compute the optimal cue order? The problem of finding the optimal (best) cue order turns out to be NP-hard (Schmitt & Martignon, 2006), that is, when the number of cues vastly increases, determining the optimal order quickly becomes too time-consuming for minds and computers. Tractability, in our view, is an important and necessary condition for what Dougherty et al. (2008) called "psychological plausibility."

Now consider situations with a small number of cues in which the optimal order can be determined. Should a rational mind try to compute the best cue order from the existing data, or could a satisfactory order be preferable to the optimal one? Using crossvalidation, Martignon and Hoffrage (2002) showed that the optimal cue order determined from a learning sample drawn from a

population was, in fact, not necessarily the optimal one in new samples from the same population. Surprisingly, orderings based on validity produced, on average, better performance in the new samples as compared with those that were optimal in the learning sample. The optimal order was less robust than the validity order. This leads to what we see as a second necessary condition for psychological plausibility: robustness. A cue order should be robust, not optimal (relative to the learning sample).³

Similarly, Brighton (2006) showed in dozens of real-world problems that TTB often outperformed neural networks, exemplar models, and other complex procedures that compute conditional weights. TTB's surprising predictive accuracy is partly due to its ignoring dependencies between cues (just like naive Bayes) rather than its relying on the apparently rational calculation of covariation matrices. Armelius and Armelius (1974) documented that people are at a loss when they are confronted with the task of estimating partial correlations between cues. This cognitive limitation, however, can now be put into a different functional perspective. In a range of situations, our minds can make more accurate inferences when they ignore the dependencies between cues; less can be more. This finding aligns with observations that cognitive limitations can actually promote better cue orders and better inferences (see also Hertwig & Todd, 2003).

Given the predictive power of simple, unconditional weights, new research questions need to be posed: How can we characterize situations in which unconditional cue orderings will lead to more accurate inferences? When do conditional weights, as used by all rational theories, begin to pay? What role do cognitive limitations play in enabling simple and robust inference processes?

How Do Minds Adapt Cues to the Task Environment?

In Gigerenzer and Goldstein (1996), we distinguished three kinds of search: search by validity, search by recency, and random search. In a stable environment with sufficient and reliable feedback, search by validity is most accurate; in the absence of these features, recency can be better because it is faster and more frugal. If one has to learn from scratch, then random search seems to be the only option. Another environmental feature is the discrimination rate of the cues. When these rates vary substantially, it pays to base search on a combination of validity and discrimination rate, such as success, as proposed by Martignon and Hoffrage (2002) and studied by Newell et al. (2004). Moreover, sensory and environmental constraints often impose order. Consider two competing male deer during rutting season. If the harem holder roars more impressively, then the challenger may give up immediately and walk away; otherwise, parallel walking is initiated, which provides cues to assess each other's physical fitness at closer distance. If these cues do not stop the contest, then they begin headbutting, their riskiest activity, which can lead to serious injuries (Enquist & Leimar, 1990). Here, the cue order is dictated by the access of the senses to the environment: Roaring can be heard first, visual cues need closer distance, and tactile cues require physical contact. Finally, in gambling environments, the priority heuristic sequentially considers different attributes of the lotteries, which cannot be expressed in terms of cue validities (Brandstätter, Gigerenzer, & Hertwig, 2006). A few pioneer experiments have investigated how people adapt search rules to environmental conditions (e.g., Rakow et al., 2004, 2005). How evolutionary, social, and individual learning interact to enable (or prevent) the adaptation of cue orders, or, more generally, heuristics, to the task environment is a largely open question.

The Recognition Heuristic Is Consistent With Memory Models

Is the recognition heuristic consistent with existent memory models? Dougherty et al. (2008) questioned its psychological plausibility: "To our knowledge, the only model within the fast and frugal tool kit to specify the underlying memory processes and primitives is the fluency heuristic, which is specified within the ACT-R framework (Schooler & Hertwig, 2005)" (p. 210, Footnote 12). However, the recognition heuristic was also implemented in the ACT-R framework, as reported in the same article. If implementation in an empirically validated memory model is the definition of plausible and if ACT-R is considered as such a model, then the recognition heuristic is psychologically plausible. Similarly, but more recently, the recognition heuristic was implemented in a signal-detection model (Pleskac, 2007).

The recognition heuristic is not a model of memory processes, rather, it models how inferences are made on the basis of the output of memory processes. It draws on an all-or-none recognition judgment, which Dougherty et al. (2008) claimed "is counter to the literature on recognition memory" (p. 204). However, one needs to distinguish between continuous trace activation and a binary recognition judgment (e.g., Malmberg, 2002; Pleskac, 2007). For instance, Schooler and Hertwig (2005) assumed that activation of a memory record (the continuous underlying value) cannot be accessed directly. However, its activation does govern sensations of which people can be aware, namely yes-no recognition judgments (retrieval of a memory record) and experienced recognition time (recognition latency). The assumption of being able to arrive at yes-no recognition judgments is intuitively plausible; we classify people as recognized or unrecognized with ease and hesitate only rarely. In many experiments, however, the task is set up to make discrimination intentionally difficult. Furthermore, yes-no recognition is consistent with how recognition is modeled in ACT-R and other theories (see Schooler & Hertwig, 2005, pp. 617-618, 625).

Schooler and Hertwig (2005) showed that an intermediate degree of forgetting enables both the recognition heuristic and the fluency heuristic to make more accurate inferences. This result is

³ The same robustness argument applies to orders established by conditional cue validities (that are sequentially computed on the subset of pairs for which higher ranked cues did not discriminate). In the learning sample, ordering cues by conditional validity was more accurate than by (unconditional) validity and almost as good as the optimal order. However, when tested in new samples, ordering by conditional validity led to worse inferences than when simply ordering by validity (Martignon & Hoffrage, 2002).

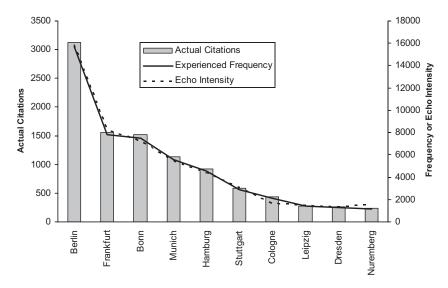


Figure 1. Familiarity, defined by Minerva 2's echo intensity, mirrors actual frequency (citations) and experienced frequency. The bars show the actual number of citations of the 10 most frequently cited German cities in the *Chicago Tribune* (left-hand axis). The two lines show how often a simulated reader encountered a particular city in the newspaper (experienced frequency) and this city's echo intensity, as modeled by Minerva 2. The simulation was performed using Dougherty, Franco-Watkins, and Thomas's (2008) code, which they were kind enough to make available.

an extension of the less-is-more effect that can arise from the recognition heuristic. Dougherty et al. (2008) could have discussed how their familiarity model differs from the fluency heuristic but instead presented their model as if it were an alternative to the toolbox approach.

What we do believe to be an interesting contribution is Dougherty et al.'s (2008) new model for a less-is-more effect. Such a between-domain effect is illustrated by our finding that American students were slightly more accurate in judgments about the largest German cities as opposed to the largest American cities (the domains being the two countries), despite knowing less about German cities. In contrast, a within-domain less-is-more effect is illustrated by Figure 2 in Goldstein and Gigerenzer's (2002) article, where three sisters differing in recognition knowledge make judgments about the same domain of objects. We have distinguished between three kinds of lessis-more effects: within domain, between domain, and during knowledge acquisition (Goldstein & Gigerenzer, 2002, p. 83).⁴ However, Dougherty et al. did not mention this, instead suggesting incorrectly that we "treated participants from different ecologies [i.e., domains] as points on the same curve" (p. 205). This is an oversight that could have been easily avoided. We might add that Dougherty et al.'s new model is not the first alternative explanation. In Gigerenzer and Goldstein (1996, Figure 6), we showed that three other models that do not involve the recognition heuristic but instead treat recognition as a cue can produce a nonmonotonic pattern of accuracy. Note that the recognition heuristic does not assume that too much knowledge leads to poor inference; it predicts when it will and will not happen. The task for the alternative models is then to specify the conditions under which the effect occurs.

Dougherty et al.'s (2008) simulations involve two models that explicitly (frequency model) or implicitly (familiarity model) en-

code frequency information. To make this clear, consider Figure 1. We had a simulated participant sample 50,000 times from the 10 most cited German cities in Dougherty et al.'s *Chicago Tribune* data. The simulation's memory processes are those of the familiarity model (Minerva 2). It becomes evident that the echo intensity of the familiarity model mirrors with impressive accuracy the frequency information of the automatic frequency counter. This reflects the fact that Minerva 2, though not explicitly a frequency counter, stores and processes every instance and creates an "obligatory, analog representation of frequency" (Hintzman, 1988, p. 547). Oddly enough, frequency counters are what our critics dismissed as psychologically implausible.

The familiarity model Dougherty et al. (2008) proposed has its merits and limitations. On the one hand, it treats paired comparisons with one simple decision rule (choose the more familiar). On the other hand, it cannot account for the evidence that people make inferences based on cues other than familiarity—although this problem would be resolved if Dougherty et al. were willing to consider the familiarity model as one of the tools in the adaptive toolbox rather than as a universal strategy. We would be curious to hear their thoughts on the connection between the adaptive toolbox and their Minerva-DM approach. The adaptive toolbox approach is taken by Schooler and

⁴ Dougherty et al. (2008) presented their simulation finding that the recognition heuristic, combined with TTB, can predict both the presence and the absence of a less-is-more effect as something new and as suggestive of excessive model flexibility. However, they neglected to mention that we have specified the exact conditions under which the effect does and does not occur—such as when the recognition validity exceeds the knowledge validity and other assumptions hold (Goldstein & Gigerenzer, 2002, pp. 79–80).

Hertwig's (2005) ACT-R implementation of the fluency heuristic (which corresponds to the familiarity model) and the recognition heuristic. If recognition does not discriminate, then an inference could be made by the fluency heuristic, TTB, or another heuristic, as the constraints of the environment dictate.

Criteria for Psychological Plausibility

We strongly agree with Dougherty et al. (2008) that psychology needs models rather than labels for cognitive processes. We also agree that these models should be psychologically plausible. However, what does this term mean? We conclude with what we believe are important criteria for assessing the psychological plausibility of a model.

Tractability

The computations postulated by a model of cognition need to be tractable in the real world in which people live, not only in the small world of an experiment with only a few cues. This eliminates NP-hard models that lead to computational explosion, such as probabilistic inference using Bayesian belief networks (Cooper, 1990), including its approximations (Dagum & Luby, 1993). Tractability is one condition for psychological plausibility that is easily met by simple heuristics.

Robustness

The cognitive inferences and estimations (e.g., of cue orders) involved should be robust rather than optimal (on the learning sample; see Roberts & Pashler, 2000). In other words, cognition should be successful in foresight rather than in hindsight and not waste effort on computations and estimations that deteriorate performance. Robustness increases with the simplicity of a model and decreases with the number of free parameters.

Frugality

Plausible models of cognitive processes need to specify not only how information is integrated but also when information search is stopped in the first place. Stopping rules, in turn, motivate search rules, which can order cues such that good cues are likely to be encountered first. In an age of overwhelming information, it is clear that a crucial issue is knowing when to ignore further information.

Speed

The cognitive processes assumed by a model should be able to be executed quickly. Speed is often, albeit not always, an important feature of everyday cognition and action.

Evidence

Models that satisfy these four requirements should be consistent with what we know about cognition. For Schooler and Hertwig (2005), implementing heuristics in cognitive architectures, such as ACT-R, fulfills this requirement. We believe that fast and frugal heuristics are among those cognitive models that satisfy these criteria well.

References

- Armelius, B., & Armelius, K. (1974). The use of redundancy in multiplecue judgments: Data from a suppressor-variable task. *American Journal* of Psychology, 87, 385–392.
- Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 107–129.
- Boyd, R., & Richerson, P. J. (2001). Norms and bounded rationality. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive tool-box* (pp. 281–296). Cambridge, MA: MIT Press.
- Boyd, R., & Richerson, P. J. (2005). The origin and evolution of cultures. New York: Oxford University Press.
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: A process model of risky choice. *Psychological Review*, 113, 409-432.
- Brighton, H. J. (2006). Robust inference with simple cognitive models. In C. Lebiere & R. Wray (Eds.), Between a rock and a hard place: Cognitive science principles meet AI-hard problems. Papers from the AAAI Spring Symposium (pp. 17–22). Menlo Park, CA: American Association for Artificial Intelligence Press.
- Bröder, A. (2000). Assessing the empirical validity of the "take-the-best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1332–1346.
- Bröder, A. (2003). Decision making with the "adaptive toolbox:" Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cogni*tion, 29, 611–625.
- Bröder, A., & Gaissmaier, W. (in press). Sequential processing of cues in memory-based multi-attribute decisions. *Psychonomic Bulletin & Review*.
- Bröder, A., & Schiffer, S. (2003). Take The Best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, 132, 277–293.
- Brunswik, E. (1937). Psychology as a science of objective relations. *Philosophy of Science*, 4, 227–260.
- Cooper, G. F. (1990). Computational complexity of probabilistic inference using Bayesian belief networks. Artificial Intelligence, 42, 393–405.
- Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. M. Todd, & the ABC Research Group, Simple heuristics that make us smart (pp. 97–118). New York: Oxford University Press.
- Dagum, P., & Luby, M. (1993). Approximating probabilistic inference in Bayesian belief networks is NP-hard. Artificial Intelligence, 60, 141–153.
- 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. P., Gettys, C. F., & Ogden, E. (1999). Minerva-DM: A memory process model for judgments of likelihood. *Psychological Review*, 106, 180–209.
- Enquist, M., & Leimar, O. (1990). The evolution of fatal fighting. *Animal Behaviour 39*, 1–9.
- Fischer, J. E., Steiner, F., Zucol, F., Berger, C., Martignon, L., Bossart, W., et al. (2002). Using simple heuristics to target macrolide prescription in children with community-acquired pneumonia. Archives of Pediatrics and Adolescent Medicine, 156, 1005–1008.
- Garcia-Retamero, R., Hoffrage, U., & Dieckmann, A. (2007). When one cue is not enough: Combining fast and frugal heuristics with compound cue processing. *Quarterly Journal of Experimental Psychology*, 60, 1197–1215.

- Garcia-Retamero, R., Takezawa, M., & Gigerenzer, G. (2006). How to learn good cue orders: When social learning benefits simple heuristics. In R. Sun & N. Miyake (Eds.), Proceedings of the 28th annual conference of the Cognitive Science Society (pp. 1352–1358). Mahwah, NJ: Erlbaum
- Gibson, R. M. (1996). Female choice in sage grouse: The roles of attraction and active comparison. *Behavioral Ecology and Sociobiology*, 39, 55–59.
- Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie* [Measurement and modeling in psychology] (Uni-Taschenbücher No. 1047). Munich, Germany: Reinhardt.
- Gigerenzer, G. (2002). Calculated risks: How to know when numbers deceive you. New York: Simon & Schuster.
- Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. In D. Koehler & N. Harvey (Eds.), Blackwell handbook of judgment and decision making (pp. 62–88). Oxford, England: Blackwell.
- Gigerenzer, G. (2007). Gut feelings: The intelligence of the unconscious. New York: Viking.
- Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: From innumeracy to insight. *British Medical Journal*, 327, 741–744
- Gigerenzer, G., & Engel, E. (Eds.). (2006). *Heuristics and the law*. Cambridge, MA: MIT Press.
- 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., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506–528.
- Gigerenzer, G., & Selten, R. (Eds.). (2001). Bounded rationality: The adaptive toolbox. Cambridge, MA: MIT Press.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999).
 Simple heuristics that make us smart. New York: Oxford University Press.
- Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109, 75–90.
- Green, L., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *The Journal of Family Practice*, 45, 219–226
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356–388.
- Hertwig, R., & Todd, P. M. (2003). More is not always better: The benefits of cognitive limits. In D. Hardman & L. Macchi (Eds.), *Psychological* perspectives on reasoning, judgment, and decision making (pp. 213– 232). Chichester, England: Wiley.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95, 528–551.
- Hoffrage, U., & Hertwig, R. (2006). Which world should be represented in representative design? In K. Fiedler & P. Juslin (Eds.), *Information* sampling and adaptive cognition (pp. 381–408). New York: Cambridge University Press.
- Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 566–581.
- Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000, December 15). Communicating statistical information. Science, 290, 2261–2262.
- Hogarth, R. M., & Karelaia, N. (2007). Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review*, 114, 733–758.

Hutchinson, J. M. C., & Gigerenzer, G. (2005). Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes*, 69, 97–124.

- Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94, 211–228.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211–240.
- Malmberg, K. J. (2002). On the form of ROCs constructed from confidence ratings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 380–387.
- Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision making work? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, Simple heuristics that make us smart (pp. 119–140). New York: Oxford University Press.
- Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, 52, 29–71.
- McCammon, I., & Hägeli, P. (2007). An evaluation of rule-based decision tools for travel in avalanche terrain. *Cold Regions Science and Technol*ogy, 47, 193–206.
- 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.
- Oaksford, M., & Chater, N. (Eds.). (1998). Rational models of cognition. New York: Oxford University Press.
- Pinker, S. (2002). The blank slate: The modern denial of human nature. New York: Viking.
- Pleskac, T. J. (2007). A signal detection analysis of the recognition heuristic. Psychonomic Bulletin & Review, 14, 379–391.
- Rakow, T., Hinvest, N., Jackson, E., & Palmer, M. (2004). Simple heuristics from the adaptive toolbox: Can we perform the requisite learning? Thinking and Reasoning, 10, 1–29.
- Rakow, T., Newell, B. R., Fayers, K., & Hersby, M. (2005). Evaluating three criteria for establishing cue-search hierarchies in inferential judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1088–1104.
- Rieskamp, J., & Hoffrage, U. (in press). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*.
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135, 207–236.
- Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, 107, 358–367.
- Schmitt, M., & Martignon, L. (2006). On the complexity of learning lexicographic strategies. *Journal of Machine Learning Research*, 7, 55–83.
- Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, 112, 610–628.
- Tinbergen, N. (1958). Curious naturalists. New York: Basic Books.
- Todd, P. M., & Dieckmann, A. (2005). Heuristics for ordering cue search in decision making. In L. K. Saul, Y. Weiss, & L. Bottou (Eds.), Advances in neural information processing systems (Vol. 17, pp. 1393– 1400). Cambridge, MA: MIT Press.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. Psychological Review, 79, 281–299.

Received May 10, 2007
Revision received August 8, 2007
Accepted August 9, 2007

Postscript: Fast and Frugal Heuristics

Gerd Gigerenzer

Max Planck Institute for Human Development

Ulrich Hoffrage University of Lausanne

Daniel G. Goldstein London Business School

In their postscript, Dougherty, Franco-Watkins, and Thomas (2008) asserted that models of fast and frugal heuristics have been vaguely specified. We strongly reject this claim. The computational models of search, stopping, and decision rules allow for precise predictions. In fact, many researchers have tested under which conditions people follow which heuristic (e.g., Bröder & Schiffer, 2003), compared the predictions of heuristics to those of rational models (e.g., Bergert & Nosofsky, 2007), and studied the ecological rationality of different heuristics through statistical analysis and computer simulation (e.g., Hogarth & Karelaia, 2007).

Our point that ecological validity (the relationship between cue and criterion in the environment) is not the same as cue validity (the perceived relationship between cue and criterion in a person's mind) is not a "new found clarity;" it was made in Gigerenzer, Hoffrage, and Kleinbölting's (1991) article. It has nothing to do with the accuracy of predictions. As in Brunswik's (1955) lens model, ecological validity refers to what is in the environment, and cue validity (Brunswik's cue utilization) refers to what is in the mind. Because people typically have imperfect knowledge of environmental structures, they accordingly rely on samplebased estimates. Dougherty et al. (2008, p. 212) took this to imply that Take The Best "operates on ANY subjective cue order, even if it were completely idiosyncratic." This claim is incorrect. A proper test of search rules used by people is fairly straightforward. First, the predictions of several models (not just one) for ordering cues—say, validity, success (Bayesian expected information gain), and beta weights (Brunswik's intuitive statistician)—are derived for the experimental task. Each prediction is based on the ecological measures of validity, success, and beta weights, or, if learning samples are small, on the sample-based measures. Next, each individual pattern of judgment (rather than the aggregate) is tested against the predictions of each model. If an individual's cue order is, for instance, closer to the ecological measures of success than to those of validity and beta weights, then he or she would be classified as relying on success (e.g., Rakow, Newell, Fayers, & Hersby, 2005). In the same way, competing models of stopping rules and decision rules can be tested. We are interested in knowing which cue orders are elicited by which situation and not in proving that everyone always orders cues by validity.

Now consider simulations. Here, the situation is different because, unlike the participants in an experiment, the researcher knows the exact ecological validities (or other measures of correlation). Many simulations tested Take The Best with the ecological validities and then compared it with other models, such as multiple regression, with the ecological beta weights. For instance, Garcia-

Retamero, Takezawa, and Gigerenzer (2006) used Take The Best with ecological validities as a benchmark and showed by means of simulation that social learning can boost accuracy beyond that reached with ecological validities alone. Dougherty et al. (2008) quoted from this article (and two other articles using simulations) and incorrectly claimed that we also wrongly equate ecological validities with cue validities. This is not the case. In simulations, one can test every model (not just Take The Best) with the ecological weights, but this does not imply that a real person would have exact knowledge of these weights.

Dougherty et al. (2008) also argued that the recognition heuristic might be vague because "to derive predictions based on the recognition heuristic, one needs to instantiate it at the level of a recognition memory model, as has been done by Pleskac (2007) and Schooler and Hertwig (2005)" (p. 213). We would like to mention that this work is in fact from our research group: Schooler and Hertwig's (2005) article is from our lab at the Max Planck Institute, and Pleskac worked in Hertwig's lab while he wrote his recognition article. We wish that Dougherty et al. (2008) had instead dealt at greater length with the fundamental questions that arise in their postscript.

General Purpose or Domain Specific?

Leibniz (1677/1951) hoped to reduce rational thinking to a single, universal calculus. Although he failed to realize it, his beautiful dream persists in many forms in current cognitive psychology, including formal logic, expected utility theory, and Bayesian inference. By definition, a single calculus is general purpose, so theories of cognition based on Leibniz's ideal do not have to address the question of ecological rationality (i.e., the question of which cognitive strategies match which environmental structures). However, if—like Dougherty et al. (2008)—one assumes a small number of general-purpose strategies rather than Leibniz's one, then this question must be addressed. Because their proposed "general-purpose" strategy of choosing the most familiar object is not viable in all situations, the question is as follows: How do minds decide when to make a judgment by familiarity and when to switch to another "general-purpose" strategy? That requires research on the ecological rationality of the familiarity heuristic and, more generally, on how people select between several general-purpose strategies.

How Do People Select Between Heuristics?

We listed this important question as a topic of future research in Table 1 of our reply, and it is also essential for understanding how a mind would operate with several general-purpose heuristics. Had Dougherty et al. (2008) argued that the present knowledge of heuristic selection, as opposed to models of heuristics, is rather vague, then they would have made a fair point. However, there is progress on the issue of selection as well. The study of ecological rationality has identified environmental structures in which, for instance, tallying is more accurate than Take The Best, and this provides testable conditions for when people switch between these heuristics. Moreover, in the case of individual learning by feedback, members of our

group have developed a formal model of strategy selection (Rieskamp & Otto, 2006).

Criticism can be fruitful, even more so if several theories were to be evaluated and compared using the same criteria. Moreover, understanding the respective advantages and blind spots should serve not only criticism of other theories but also theory integration. What psychology lacks in comparison with economics or physics is an integrated system of theories. Now is the time to ask what we can learn from other points of view and how we can integrate disparate theories to secure the future of psychology. Cumulative progress can hardly be achieved otherwise.

References

- Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 107–129.
- Bröder, A., & Schiffer, S. (2003). Take the best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, 132, 277–293.
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193–217.
- Dougherty, M. R., Franco-Watkins, A. M., & Thomas, R. (2008). Psycho-

- logical plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review*, 115, 199–213.
- Garcia-Retamero, R., Takezawa, M., & Gigerenzer, G. (2006). How to learn good cue orders: When social learning benefits simple heuristics. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 1352–1358). Mahwah, NJ: Erlbaum.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506–528.
- Hogarth, R. M., & Karelaia, N. (2007). Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review*, 114, 733–758.
- Leibniz, G. W. (1951). Toward a universal characteristic. In P. P. Wiener (Ed.), *Selections* (pp. 17–25). New York: Scribner's Sons. (Original work published 1677)
- Rakow, T., Newell, B. R., Fayers, K., & Hersby, M. (2005). Evaluating three criteria for establishing cue-search hierarchies in inferential judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1088–1104.
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135, 207–236.
- Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, 112, 610–628.

ORDER FORM Start my 2008 subscription to <i>Psychological Review</i> ! ISSN: 0033-295X		Check enclosed (make payable to APA)	
		Charge my: VISA MasterCard American Express	
\$7	1.00, APA Member/Affiliate	Cardholder Name	
	56.00, Individual Nonmember	Card No Exp. Date	
In DC add 5.75% / In MD add 5% sales tax TOTAL AMOUNT ENCLOSED \$		Signature (Required for Charge) BILLING ADDRESS:	
Subscription orders must be prepaid. (Subscriptions are on a calendar year basis only.) Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.		Street	
		City State Zip	
SEND THIS ORDER FORM TO: American Psychological Association		Daytime PhoneE-mail	
Subscriptions 750 First Street, NE	1	MAIL TO:	
	750 First Street, NE	Name	
AMERICAN	Washington, DC 20002-4242	Address	
PSYCHOLOGICAL Or call 800-374-2721 , fax 202-336-5568 .			
Association 1	TDD/TTY 202-336-6123.	City State Zip	
	For subscription information, e-mail: subscriptions@apa.org	APA Member # REVAIT	