1

Re-visions of rationality?

Ben R. Newell

University of New South Wales, Sydney, Australia.

&

Centre for Economic Learning and Social Evolution.

Key words: rationality, process models, judgment heuristics.

Teaser sentence: Empirical evidence suggests proponents of the 'adaptive toolbox' framework of human judgment need to rethink their vision of rationality.

Address correspondence to:

Ben R Newell

School of Psychology

The University of New South Wales

Sydney 2052

Australia

Email: ben.newell@unsw.edu.au

Tel: +61 (02) 9385 1606

Fax: +44 (02) 9385 361 4

The appeal of simple algorithms that take account of both the constraints of human cognitive capacity and the structure of environments has been an enduring theme in cognitive science. A novel version of such a boundedly rational perspective views the mind as containing an 'adaptive toolbox' of specialized cognitive heuristics suited to different problems. Although intuitively appealing, when this version was proposed empirical evidence for the use of such heuristics was scant. I argue that in the light of empirical studies carried out since then, it is time this 'vision of rationality' was revised. An alternative view based on integrative models rather than collections of heuristics is proposed.

Human judgment can be evaluated by the degree to which it 'coheres' with a formal model such as Bayes' theorem, and by the degree to which it 'corresponds' with the properties of the environment [1]. Some researchers have argued that demonstrating departures from coherence provides windows on the underlying heuristic processes commonly used in judgment, in much the same way as the study of visual illusions elucidate the workings of our perceptual machinery [2,3]. Others contend that focusing on such biases paints an unduly pessimistic picture of human judgment [4], and some go further by suggesting that coherence criteria are simply the wrong standards for evaluating the quality of human judgment [5,6].

This latter view, perhaps most prominently proposed by Gigerenzer and colleagues, argues for an ecological standard of rationality, that is judgment algorithms that perform well in the real world, regardless of their adherence to formal inference methods (e.g., multiple regression, Bayes' theorem). Ecological rationality is thus concerned solely with the correspondence criterion of rationality. The approach is an extension of the ideas of Brunswik [7] and in particular Simon [8] who argued that when defining optimal behaviour it is imperative to consider both the cognitive limitations of an organism and the role played by the environment in which the organism finds itself.

The centerpiece of the ecological approach is a new breed of 'fast-and-frugal' heuristics, which emphasize detailed process models, make clear predictions and are readily testable [5,6]. The heuristics acknowledge the existence of cognitive limitations by utilizing the least necessary amount of information (frugal) in the shortest time (fast). They also save us from the ignominy of poor and biased judgments by capitalizing on the fit between a heuristic and an environment – i.e., they are 'ecologically rational' [9].

The appeal of this framework to cognitive scientists is the promise of simple, psychologically plausible algorithms that, counter-intuitively, perform as well – or sometimes even better than more complex ones (See Box 1 for discussions of the psychological plausibility of the heuristics). For the broader community the temptation of easy shortcuts "that make us smart" has proved too much to resist.

Consequently, the fast and frugal framework has generated extensive debate in the literature [10-30] as well as being examined in a number of applied contexts – evidence-based medicine, [31] and magistrates' decisions [32], to take just two examples.

Despite this impact and the appeal of a set of readily testable simple models, when proposed, this vision of rationality was supported by very little empirical evidence demonstrating the use of the heuristics in the environments in which they are claimed to operate – a shortcoming that was duly noted by a number of commentators (e.g., 27-30).

Since the original framework was introduced the gulf between the bold vision and the empirical reality has begun to be filled. Researchers have started to question whether there is any evidence for a set of fast and frugal heuristics contained within an 'adaptive toolbox'; whether we can determine how one heuristic is selected or 'triggered' over another in particular situations; and whether it is possible to distinguish between a toolbox of strategies and a single evidence-accrual strategy.

To address these questions, I examine some of the empirical evidence for the heuristic that has received the most attention in the literature – the "Take-the-Best" heuristic. Whilst acknowledging that focusing on a single heuristic limits the degree to which the whole 'vision' can be evaluated, I argue that the questions raised about the empirical validity of Take-the-Best point to some fundamental problems for the

thesis underlying the fast and frugal framework, and furthermore that existing evidence is at least as consistent with a different vision of rationality.

Do people "Take-the-Best"?

Imagine you are facing a choice between two alternatives – such as two companies to invest in – and your task is to pick the one that is 'better' with regard to some criterion (e.g., future returns on investments). "Take-the-Best" (TTB) is designed for such a choice and exemplifies fast and frugal judgment by simply using the 'best' piece of information applicable in a given situation [10].

TTB operates according to two principles. The first – the recognition principle – states that in any given decision made under uncertainty, if only one amongst a range of alternatives is recognized, then the recognized alternative will be chosen [11] (see 12, 19, 21-23, 33-34 for conflicting opinions about the use of a "recognition heuristic" in both laboratory and field settings). The second principle is invoked when more than one of the alternatives are recognized and the recognition principle cannot provide discriminatory information. In such cases, people are assumed to have access to a reference class of cues or features, which are searched in descending order of feature validity until a feature that discriminates one alternative from the other is discovered. Search then stops and this single 'best' discriminating feature is used to make the choice. The algorithm is thus not rational in a formal sense because rather than using all pieces of information (as for example linear regression would) it bases its choice on a single piece of information [10].

Despite the impressive success of this simple strategy in computer simulations [10], finding evidence in the laboratory for the use of the deterministic search, stopping and decision rules has proved elusive. In a series of studies, Newell and

colleagues [12-15] used a simple share prediction task to examine peoples' adherence to the search, stopping and decision rules of TTB. In the experiments a number of different factors were varied such as the cost of information, the number of cues in the environment, the underlying structure of the task (deterministic or probabilistic), the informativeness of cues, and the provision of hints concerning the validity ordering of the cues.

Despite designing environments strongly constrained to promote the use of TTB, in all the experiments the over-all pattern of results was similar: simply stated – some of the people made choices consistent with TTB some of the time. In all experiments, a significant proportion of participants adopted strategies that violated all or some of TTB's rules. Consistent with these results, a number of other researchers have also found evidence for similar departures from TTB's search, stopping and decision rules [16-18, 20, 23-25].

Some might argue that these types of experiments are simply looking in the wrong place for evidence – the tasks are just not suited for promoting TTB (see Box 2). It could also be argued that evidence of *some* TTB-consistent behaviour provides partial validation for the psychological reality of the model and that it is unrealistic to expect *every* participant to behave according to one model [17]. These claims may well have some substance, but perhaps, rather than dismissing the evidence or somehow shaping it to fit a more moderate version of the framework it could be used to support an alternative model.

How many tools in the adaptive toolbox?

The wide individual differences in behavior that show systematic deviations from the deterministic rules of heuristics present a fundamental problem for the fast-

and-frugal framework. Although the framework allows for people to have access to numerous strategies or heuristics in their 'adaptive toolbox', it assumes that it is the environment, which to a large extent determines strategy selection and not the individual. Without such an assumption the framework necessarily falls foul of the homunculus problem of needing a meta-heuristic to select the appropriate 'tool for the job' [35]. The individual differences in strategy use are not easily reconciled with the proposed role for the environment in triggering particular strategies. Why do people with the same cognitive apparatus, operating in the same environment – (for which there is often a single 'ecologically rational' strategy) - use widely different strategies? [Cf. 25]

The reason this question poses a problem for the fast-and-frugal framework is that although the models are clearly specified, there is no indication of the degree of empirical deviation permissible from the deterministic search, stopping and decision rules. Any deviations are merely ignored as noise, allowing proponents to claim 64% of choices consistent with TTB as an 'excellent result' while others might regard the failure to fit 36% as a considerable problem (See Box 2). To advance the debate and increase the testability of the heuristics, clearer specification as to what constitutes a good fit between a heuristic and data is required. Furthermore, it is essential to incorporate an error theory into the 'toolbox' to account for the stochastic deviation from the heuristics' deterministic rules. Without such advances we risk getting mired in arguments about the 'fullness of the glass' [14].

An adjustable spanner?

Perhaps there is another way? Could these patterns of individual variability be reconciled within a single model rather than attributing behavior to different

heuristics? Lee and Cummins [25] who recently reported similar patterns of individual variability in a test of TTB, suggested that the beginnings of a unifying explanation could come from recognizing that behavior that conforms to heuristics are special cases of a more general approach to decision-making. They developed this argument by presenting an evidence-accrual model of decision-making, which relies on a random walk sequential sampling process. Such sequential sampling models have been applied to a wide range of tasks, and although specific mechanisms differ, in general the models assume that rather than "taking" a predetermined quantity of information, sampling of each option occurs until evidence sufficient to favor one option over the other has been accumulated [e.g., 25, 36-40].

The important feature of an evidence accumulation model in the context of a binary choice problem is that it can "mimic" the performance of TTB's stopping rule or the recognition heuristic, or a strategy that incorporates more evidence (e.g., a weighted additive rule) through adjustments of the evidence required before a decision is made (see Figure 1.) Thus one way of explaining individual variability is to suggest that all participants use an evidence-accrual strategy but that some participants require greater amounts of evidence than others before making their decisions [25]. Lee and Cummins [25] found that such a unified model accounted for 84.5% of the decisions made by participants – more than that accounted for by either TTB or a compensatory strategy alone. Importantly, through the application of model selection criteria, it was demonstrated that the improved accuracy was *not* due to the additional complexity of the unified model (i.e., its two free parameters compared to the parameter-free TTB). Maybe we are all using the same tool – an *adjustable spanner* perhaps?

Of course, for such a unified model to become a favored alternative explanation, we need to develop empirical techniques that allow us to distinguish between the two accounts [e.g., 25], and also be able to specify how the evidence threshold is affected by factors such as the cost of information, time pressure, the cost and benefits of correct and incorrect decisions, and perhaps the effect of personality characteristics such as intelligence [18, 25, 38, 40].

Furthermore, even if this empirical challenge proves difficult, the parsimony afforded by such a general-purpose 'adjustable spanner' over a 'toolbox of heuristics' should not be underestimated. If one of our goals in science is data reduction, then arguably we are more likely to achieve this goal by constraining a model through the empirical specification of a parameter, than through generating a 'periodic table' of heuristics.

A similar call for unification of the diverse range of heuristics, biases and modes of processing regularly discussed in social cognition into a 'unimodel' has recently been made by Kruglanski and his colleagues [41, 42]. The central component of this 'unimodel' is the accumulation of information to serve as evidence for an assessment. An exciting challenge is to explore the commonalties between these unifying approaches.

A revision?

The attraction of simple algorithms that incorporate the constraints of human cognitive capacity and the structure of the environment has endured in cognitive science arguably since its inception as a discipline [43]. The recent 'vision' of Gigerenzer and colleagues has continued this eminent tradition. However, with largely equivocal evidence for the only heuristic to undergo detailed empirical

scrutiny, claims for a 'toolbox' of ecologically rationally heuristics that are constructed and *triggered into action by the properties of particular environments* have been weakened.

Where does this leave the vision of rationality? Certainly the principle tenets of a boundedly rational perspective – cognitive and environmental limitations – are not undermined by this critique but it seems that the formulation of the heuristics and the criteria used to judge their 'success' requires considerable revision. A starting point, perhaps, would be to use the techniques adopted by Anderson and colleagues in their rational analysis approach to human cognition [44, 45, 46]. The goals of Gigerenzer's and Anderson's approaches are similar – understanding cognition given environmental and human constraints – but the methods adopted differ crucially.

Anderson et al. use formal models (e.g., Bayes' theorem) to derive "optimal" behavior functions (given cognitive and environmental limitations), *prior* to making predictions about cognitive algorithms that might be appropriate for approximating the "optimal" function. This technique has proved successful in elucidating *general mechanisms* in cognition [e.g., 46] – and, because, the "optimal" function is derived *apriori*, an explanation of *why* an algorithm performs well [Cf. 26]. In contrast, Gigerenzer et al. propose heuristics *first* – in the absence of empirical evidence – and then, if the heuristics do well, 'explain' this success after the fact by the potentially circular reference to the 'fit' with an environment.

Perhaps our vision of rationality would be clearer if we understood both *which* algorithms do well and *why*. That is, if we applied the correspondence criterion of rationality to find out *which*, and the coherence criterion to find out *why*, our view of rationality might be much improved. We might even discover that a single 'tool' can do rather well – and why.

Box 1.

Simulating Plausibility?

The argument for the psychological reality of the tools in the adaptive toolbox is based on an appeal to their plausibility as mechanisms for inference. The argument goes something like this: The methods of classical rationality are computationally intractable and time consuming and thus beyond the bounds of human decision makers; in contrast simple mechanisms like TTB can be carried out under conditions of limited time and knowledge; simulations showing that simple models like TTB often match or outperform competing rational models in terms of accuracy are thus 'proof' that the fast and frugal models provide better accounts of human decision processes [5, 14, 26, 27].

A number of questions have been raised regarding this appeal to plausibility. First, how simple are the models? It is true that TTB can be *described* as a simple three-step algorithm but its successful *execution* relies on a large amount of precomputation. For example, before search can begin cues need to be hierarchically organized in validity order – how is such a hierarchy constructed? [24, 25] Second, should we be persuaded by the argument that TTB is fast because it searches for fewer pieces of information? The interpretation of speed relies on assumptions about the architecture of the cognitive system. A serial architecture would show an advantage for TTB, but in a parallel architecture (for which there is ample converging evidence [49]) large amounts of information can be searched simultaneously and therefore speed and amount of information will be unrelated [27].

More generally, there are a number of 'equally plausible' cognitive algorithms that perform just as well if not better than TTB on the types of binary decision

problems for which TTB is designed, in addition to being applicable to a range of other problems. Many of these models use completely different strategies to TTB such as relying on the similarity between current and previously seen exemplars [23, 26, 48].

Given the performance of these other models, some of which appear to have more convincing empirical support than TTB (e.g.,[23]), it seems we need to be cautious in accepting the prima facie plausibility of TTB as evidence of its psychological reality. Exploring the relation between exemplar-based models, TTB and evidence-accumulation models is an important and exciting avenue for future research.

Box 2.

Inferences from memory or 'givens'?

Gigerenzer and Todd [5] described tasks involving 'inferences from givens' - in which all relevant information is provided by the experimenter - as unsuitable for testing models of ecological rationality. This is because information search *in memory* and its accompanying cognitive costs are claimed to be crucial triggers for the adoption of simple heuristics. Many studies have attempted to simulate this costly memory search by requiring participants to pay real money to acquire information about alternatives under consideration [e.g., 12-15].

However, in an ingenious set of studies Bröder and Schiffer [17] set out to test this *memory-search hypothesis* more directly. Participants were asked to solve a murder. In an extensive learning phase they acquired knowledge about the attributes of 10 suspects (e.g., type of clothing, breed of accompanying dog, etc.). Following learning, a cue validity-hierarchy was established by telling participants how many witnesses had agreed about certain attributes. In the inference phase two suspects were presented without any accompanying attribute information and participants had to infer which was more likely to have committed the murder. Thus during inference all information pertaining to the suspects had to be retrieved from *memory*.

Claims for memory search being a crucial component for simple heuristic use were corroborated. In four experiments Bröder and Schiffer found more inferences consistent with the use of TTB – up to 64% in one experiment – than other compensatory decision strategies. In addition there was some evidence that TTB is relied on more when attribute information is presented in word lists than as images, suggesting that representational format is an important moderating variable for strategy use.

More of this type of research is necessary to establish additional boundary conditions on the adaptive toolbox framework. Without such conditions, it is impossible to evaluate the adequacy of the proposed models of decision processes. If the framework *only* applies for inferences from memory then this needs to be clearly stated. Without knowing the range of applications it is hard to falsify the theory [17].

Acknowledgments

The writing of this article was supported by a Faculty Research Grant from The University of New South Wales. The work was part of the program of the ESRC Research Centre for Economic Learning and Social Evolution (ELSE). I am indebted to the ELSE fellows for many illuminating discussions especially David Shanks, David Lagnado, and Mark Johansen. I also thank Tim Rakow for his numerous insights, and for suggesting the label 'adjustable spanner', and the anonymous reviewers for extremely helpful comments on an earlier draft of the manuscript.

References

- 1. Hammond, K.R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. Oxford University Press.
- 2. Tversky, A and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and Biases. *Science*, *185*, 1124-1131.
- 3. Kahneman, D and Tversky, A. (1996). On the reality of cognitive illusions. *Psychol. Rev.* 103, 582-591.
- 4. McKenzie, C.R.M (2003). Rational models as theories not standards of behaviour. *Trends Cogn. Sci.* 7, 403-406.
- Gigerenzer, G. and Todd, P.M. (1999). Fast and frugal heuristics: The adaptive toolbox. In *Simple heuristics that make us smart* (Gigerenzer, G., Todd, P.M. and the ABC Research Group) pp 3-34, Oxford University Press.
- 6. Chase, V. et al. (1998) Visions of rationality. Trends Cogn. Sci. 6, 206-214.
- 7. Brunswik, E., et al. (2000). *The essential Brunswik: Beginnings, explications, and applications*. Oxford University Press.
- 8. Simon, H.A. (1956) Rational choice and the structure of the environment. *Psychol. Rev.* 63, 129-138.
- 9. Gigerenzer, G. (2001) The adaptive toolbox. In *Bounded Rationality: The adaptive toolbox* (G. Gigerenzer and R. Selten, eds.), pp 37-51, MIT Press.
- 10. Gigerenzer G. and Goldstein D. (1996) Reasoning the fast and frugal way: models of bounded rationality. *Psychol. Rev.* 103, 650-669.
- 11. Goldstein D. and Gigerenzer G. (2002) Models of ecological rationality: The recognition heuristic. *Psychol. Rev.* 109, 75-90.

- 12. Newell, B.R. and Shanks, D.R. (2004) On the role of recognition in decision making. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 923-935.
- 13. Newell B.R. and Shanks, D.R. (2003) Take the best or look at the rest? Factors influencing "one-reason" decision making. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 53-65.
- 14. Newell B.R., *et al.* (2003) Empirical tests of a fast and frugal heuristic: Not everyone "takes-the-best". *Organ. Behav. Hum. Decis. Process.* 91, 82-96.
- 15. Newell, B.R., *et al.* (2004) Search strategies in decision making: The success of success. *J. Behav. Decis. Mak.* 17, 117-137.
- 16. Bröder, A. (2000). Assessing the empirical validity of the "Take-the-Best" heuristic as a model of human probabilistic inference. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1332-1346.
- 17. Bröder, A. and Schiffer S. (2003) "Take-the-best" versus simultaneous feature matching: Probabilistic inferences from memory and the effects of representation format. *J. Exp. Psychol. Gen.* 132, 277-293.
- 18. Bröder, A. (2003) Decision making with the adaptive toolbox: Influence of environmental structure, personality, intelligence, and working memory load. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 611-625.
- 19. Oppenheimer, D.M. (2003) Not so fast (and not so frugal!) Rethinking the recognition heuristic. *Cognition*, 90, B1-B9.
- 20. Juslin, P.J., Jones, S., Olsson, H., & Winman, A. (2003). Cue abstraction and exemplar memory in categorization. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 924-941.
- 21. Boyd M. (2001) On ignorance, intuition and investing: A bear market test of the recognition heuristic. *J. Psychol. Finan. Market.* 2, 150-156.

- 22. Over, D.E. (2000) Ecological rationality and its heuristics. *Think. Reason.* 6, 182-192.
- 23. Juslin, P. and Perrson, M. (2002) PROBabilities from EXemplars (PROBEX):
 A "lazy" algorithm for probabilistic inference from generic knowledge. *Cogn. Sci.* 95, 1-45.
- 24. Rakow, T. *et al.* (2004) Simple heuristics from the adaptive toolbox: Can we perform the requisite learning? *Think. Reason.* 10, 1-29.
- 25. Lee M.D. and Cummins T.D.R. (2004) Evidence accumulation in decision making: Unifying the 'Take-the-best' and the 'rational' models. *Psychon. Bull. Rev.*11, 343-352.
- 26. Chater N., *et al.* (2003) Fast, frugal and rational: How rational norms explain behaviour. *Organ. Behav. Hum. Decis. Process.* 90, 63-80.
- 27. Chater, N. (2000) How smart can simple heuristics be? *Behav. Brain Sci.*, 23, 745-746.
- 28. Cooper, R. (2000) Simple heuristics could make us smart but which heuristics do we apply when? *Behav. Brain Sci.*, 23, 746.
- 29. Oaksford, M (2000) Speed, frugality and the empirical basis of Take-the-Best. *Behav. Brain Sci.*, 23, 760 -761.
- 30. Shanks, D.R. and Lagnado, D.A. (2000) Sub-optimal reasons for rejecting optimality. *Behav. Brain Sci.*, 23, 761-762.
- 31. Elwyn, G. et al. (2001). Decision analysis in patient care. *The Lancet*, 358, 571-574.
- 32. Dhami, M-K. & Ayton, P. (2001). Bailing and jailing the fast-and-frugal way. *J. Behav. Decis. Mak.* 14, 141-168.

- 33. Borges, B. *et al.* (1999) Can ignorance beat the stock market? In *Simple heuristics that make us smart* (Gigerenzer, G., Todd, P.M. and the ABC Research Group) pp 59-72 Oxford University Press.
- 34. Ortmann, A. et al. (in press) The recognition heuristic: A fast and frugal way to investment choice? In *Handbook of Experimental Economics Results* (C.R. Plott & V.L. Smith, eds), Elsevier.
- 35. Goldstein *et al.* (2001) Why and when do simple heuristics work? In *Bounded Rationality: The adaptive toolbox* (G. Gigerenzer and R. Selten, eds.), pp 173-190, MIT Press.
- 36. Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59-108.
- 37. Nosofsky, R.M. and Palmeri, T.J. (1997) An exemplar-based random walk model of speeded classification. *Psychol. Rev.* 104, 266-300.
- 38. Busemeyer, J.R. and Townsend, J.T. (1993) Decision field theory: A dynamic cognition approach to decision making. *Psychol. Rev.* 100, 432-459.
- 39. Smith, P.L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *J. Math. Psychol.* 44, 408-463
- 40. Bockenholt, U. et al. (1991) The effects of attractiveness, dominance and attribute differences on information acquisition in multiattribute binary choice.
 Organ. Behav. Hum. Decis. Process. 49, 258-281.
- 41. Erb, H.P, *et al.* (2003). Searching for commonalities in human judgment: The parametric unimodel and its dual mode alternatives. *Eur. Rev. Social Psychol.* 14, 1-47.
- 42. Kruglanski, A.W. (2004) The quest for the gist: On challenges of going abstract in social and personality psychology. *Pers. Soc. Psychol Rev.* 8, 156-163.

- 43. Miller, G.A. (2003). The cognitive revolution: a historical perspective. *Trends Cogn. Sci.* 7, 141-144.
- 44. Anderson, J.R. (1990) The adaptive character of thought. Erlbaum.
- 45. Anderson, J.R. and Milson, R. (1989) Human memory: An adaptive perspective. *Psychol. Rev.* 96, 703-719.
- 46. Anderson, J.R. and Schooler, L.J. (1991) Reflections of the environment in memory. *Psychol. Sci.* 2, 396-408.
- 47. Rumelhart, D.E., McClelland, J. and the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the microstructure of cognition (Vols 1 & 2)*, MIT Press.
- 48. Dougherty, M.R.P. *et al.* (2003) Memory as a fundamental heuristic for decision making. In *Emerging perspectives on judgment and decision research* (S.L. Schneider and J. Shanteau, eds) pp 125-164, Cambridge University Press.

Figure Caption

Figure 1. Perspectives on the 'tools' of decision making

The 'adaptive toolbox' perspective proposes discrete heuristics depicted by individual boxes in the diagram (e.g., 'recognition heuristic', 'take-the-best heuristic', 'take-the-best-two' and so on). A different heuristic is used for different problems, and the 'selection' of the heuristic is thought to be largely driven by the environment. The 'adjustable spanner' perspective suggests that only one tool is used and that different thresholds of accumulated evidence give rise to patterns of data that 'mimic' the stopping rule of the heuristics. This mimicking is illustrated by the red-dashed line where the amount of accumulated evidence is consistent with the use of "Take-the-best".

Figure 1

