Available online at www.sciencedirect.com
SCIENCE DIRECT®
ACADEMIC PRESS
Organizational Behavior and Human Decision Processes 90 (2003) 63–86
ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
www.elsevier.com/locate/obhdp

Fast, frugal, and rational: How rational norms explain behavior

Nick Chater,[a,*] Mike Oaksford,[b] Ramin Nakisa,[c] and Martin Redington[c]

[a] *Department of Psychology, University of Warwick, Coventry CV7 4AL, UK*
[b] *School of Psychology, Cardiff University, P.O. Box 901, Cardiff CF1 3YG, UK*
[c] *Department of Experimental Psychology, University of Oxford, Oxford, UK*

Abstract

Much research on judgment and decision making has focussed on the adequacy of classical rationality as a description of human reasoning. But more recently it has been argued that classical rationality should also be rejected even as normative standards for human reasoning. For example, Gigerenzer and Goldstein (1996) and Gigerenzer and Todd (1999a) argue that reasoning involves "fast and frugal" algorithms which are not justified by rational norms, but which succeed in the environment. They provide three lines of argument for this view, based on: (A) the importance of the environment; (B) the existence of cognitive limitations; and (C) the fact that an algorithm with no apparent rational basis, Take-the-Best, succeeds in an judgment task (judging which of two cities is the larger, based on lists of features of each city). We reconsider (A)–(C), arguing that standard patterns of explanation in psychology and the social and biological sciences, use rational norms to explain why simple cognitive algorithms can succeed. We also present new computer simulations that compare Take-the-Best with other cognitive models (which use connectionist, exemplar-based, and decision-tree algorithms). Although Take-the-Best still performs well, it does not perform noticeably better than the other models. We conclude that these results provide no strong reason to prefer Take-the-Best over alternative cognitive models.

1. Introduction

Research on human judgment and decision making has frequently focussed on the relationship between observed human reasoning and classical rational norms. Instances where actual reasoning and classical norms diverge have been taken to exemplify cognitive biases; human performance is viewed as failing to measure up, in some ways and under some circumstances, to classical rational norms. For example, people appear to persistently fall for logical blunders (Evans, Newstead, & Byrne, 1993), probabilistic fallacies (e.g., Tversky & Kahneman, 1974), to make inconsistent decisions (Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1986), and to make irrational moves in games (Colman, 1995). Indeed, the concepts of logic, probability, decision theory and the like do not appear to mesh naturally with our everyday reasoning strategies: these notions took centuries of intense intellectual effort to construct, and present a tough challenge for each generation of students. From this perspective, the gap between observed decision making behavior and classical rationality may appear a yawning gulf.

Various factors have been viewed as contributing to this gulf: performance errors, computational limitations, and differences between the understanding of the task employed by experimenter and experimental participant (e.g., Ayton & Hardman, 1997; Cohen, 1981; Oaksford & Chater, 1993; Stanovich, 1999; Stanovich & West, 2000; Stein, 1996). There have also been persuasive arguments that individual differences concerning cognitive ability and/or educational background can substantially affect the gap between observed behavior and classical norms. Where cognitive ability is high and/or the task constraints do not severely impact cognitive limitations of memory and attention, people may on occasion conform quite well to classical rational norms, even on tasks where performance is typically viewed as systematically irrational (Stanovich & West, 1998a, 1998b, 1998c). This line of argument suggests that the gulf between performance

* Corresponding author.
E-mail address: nick.chater@warwick.ac.uk (N. Chater).

0749-5978/03/$ - see front matter
doi:10.1016/S0749-5978(02)00508-3

and classical rationality may be bridged, at least for some individuals, in some circumstances (Stanovich, 1999, although see, for example, Ayton, 2000; Hertwig, 2000).

Recently, a number of theorists have suggested a more radical alternative—that comparing human behavior against classical rationality is like comparing apples and oranges. Whereas the traditional viewpoint with psychology and economics is that the comparison is appropriate, but that the disparity between the two may be substantial, the radical viewpoint suggests that the comparison itself is misconceived.

Evans and Over (1996, 1997), for example, distinguish between two notions of rationality:

> Rationality$_1$: Thinking, speaking, reasoning, making a decision, or acting in a way that is generally reliable and efficient for achieving one's goals
>
> Rationality$_2$: Thinking, speaking, reasoning, making a decision, or acting when one has a reason for what one does sanctioned by a normative theory. (Evans & Over, 1997, p. 2)

Crucially, Evans and Over argue that these two kinds of rationality are largely independent: "people are largely rational in the sense of achieving their goals (rationality$_1$) but have only a limited ability to reason or act for good reasons sanctioned by a normative theory (rationality$_2$)" (Evans & Over, 1997, p. 1). If this is right, then achieving one's goals can be achieved without following the precepts of classical rationality—i.e., without there being a justification for the actions, decisions or thoughts which lead to success: rationality$_1$ does not require rationality$_2$. That is, Evans and Over are committed to the view that thoughts, actions or decisions which cannot be normatively justified using classical rational norms can, nonetheless, consistently lead to practical success.

Relatedly, Gigerenzer and his colleagues (e.g., Gigerenzer, 2000; Gigerenzer & Goldstein, 1996; Gigerenzer & Todd, 1999a) have developed a major research program on human judgment that attempts to break out of the restrictive mould of comparing human performance with classical rationality. Like Evans and Over (1996, 1997), Gigerenzer and colleagues have argued that, instead, inference should be assessed in terms of its success in solving ecologically relevant problems in natural environmental contexts: "the minds of living systems should be understood relative to the environment in which they evolved rather than to the tenets of classical rationality [i.e., probability theory, expected utility theory and so on]..." (Gigerenzer and Goldstein, p. 651) (emphasis added). The proposal is that the point of reasoning is to allow people to deal with the everyday world, rather than conforming with rational norms.

The aim of this paper is to consider the viability of this radical proposal. For concreteness, we focus on a particularly influential formulation, by Gigerenzer and Goldstein (1996), but we intend our analysis to apply more generally to accounts of ecological and classical norms of rationality.

Gigerenzer and Goldstein (1996) give three lines of argument that an ecological standard of rationality (i.e., that reasoning gets good results in the real world) should replace classical norms as the appropriate comparison for human reasoning, based on: (A) the importance of the environment; (B) the existence of cognitive limitations; and (C) an 'existence proof,' i.e., a specific algorithm, Take-the-Best that exemplifies their approach. Take-the-Best succeeds in a real environment, even though it has no apparent rational justification.

Gigerenzer and colleagues also argue that Take-the-Best is also more than a mere existence proof—it is intended to be a cognitively plausible model of a specific kind of cognitive estimation (e.g., Gigerenzer & Goldstein, 1996; Gigerenzer, 2000). The task Gigerenzer and Goldstein consider is that of two alternative forced choice concerning which of two German towns has the larger population, based on a set of nine binary 'features' of each town (e.g., "has a soccer team," "is a state capital," etc.). Gigerenzer and Goldstein (1996) consider the computational problem of learning how to predict which of two cities is the larger, from a 'training set' of cities, their features, and populations. From the point of view of conventional statistics, an 'obvious' way to proceed in such a task is to attempt to use some form of regression (e.g., linear regression) to assess the influence of each of the features on city size. When presented with a forced choice test, the regression might then be used to integrate all the features of the two cities, to come to an overall conclusion concerning which is likely to be the larger.

Gigerenzer and Goldstein's Take-the-Best algorithm, however, takes a radically different approach. It has two steps. The first routine, the recognition principle states that, if a reasoner recognises the name of one city but not the other, then the first city should be assumed to be the larger–no further memory search is carried out. If the recognition principle does not decide the issue, Take-the-Best moves to a second and more elaborate routine (on which we concentrate below). Features of the cities are considered in order, one-by-one, from the feature most diagnostic of city size to the feature that is least diagnostic of city size (where diagnosticity is calculated as the probability that the feature will correctly signal which is the larger of two randomly chosen cities which differ on this feature). As soon as a feature is found on which the cities differ (e.g., one city has a soccer team but the other does not), then the feature is used to decide which city is the larger (the city with the soccer team) and the calculation terminates. This means that the decision is based on a single feature, rather than attempting to 'integrate' all the different features of the two cities; and indeed many of the features of the cities are not even considered in the decision. Gigerenzer and Goldstein (1996) showed that, despite this very 'frugal' use of information, Take-the-Best performs impressively. In a computational 'competition' using real features of German cities, Take-the-Best performs as well as linear regression and a range of approximations to linear regression. Subsequent computational simulation work has successfully generalized these findings to a remarkable range of domains, ranging from judgments of levels of homelessness based on features of US cities, to judgments of house prices, professors' salaries, the amount of time 35 species of mammal spend asleep based their biological features, and many more (Czerlinski, Gigerenzer, & Goldstein, 1999).

Take-the-Best has interesting antecedents in previous work in the literature of behavioral decision making. Non-compensatory strategies (i.e., those that use aspects of individual features to make a decision, rather than integrating all the features given) have been widely discussed (e.g., Einhorn, 1970, 1971; Ganzach, 1995), including 'elimination by aspects' (Tversky, 1972) and the lexicographic heuristic (Tversky, 1969).

More broadly, the Adaptive Decision Maker research program of Payne and colleagues (Payne, 1976; Payne, Bettman, & Johnson, 1988, 1990, 1993; Payne, Bettman, & Luce, 1996), has emphasized that the decision maker can strategically choose between a range of decision making methods—many of which will be 'fast and frugal.' In the Adaptive Decision Maker framework, though, fast and frugal algorithms are one end of a continuum of options from which the decision maker may choose—given sufficient time, cognitive resources and motivation, participants may choose strategies which integrate the information that they have been given in more elaborate ways (although not necessarily with better decision making results). Work within this tradition focuses in considerable detail on the conditions under which particular 'fast and frugal' methods are applied, and under what conditions these methods are successful (Payne et al., 1993). One difference of emphasis between the two approaches is that the Adaptive Decision Maker program has been concerned primarily with understanding how people make choices between options, where Gigerenzer and colleagues have focussed primarily on questions of judgment: i.e., tasks in which people have to judge, on limited information, which of two states of the world holds. Although, from a normative point of view, the domains of choice and judgment are very different (roughly, the normative theory for choice is utility theory; the normative theory for judgment is decision theory), it is quite possible that some of the underlying cognitive algorithms used the two cases are closely related. Indeed, Tversky's (1969) lexicographic heuristic, mentioned above, is very closely related to Take-the-Best.[1]

The Adaptive Decision Maker framework stresses the flexible use of cognitive heuristics and strategies. By contrast, Gigerenzer and Goldstein (1996) do not explicitly discuss whether they see Take-the-Best as a universal cognitive algorithm, or as being selected dynamically by decision makers from a range of decision making methods. But the latter position appears to be embodied in the idea of the 'adaptive toolbox' (Gigerenzer & Todd, 1999b; Gigerenzer, 2000, 2001; Gigerenzer & Selten, 2001). The adaptive toolbox is "the collection of specialized cognitive mechanisms that evolution has built into the human mind for specific domains of inference and reasoning, including fast and frugal heuristics" (Gigerenzer & Todd, 1999b, p. 30). Gigerenzer and Todd (1999b) suggest that, for example, there may be adaptive selection from within the family of Take-the-Best-type algorithms (and presumably also outside this family), depending on, among other things, the kinds of factors analysed in the Adaptive Decision Maker framework (Payne et al., 1993). We shall see, at the end of this paper, that the empirical evidence concerning Take-the-Best is most consistent with this type of interpretation.

The two parts of this article focus on general and specific issues in turn. In the first part, we begin by outlining the central role for rational norms in the explanation of human behavior. This role for rational norms is quite different from that embodied in "classical" or "unbounded" rationality (Gigerenzer & Goldstein, 1996; Gigerenzer & Todd, 1999b), but we believe it underlies explanation throughout much of the social and biological sciences. In particular, we argue that norms of classical rationality are crucially involved in explaining why a particular behavior is ecologically successful. Thus, we argue that classical and ecological notions of rationality are complementary, rather than standing in competition. With this analysis in mind, we reevaluate and counter the three arguments (A)–(C) that Gigerenzer and Goldstein (1996) give for the thesis that an ecological notion of rationality should replace classical rationality. The second part of this article conducts a new 'competition,' between Take-the-Best and a range of algorithms based on existing cognitive architectures widely used in cognitive science and artificial intelligence,

[1] Specifically, the lexicographic rule compares two choice options by considering features of each option, one by one, in descending order of importance. If one option 'wins' on a particular feature, it is chosen; if there is a tie, the next most important feature is chosen, and so on. The core component of Take-the-Best makes a judgment between two options—e.g., which is the largest city—by considering cues in descending order of 'cue validity' concerning that judgment; again, if one option 'wins' on a cue, it is chosen (e.g., judged to be the largest city); if there is a tie, the next most 'valid' cue is chosen.

and considers theoretical arguments which appear to favor Take-the-Best. We conclude that, while Take-the-Best's performance is again impressive, there is no strong reason to view Take-the-Best as having greater cognitive plausibility than a range of other algorithms. This competition also raises questions concerning the relationship between research on judgment and decision making and computational and experimental work on basic mental processes in cognitive psychology (e.g., Dougherty, Gettys, & Ogden, 1999; Weber, Goldstein, & Barlas, 1995).

2. Rationality and the explanation of behavior

We have seen that some leading theorists (e.g., Evans & Over, 1996, 1997; Gigerenzer & Goldstein, 1996; Gigerenzer & Todd, 1999b) have argued that classical rational principles have no useful role in explaining everyday behavior; that classical rationality and 'ecological' rationality are entirely separate domains. One implication is that classical rational principles, of logic, probability, decision theory, and game theory do not even provide normative standards against which everyday behavior can be compared—because to make such a comparison is to compare apples and oranges. Everyday behavior is properly judged by its results, rather than conformity to abstract standards of reasoning. If their analysis is correct, then it appears to have large ramifications across the social and biological sciences, where classical rational principles are frequently used to explain everyday behavior. But, we suggest, this apparently radical perspective is rooted in an incorrect characterization of the role that classical rational principles play in explanation in the social and biological sciences. Specifically, critics of the application of classical rational principles typically assume that such principles explain behavior by assuming that the mind performs rational 'calculations.' We believe that is a crucial mischaracterization of the project of rational explanation in the social and biological sciences, which aim for rational description, without ascribing rational calculations to the cognitive system.

2.1. Rational calculation vs. rational description

Gigerenzer and Goldstein's characterization of the classical view—that rational norms are the laws of thought—is intended to encompass the broad sweep of rational explanation of behavior across several disciplines.

> From Pierre Laplace to George Boole to Jean Piaget, many scholars have defended the now classical view that the laws of human inference are the laws of probability and statistics... Following this time-honored tradition, much contemporary research in psychology, behavioral ecology, and economics assumes standard statistical tools to be the normative and descriptive models of inference and decision making. (p. 650)

Gigerenzer and Todd (1999b, p. 9) amplify the point:

> Unbounded rationality is a strange and demanding beast. On the one hand, researchers who envision rationality in this way accept the difference between God, or Laplace's [hypothetical] superintelligence, and mere mortals. Humans must make inferences from behind a veil of uncertainty, but God sees clearly; the currency of human thought is probabilities, whereas God deals in certitude. On the other hand, where it comes to how they think these uncertain inferences are executed, those who believe in unbounded rationality paint humans in God's image. God and Laplace's superintelligence do not worry about limited time, knowledge, computational capacities. The fictional, unboundedly rational human mind does not either...

The idea that rational explanation presupposes that rational calculation (and implausibly vast amounts of such calculation) is conducted by the human mind fits well with Evans and Over's (1996, 1997) viewpoint that classical rational norms explain an agent's behavior only when the agent understands the relevant normative justification—this is their rationality$_2$, above. Their view is that rationality$_2$ explanation requires that an agent possesses rational norms and can calculate their implications for the particular decision being faced. Evans and Over point out, echoing the quotes above, that the complexity of these calculations is likely to exceed the capacity of the cognitive system, for most interesting real world reasoning problems (see also Oaksford & Chater, 1991, 1993).

We suggest that the view that rational explanation requires that people themselves carry out the relevant rational calculations is a fundamental mischaracterization of how rational principles are used to explain thought and behavior in behavioral ecology, economics, and psychology. Instead of being committed to what we shall call rational calculation, researchers in these disciplines are actually committed to a very different and more modest claim: rational description.[2,3]

Rational calculation is the view that the mind works by carrying out probabilistic, logical, or decision-theoretic operations. Gigerenzer and Goldstein (1996) endorse this reading of the role of rationality in the social

[2] The terminological picture is further complicated by the fact that in discussing their competition between algorithms, Gigerenzer and Goldstein (1996) carefully define algorithms to be "rational"—with quotation marks—if they use all informational available to them, and if they combine all this information together (p. 657). Gigerenzer and Goldstein quite deliberately set this usage apart against general notions of rationality, such as we are discussing here. To minimize confusion, we shall not refer to "rational algorithms" here.

[3] There are direct relations between this distinction and the general distinction between rule-described and rule-following behavior (Chomsky, 1980; Hahn & Chater, 1998; Kripke, 1982).

and biological sciences, arguing that theorists in these domains are implicitly committed to the view that the mind is a "supercalculator...—carrying around the collected works of Kolmogoroff, Fisher, or Neyman" (p. 651). Rational calculation is explicitly avowed by relatively few theorists, though it has clear advocates with respect to logical inference: Mental logicians propose that much of cognition is a matter of carrying out logical calculations (e.g., Braine, 1978; Inhelder & Piaget, 1958; Rips, 1994).

Rational description, by contrast, is the view that behavior can be approximately described as conforming with the results that would be obtained by some rational calculation. This view does not assume (though it does not rule out) that the thought processes underlying behavior involves any rational calculation. An analogy may be useful: the wings of a bird may approximate the results of a rational calculation of optimal aerodynamic design. Moreover, this observation helps explain why the wing has the structure that it does; but there is, of course, no presumption that the bird conducts any calculations in designing its wing.

Behavioral ecologists extend this pattern of biological explanation from anatomy and physiology to behavior. They attempt to explain an animal's strategies for foraging, defending territory, or choosing mates, by showing that these can be approximately described as the results of a rational calculation of optimal behavior. There is no presumption that the animal carries out complex probabilistic or decision-theoretic calculations underlying this behavior, any more than it rationally calculates the optimal length of its incisors, or the optimal structure of its lungs. Indeed, behavioral ecologists expressly disavow a rational calculation interpretation of their theories, as being patently at variance with the cognitive limitations of the animals they study (e.g., McFarland & Houston, 1981).

Contemporary economics also aims to explain behavior by rational description, rather than rational calculation. Indeed, the emphasis in economic explanation in the middle and latter part of the past century has been to attempt to minimize psychological claims as far as possible (Loewenstein, 1992). For example, many nineteenth century theorists, drawing on the philosophical tradition of utilitarianism, viewed utility as a psychological construct; and they hypothesized that economic agents act in order to maximize their utility—hence at least some calculations concerning utility would seem to be inevitably attributed to the agent. But the twentieth century has seen the development of the 'revealed preference' interpretation of utility (Samuelson, 1937)—utilities are defined in terms of people's patterns of preferences. In this interpretation, utility is a behavioral (and economically observable) construct, rather than a component of internal mental calculations. This behavioral approach has been extended to expected utility theory (von Neumann & Morgenstern, 1944) and to subjective probability (Anscombe & Aumann, 1963; Savage, 1954). Patterns of observable choice behavior are used to attribute utilities and probabilities to an economic agent. Crucially, there is no assumption that these utilities or probabilities are internal to the agent; and hence, a fortiori, there is no assumption that agents actually engage in probabilistic or decision-theoretic calculations.

There is, moreover, a recognition in economics that applying rational theories, such as probability theory, expected utility theory, and game theory will only provide an approximation model of people's behavior. Economists allow that "Faced with complexity, individuals resort to rules of thumb, to 'back of the envelope' calculations, to satisficing behavior..." (Kreps, 1990, p. 119). Economists thus recognise that behavior only approximates to rationality; but economic theory typically idealises away from such limitations. Various justifications for these idealizations have been proposed, ranging from making a virtue of severe idealization, so long as the resulting theory makes successful predictions (Friedman, 1953), to the view that 'errors' in individual agents will cancel out on aggregate, to the view that errors be gradually eliminated in contexts where individuals can learn (e.g., Cyert & de Groot, 1974; de Canio, 1979), or eliminated in competitive markets (see Akerlof & Yellen, 1985, for theoretical analysis). There are also those skeptical of much economic theory, who are unpersuaded by any of these arguments (e.g., Nelson & Winter, 1982; Simon, 1959). For our purposes, the important point is that most economists interpret their theories as about rational description, but not rational calculation (at best, people act 'as if' they made rational calculations; but they do not actually perform such calculations); and they agree furthermore that actual behavior is only an approximation to rational standards.[4]

This leads us naturally to consider the psychology of human reasoning, judgment and decision making. The

[4] It is also true, of course, that there are often many aspects of purely descriptive psychological models, e.g., of processes in categorization or memory, in which some aspects of the machinery of the model is not assumed to be calculated by the cognitive system, but is instead a description of cognitive processes. For example, in models that involve retrieval of stored instances or traces from memory, a measure of mental 'distance' between present and stored item is often used to predict categorization or memory performance (e.g., Nosofsky, 1986). But it is typically assumed that the cognitive system does not itself calculate this distance, any more than planets need to compute the 'distance' between them to calculate the gravitational forces between them. Moreover, many cognitive models are formulated in a way that leaves vague the distinction between aspects of the model that are intended to be calculated by the cognitive system, rather than as descriptions of its operation. We thank an anonymous reviewer for raising this point.

question at issue is how well people's behavior fits with rational models based on probability, expected utility theory, or game theory—and hence how far the rationality assumptions in economic models idealise away from actual human behavior. Thus, crucially, as in economics, the issue again concerns rational description—how well behavior conforms to rational description—rather than how far people make rational calculations. Only rarely is any attempt made to assess whether people are actually carrying out internal probabilistic, decision- or game-theoretic calculations (e.g., Gigerenzer & Hoffrage, 1995)—generally, the working hypothesis seems to be that they are not.[5]

We have seen that theoretical economists have assumed that the predictions of rational theories are only followed to an approximation. Laboratory studies of individual behavior, by both psychologists and experimental economists, have confirmed this assumption. Indeed, they have revealed that behavior can show many substantial and systematic departures from normative theories. People appear to fall for numerous probabilistic fallacies (Kahneman et al., 1982), make predictable logical blunders (Evans, 1982, 1989), show inconsistent preferences with (Allais, 1953; Ellsberg, 1961) and without (May, 1954) uncertainty, and to fail to adopt the predicted "Nash equilibria" (Nash, 1950) in game theory (Flood, 1958; Ledyard, 1995). The practical question of interest for theorists applying rational principles to explain behavior is not whether the mind is carrying out the rational calculations that they postulate—in most cases, at least, it seems inconceivably unlikely that this is true. Instead, the real question is whether and how the broad outlines of human behavior in everyday contexts can usefully be described in rational terms at all.

So, we suggest, across economics, biology, and psychology, the working assumption of those producing rational explanations is that these explanations aim to explain observed behavior, in terms of its optimality in relation to goals, environment, and computational resources.

Todd and Gigerenzer (1999, p. 365), by contrast, sum up the findings of their research program with the statement that "A bit of trust in the abilities of the mind and the rich structure of the environment may help us to see how thought processes that forgo the baggage of the laws of logic and probability can solve real-world adaptive problems quickly and well." But we would suggest that proponents of rational explanation were never committed to the idea that the thought processes themselves are weighed down by the baggage of normative models. And we shall suggest that where a simple heuristic is shown to succeed in a real-world environment, the question of why it succeeds still remains to be answered—and it is hard to see how this question can be addressed without taking up the 'baggage' of normative theory once more.

We shall argue that the mischaracterization of rational explanation in the social and biological sciences as involving rational calculation, rather than just rational description, is of critical importance. Moreover, once the rational description viewpoint is properly understood, the fundamental attacks on the relevance of rational norms in explaining behavior lose their force. We therefore now turn to spelling out a methodology for explanation by rational description, which has been crisply expressed in Anderson's (1990) methodology of "rational analysis."

3. Rational analysis: A methodology for the rational description of behavior and cognition

3.1. Rational and algorithmic explanations of cognition

Human inferential behavior is spectacularly successful. In the face of severe memory and time constraints, the cognitive system vastly outperforms the most sophisticated artificial intelligence systems in almost every real-world domain (see, e.g., McDermott, 1987; Pearl, 1988). Two fundamental questions must be addressed. First, why is inference successful? Following Anderson (1990), we characterize answers to this question as explanations at the rational level. Second, how is this success achieved? Again following Anderson (1990, 1991), we characterize answers to this question as explanations at the algorithmic level (Marr, 1982)—i.e., by specifying the computational procedures involved in inference, such as Take-the-Best and the other algorithms in Gigerenzer and Goldstein's competition.

3.2. Rational analysis

The central idea underlying explanation at the rational level is that, if cognition is well adapted to achieving a goal in some environment, then it can be described as approximating, to some degree, the optimal solution to achieving that goal in that environment. Providing a rational level explanation requires specifying the goals of the system, the structure of the environment, and formally deriving an optimal solution for achieving the goal in that environment. An elegant methodology for constructing such explanations has been formulated in the context of cognitive psychology

[5] There is a telling contrast here with the empirical testing of the mental logic theory of deductive reasoning mentioned above, a theory that does concern rational calculation. In the literature on this theory, there has been considerable emphasis on attempting to show, using error and reaction time data, that the internal processes underlying deductive reasoning involve carrying out steps in a logical calculation (see Rips, 1994).

in Anderson's concept of rational analysis.[6] Anderson's approach involves six steps:

1. Specify precisely the goals of the cognitive system.
2. Develop a formal model of the environment to which the system is adapted.
3. Make minimal assumptions about computational limitations. [Sometimes 'minimal' assumptions will be no assumptions at all—this step is therefore optional.]
4. Derive the optimal behavior function given 1–3 above. [This requires mathematical analysis using rational norms, such as probability theory and expected utility theory.]
5. Examine the empirical evidence to see whether the predictions of the behavior function are confirmed. [Note crucially that the goal is to use the rational analysis to predict and describe behavior. Thus, the rational calculations in Step 4 are not assumed to be carried out by the cognitive system. This is what makes rational analysis a theory of rational description, rather than rational calculation.]
6. Repeat, iteratively refining the theory.

As we will see below, this pattern of explanation is equally appropriate to characterizing classical rational explanation in economics or behavioral ecology.

In the following three sections, we illustrate rational and algorithmic explanation by reference to specific examples from psychology and the social and biological sciences, showing how such explanation appears to undermine each of Gigerenzer and Goldstein's (1996) arguments (A)–(C). For each argument, we outline its basis in Anderson's account of rational analysis (1–6 above), and exemplify these points from psychology and other areas of the social and biological sciences.

3.3. Ecological considerations

Advocates of ecological views of rationality (Evans & Over, 1996, 1997; Gigerenzer & Goldstein, 1996; Gigerenzer & Todd, 1999a) make much of the contrast between everyday human behavior, the success of which must be judged in the context of a specific and complex environment, and abstract classical principles of rationality, which appear to be justified a priori, and hence to embody no constraints concerning the reasoning environment. In short, the concern is that classical principles of rationality are 'unecological,' and hence inappropriate as standards of real-world reasoning.

But, as we have seen, a central element of rational analysis is modeling the environment at an appropriate level of idealization (Anderson's Step 2). The environmental success of inference is explained to the extent that it approximates the optimal behavior function (Anderson's Step 4) derived by applying rational principles to the environmental problem. Consequently, on this view, rational principles and environmental success are complementary, and not in opposition. In psychology, this is familiar from perception, where rational level theory (Marr's computational level) involves detailed modeling of the visual environment. Only then can optimal models for visual processing of that environment be defined. Indeed, Marr (1982) explicitly allies this level of explanation with Gibson's "ecological" approach to perception (Gibson, 1979), where the main focus is on environmental structure.

Similarly, in behavioral ecology, environmental idealizations of resource depletion and replenishment of food stocks, patch distribution and time of day are crucial to determining optimal foraging strategies (Gallistel, 1990; McFarland & Houston, 1981; Stephens & Krebs, 1986). And in economics, idealizations of the "environment" are crucial to determining rational economic behavior (McCloskey, 1985). In microeconomics, modeling the environment (e.g., game-theoretically) involves capturing the relation between each actor and the environment of other actors and exogenous variables (Kreps, 1990). In macroeconomics, explanations using rational expectations theory (Muth, 1961) begin from a formal model of the environment, as a set of equations governing macro-economic variables.

In summary, environmental analysis cannot replace rational norms as an explanation of why behavior is successful. In rational explanation in psychology, behavioral ecology, and economics, both environmental modeling and rational principles are required.

Whereas in some contexts, Gigerenzer and colleagues consider environment analysis as a possible alternative to explanation in terms of rational norms, in other contexts, environmental analysis is viewed as a potential 'add-on' to rational explanation. Gigerenzer and Todd (1999b, p. 11), for example, state that "in Anderson's rational analysis framework (Anderson, 1990; Oaksford & Chater, 1994) constraints from the environment are used to modify one's understanding of what is optimal behavior in a particular context." We would argue instead that, without constraints from agent's goals and environment (and, optionally, computational constraints) the notion of 'optimal behavior' is simply ill-defined—the goals and environment define the optimization problem. In particular, then, it is misleading to think of constraints from the environment modifying one's understanding of optimal behavior, as Gigerenzer

[6] Gigerenzer and Goldstein (1996) argue that Anderson's (1990) framework for rational analysis is an example of classical rationality in psychology. More specifically, Gigerenzer and Todd (1999a) classify Anderson's (1990) rational analysis as rationality subject to constraints. See Oaksford and Chater, 1998b for a collection of recent work in this tradition; and Chater and Oaksford, 2000, relating the methodology to other philosophical positions concerning rationality and behavior.

and Todd suggest. Without a specification of the structure of the environment, optimal behavior is not well defined.

Although Gigerenzer and Goldstein's (1996) and Evans and Over's (1996, 1997) arguments do not undermine rational explanation in psychology, their arguments do have a genuine target. Their emphasis on the environment does count against decontextualized accounts of human inference, which ignore content and context. Part of the reason for this is the attempt to devise empirical tests of descriptive rational theories which are independent of specific beliefs or utilities, which has led to a focus on internal consistency of behavior in highly artificial conditions, rather than on how behavior meshes with the environment. Reasoning that may appear poor in an ecologically invalid laboratory context may be highly adaptive in the natural environment, as has been extensively argued (Gigerenzer, Hell, & Blank, 1988; Gigerenzer & Hoffrage, 1995; Gigerenzer & Murray, 1987; Oaksford & Chater, 1991, 1993, 1998a; Oaksford, Chater, & Stenning, 1990). Thus, it is important to stress the environmental context in which reasoning takes place in order to understand everyday human inference (Oaksford & Chater, 1995).

3.4. Cognitive limitations

If rational explanation in the social and biological sciences is assumed to involve rational calculation, then this style of explanation seems to run into immediate problems of computational complexity. Evans and Over (1997) note that problems of computational complexity bedevil rationally-based theories in the psychology of reasoning; and Gigerenzer and Goldstein (1996) argue, as we have already noted, that classical rational explanation requires the assumption that the mind is a "supercalculator."

But if we adopt the view that we have been advocating, that rational explanation should be understood in terms of 'rational description' rather than 'rational calculation,' then these concerns about computational complexity disappear. To be sure, in rational analysis, deriving the optimal behavior function (Anderson's Step 4) is frequently very complex. Indeed, the relevant rational theories in which these calculations are made, including probability theory, expected utility theory and logic are typically computationally intractable for complex problems (Cherniak, 1986; Garey & Johnson, 1979; Good, 1971; Paris, 1992; Reiner, 1995; Stanovich & West, 2000; Stich, 1990). Intractability results imply that no computer algorithm could perform the relevant calculations given the severe time and memory limitations of a "fast and frugal" cognitive system. Thus it might appear that there is an immediate contradiction between the limitations of the cognitive system and the intractability of rational explanations.

There is no contradiction, however, because the optimal behavior function is an explanatory tool, not part of an agent's cognitive equipment. To extend our earlier analogy, the theory of aerodynamics is a crucial component of explaining why birds can fly. But clearly birds know nothing about aerodynamics, and the computational intractability of aerodynamic calculations does not in any way prevent birds from flying. Similarly, people do not need to calculate their optimal behavior functions in order to behave adaptively. They simply have to use successful algorithms; they do not have to be able to make the calculations that would show that these algorithms are successful.

This viewpoint is standard in rational explanations across a broad range of disciplines. Economists do not assume that people make complex game-theoretic or macroeconomic calculations (Harsanyi & Selten, 1988); zoologists do not assume that animals calculate how to forage optimally (e.g., McFarland & Houston, 1981); and, in psychology, rational analyses of, for example, memory, do not assume that the cognitive system calculates the optimal forgetting function with respect to the costs of retrieval and storage (Anderson & Milson, 1989; Anderson & Schooler, 1991). Such behavior may be built in by evolution or be acquired via a long process of learning—but it need not require on-line computation of the optimal solution.

In some contexts, however, some on-line computations may be required. Specifically, if behavior is highly flexible with respect to environmental variation, then calculation is required to determine the correct behavior, and this calculation may be intractable. Thus the two leading theories of perceptual organization assume that the cognitive system seeks to optimize on-line either the simplicity (e.g. Leeuwenberg & Boselie, 1988) or likelihood (von Helmholtz, 1910/1962; see Pomerantz & Kubovy, 1987) of the organization of the stimulus array. These calculations are recognized to be computationally intractable (see Chater, 1996). This fact does not invalidate these theories, but it does entail that they can only be approximated at the algorithmic level. Within the literature on perceptual organization, there is considerable debate concerning the nature of such approximations, and which perceptual phenomena can be explained in terms of optimization, and which result from the particular approximations that the perceptual system adopts (van der Helm & Leeuwenberg, 1996).

More generally, where on-line computations of 'rational' thought or behavior is required, the usefulness of traditional rational models depends on the (often implicit) assumption that theoretical predictions of rational theories are reasonably stable if the optimization assumption is weakened. For example, in economics, weakenings of rational assumptions have been argued to not just preserve the basic pattern of predictions of economic models, but to enrich them (e.g., by distinguishing viable from non-viable equilibria consistent with rationality assumptions, van Damme, 1991; Harsanyi & Selten, 1988).[7] In behavioral ecology, it has been argued that many phenomena can only be fully understood by showing how the limitations of the animal's cognitive abilities interact with the goal of optimization (see, e.g., Brunner, Kacelnik, & Gibbons, 1992; Kacelnik, 1998; Todd & Kacelnik, 1993). In sum, psychologists, economists, and behavioral ecologists using rational explanation do not claim that the mind has unlimited computational power, and they do not have to.

Rational level explanation, like all scientific explanations of complex phenomena, involves drastic simplifications. Psychologists, economists, and zoologists have pursued optimality approaches in the hope that idealizing away from cognitive limitations may provide an approximate, and insightful, description of aspects of human or animal behavior. The hope is that, by ignoring limitations on rationality, just as physicists sometimes ignore friction, a useful idealization may be possible (Roth, 1996). Whether this hope will prove justified is the real locus of debate concerning optimality models in psychology, economics and behavioral ecology (see, e.g., Arrow, Colombatto, Perlman, & Schmidt, 1996; Simon, 1992, for discussion). Both advocates and detractors of perfect rationality idealizations agree that computational limitations exist, and must form part of any complete explanation of human or animal behavior. That is, they agree that the mind is not a Laplacean demon.[8]

Interestingly, Gigerenzer and Todd (1999b) acknowledge that advocates of classical rational explanation typically accept that cognitive limitations are real: "Proponents of unbounded rationality generally acknowledge that their models assume that humans act as if they were unboundedly rational. On this interpretation, the laws of probability do not describe the process but merely the outcome of reasoning." (Gigerenzer & Todd, p. 9). This is entirely consonant with the present view—and seems to undercut their concern with the cognitive complexity of rational explanation. We presume that Gigerenzer and Todd (1999b) would argue that the 'as if' interpretation of rational explanation (which we would endorse) is, in some way, illegitimate; and showing this would seem to be of substantial importance for their position. They do not, however, discuss this issue further.

The concerns over computational complexity that Gigerenzer and Goldstein (1996), and Evans and Over (1996, 1997), raise do have a genuine target: psychological models of inference where normative theories are interpreted as models of mental calculation, not merely behavioral description. The paradigm example of such models are 'mental logic' theories in the psychology of reasoning, which regard the syntactic proof theory for logic as the basis of the algorithms that implement logical inference in the mind (e.g., Braine, 1978; Fodor & Pylyshyn, 1988; Rips, 1994). However, these algorithms are intractable and therefore cannot apply to complexities of real-world contextualized inference (Chater & Oaksford, 1990; Cherniak, 1986; McDermott, 1987; Oaksford & Chater, 1991). Consequently, considerations of cognitive limitations and computational complexity are primarily relevant at the algorithmic level, ruling out computationally intractable implementations of rational calculi such as logic as models of human inferential processes. But these considerations do not undermine the role of these calculi in rational level description.[9]

3.5. Take-the-Best as an existence proof

The main body of Gigerenzer and Goldstein's attack on the role of classical rational norms in understanding real-world reasoning and decision making is devoted to providing an existence proof "... that cognitive mechanisms capable of successful performance in the real world do not need to satisfy the classical norms of rational inference" (p. 650). They therefore conclude that providing algorithms alone can be regarded as an alternative to the standard approach to rational explanation that we have been arguing for in this article.

We argue instead that Take-the-Best illustrates that successful rational algorithms can be developed, before a rational explanation for why they work has been developed.[10] From this perspective, rather than standing

[7] The concern has been raised, however, that certain weakenings of rationality assumptions can have more drastic consequences on overall system-level predictions (e.g., Akerlof & Yellen, 1985).

[8] Indeed, to the extent that Stanovich and West's (2000) position that normative results are often obtained by a many cognitively able individuals, it could be argued computational limitations cannot be too overwhelming an obstacle to normative, or near-normative, performance. We thank an anonymous reviewer for raising this point.

[9] Oaksford and Chater (1991, 1995) point out than an even more serious problem for logic-based theories of inference is that they do not predict people's common sense reasoning behavior, and therefore fail at the rational level. Oaksford and Chater (1994) therefore propose different rational explanations which they claim provide a better explanation of common sense reasoning and data from laboratory tasks. They do not assume, however, that their rational level account is directly implemented (Oaksford & Chater, 1998a).

[10] The opposite pattern, where rational explanations of behavior have proceeded without considering how they might be approximated by cognitive algorithms, is also common, whether in social behavior (Crawford, Smith, & Krebs, 1987; Messick, 1991), economics (e.g., Muth, 1961; von Neumann & Morgenstern, 1944) or animal behavior (Maynard-Smith & Price, 1973).

as a lone 'existence proof' Take-the-Best stands in an illustrious tradition.

Let us consider three examples of psychological interest: associative learning, connectionist models and statistical tests.

First, the Rescorla–Wagner learning rule for associative learning (Rescorla & Wagner, 1972) was developed without any clear 'rational analysis' of why it leads to successful learning. But such an analysis has been provided, by showing that the algorithm asymptotically approximates the optimal solution in a normative probabilistic account of causal reasoning (Cheng, 1997; Shanks, 1995a, 1995b).

Second, connectionist models typically embody algorithms that are shown to learn some task, without a rational explanation of why that learning is successful. For example, in the study of reading, connectionist models mapping from orthography to phonology,[11] (e.g., Bullinaria, 1994; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) learn to 'read aloud' effectively, although not based on a rational theory of the orthography-phonology mapping, or how it should be learned. But these successes have led to a large research program of providing a rational explanation of connectionist learning (e.g., Chater, 1995; Mackay, 1992; McClelland, 1998; Neal, 1993), as well as theoretical analysis of the orthography-phonology mapping (Brown, 1998).

Third, the history of statistical tests used in psychology shows the widespread use of tests as calculating algorithms, before a rational analysis of the conditions under which they apply has been developed (Gigerenzer & Murray, 1987). In response to this, a range of statistical theories have been developed to provide a rational basis for practical statistical algorithms (e.g., Bernado & Smith, 1995).

In each case, where algorithms have proved practically successful in the absence of an obvious rational basis, this has triggered a search to provide a rational explanation for why the algorithm is successful. We suggest that, analogously, Gigerenzer and Goldstein's impressive demonstration of the success of Take-the-Best should lead to a search for a rational analysis of why it succeeds, rather than the conclusion that rational explanation is dispensable.

Indeed, recent important work by Gigerenzer and colleagues (e.g., Martignon & Hoffrage, 1999; Martignon & Laskey, 1999) suggests that there may be no fundamental dispute on this issue. They provide a rigorous formal analysis of the conditions under which Take-the-Best succeeds. Thus, they provide a descriptive rational explanation for the success of Take-the-Best's behavior, using standard normative principles to do so.

Thus, Take-the-Best stands as an outstanding example of how a 'fast and frugal' algorithm can succeed in the real world, and exemplifies that environmental success does not require that the cognitive system engages in rational calculation using probability or statistical theory. But this does not challenge the use of descriptive rational theories, to explain why algorithms are successful—and, as noted above, Gigerenzer and colleagues have themselves provided such an analysis of the success of Take-the-Best. Thus, once it is recognized that rational explanation in psychology, economics or behavioral ecology involves rational description rather than rational calculation, it is clear that the success of Take-the-Best does not undermine, but is consistent with, rational explanation in these disciplines.

3.6. Counterarguments and replies

We have argued that rational explanation has a central role in the social and biological sciences—it provides descriptions of behavior as approximating rationally optimal behavior, given a specific problem, environment, and (possibly) set of cognitive limitations. We close this section by briefly responding to two possible lines of counterargument: (1) that the iterative aspect of rational analysis (Step 6) betrays a fundamental difference with respect to normative rational theories under test in the psychology and economics of judgment and decision making; (2) that appeal to evolutionary considerations provides an alternative answer to questions of why cognition succeeds, obviating the need for rational explanation.

3.6.1. Iteration and rational norms

It might appear that the iterative step in Anderson's rational analysis (Step 6) betrays a fundamental difference between the program of rational analysis of behavior and norms of classical rationality, as used in economics, psychology, or behavioral ecology.[12] The norms of classical rationality have been derived by a priori analysis, typically from deriving normative theories from seemingly incontrovertible axioms or assumptions. For example, probability theory can, for example, be justified by a range of (converging) arguments, such as the Dutch book theorem, which states (given certain assumptions) that if a person's judgments violate the laws of probability theory, they will accept a bet that they will certainly lose (which seems

[11] We take no stand here on whether such models are compatible with detailed psychological and neuropsychological data. See, e.g., Coltheart, Curtis, Atkins, and Haller (1993) for discussion.

[12] We thank an anonymous reviewer for pointing out this possible concern.

undeniably irrational). If logic, probability theory, expected utility theory, and game theory, are justified by a priori analysis, then it would seem that they cannot be iteratively modified, in order to provide better fits with empirical phenomena (as Anderson's Step 6 implies).

This argument is, however, misleading. The modification involved in Anderson's Step 6 does not typically involve modification to the rational norms themselves (i.e., modification of the calculating machinery used at Anderson's Step 4)—a radical option suggested, for example, by Cohen (1981). Rather, it involves modifications at Steps 1–3—a revised formulation of the goals of the agent, the environment, or cognitive limitations. In explaining foraging behavior, for example, the behavioral ecologist may improve the rational description by noting that the animal aims not merely to maximise food intake, but to minimise variance in food intake (e.g., to avoid the risk of obtaining very little food on some occasions with perhaps life-threatening consequences). The same behavioral ecologist may incorporate a richer model of the environment, incorporating, for example, the length and temperature of the night time period (this will partly determine what lower bound on food intake is acceptable if the animal is to survive the night); and the behavioral ecologist may also note that the animal can only monitor the amount of food it has obtained with Weberian accuracy, which also impacts on its foraging strategy (see Brunner et al., 1992). Parallel examples from economics and psychology have the same form. Crucially, Anderson's Step 6 is not typically concerned with modifying, e.g., the laws of probability. Hence the suspected clash with a priori rational norms does not typically arise.

To sum up, the iterative aspect of rational descriptive modelling need not involve tinkering with apparently incontrovertible norms of reasoning. Instead, it involves iteratively modifying the empirical assumptions, concerning the agent's goals, environment, and cognitive limitations, which are an input to rational calculations.[13]

3.6.2. *Evolution as an alternative "why" explanation*

One line of research that is often perceived as allied with the 'ecological' view of cognition, and hostile to classical rationality, is evolutionary psychology (e.g., Barkow, Cosmides, & Tooby, 1992; Pinker, 1998). Indeed, both Evans and Over (1996) and Gigerenzer and Goldstein draw on evolutionary considerations in developing their accounts of ecological rationality. This raises the critical question of whether evolutionary explanation might provide an alternative explanation for why cognitive mechanisms succeed in their environments—an explanation that can replace explanations that invoke classical rationality.

Perhaps natural selection has ensured that our cognitive algorithms succeed; or perhaps our learning mechanisms have simply favored algorithms that work. But explanations in terms of evolution or learning do not explain why specific cognitive algorithms are adaptive. Instead, they explain why we possess adaptive rather than non-adaptive algorithms—essentially because adaptive algorithms, by definition, perform better in the natural environment, and processes of natural selection or learning will tend to favor algorithms that are successful.

Let us illustrate this point with an example from a domain in which evolutionary explanation is widely accepted. An account of optimal foraging in behavioral ecology may explain why particular foraging strategies are successful and others are not. Behavioral ecologists assume evolution explains why animals possess good foraging strategies, but do not take evolutionary explanation to provide an alternative to the rational level explanation given by optimal foraging theory.

There is, though, a way of sharpening this concern further, in the light of the iterative character of rational analysis, that we dealt with above.[14] This is that the process of iteratively developing a rational analysis might have much in common with the 'iterative' refinement of the algorithms employed by the cognitive system, during biological evolution. Indeed, one might even suppose that the parallel between the 'meta' level of theoretical 'evolution' and the process of biological evolution might be strong—after all, each involve successive modifications in order to provide a better fit between environment, goals and computational resources. This might suggest (although spelling out a detailed argument is not straightforward) that devising an explanation of some behavior in terms of rational analysis will only be possible where some evolutionary process can lead to the corresponding algorithm. And this might further suggest that an explanation in terms of rational analysis will work only when a parallel evolutionary explanation would serve equally well.

We suggest that, nonetheless, an evolutionary explanation is not an alternative explanation of why a

[13] In economics, there is, interestingly, work which does seek to challenge norms of rationality on the basis of empirical data, thus challenging the very premise of the claim that iterative modification does not fit with the a priori character of rational norms. The motivation underlying this challenge is that rational norms are ultimately justified only insofar as they capture human reasoning behavior (e.g., Cohen, 1981). For example, it has been argued that Allais's paradox undermines the normative status of expected utility theory, and that the theory should be revised to fit with our intuitions (see Allais, 1953; Chew, 1983; Fishburn, 1983; Kahneman & Tversky, 1979; Loomes & Sugden, 1982; Machina, 1982).

[14] We thank an anonymous reviewer for raising this point.

cognitive algorithm is successful. Evolutionary explanations are based on the fact that certain algorithms are more successful than others (where success is measured in terms of contribution to the inclusive fitness of the individual with that algorithm). But we still have to address the question of why some algorithms are successful in the environment, whereas some are not. Answering this question requires analysing the structure of the environment, the goals of the agent, and studying how these goals can be achieved given that environment. In short, it involves rational level explanation. This suggests that rational explanation addresses a different question to evolutionary explanation. Rational explanation is required to explain why a particular cognitive algorithm succeeds (given particular goals, environment, and computational limitations). Evolutionary explanation explains how the differential success of different algorithms can lead to the gradual predominance of successful algorithms, through a process of natural selection. These are fundamentally different questions, and hence it is not appropriate to view evolutionary explanation as a potential alternative to rational explanation.

3.7. Summary

We have considered three lines of argument that aim to undercut the role of rational explanation in understanding everyday judgment and decision making. For each argument, we have found points of agreement: We endorse the emphasis on the environment and on cognitive limitations, and on finding simple algorithms that work well in the real environment. But we have argued that these positive points are entirely consistent with "classical rationality," as it is used to explain behavior in the social and biological sciences. Indeed, we argue that rational explanation is always desirable: without it, the adaptive success of cognitive algorithms is left unexplained.

It is worth noting that the consequences of rejecting rational level explanation would be alarming. The very idea that human thought can be understood as reasoning rather than as a collection of uninterpreted procedures involves the assumption that some rational norms are being approximated (see Newell, 1982; Oaksford & Chater, 1995). Giving up the idea that thought involves reasoning has catastrophic implications, not just within psychology, but more broadly: Assumptions of (approximate) human rationality are at the core of "rational choice" explanations in the social sciences and economics (e.g., Elster, 1986) and appear to underpin the attribution of meaning both to mental states and to natural language (Davidson, 1984; Quine, 1960). Fortunately, these alarming consequences need not be faced. Whereas Gigerenzer and Goldstein argue that the cognitive system is fast and frugal, but does not admit of rational explanation; we argue instead that cognition is fast, frugal and can be explained in rational terms.[15]

4. How Plausible is Take-the-Best?

We now turn to the question of the plausibility of Take-the-Best as a model of cognitive estimation. We present a new competition between Take-the-Best and a range of standard algorithms used in cognitive modeling in psychology, and find that Take-the-Best's performance is impressive. We agree with Gigerenzer and Goldstein (1996, 1999) that Take-the-Best is a serious contender as a model of this kind of cognitive estimation.

We argue, though, that there are grounds for caution. Take-the-Best performs well in city size estimation (and in a wide range of other domains, Czerlinski et al., 1999), and at a level comparable with human performance. On the other hand, as we shall see later, the empirical evidence for Take-the-Best is not clear cut.

Gigerenzer and Goldstein's case for the cognitive plausibility of Take-the-Best has three components: (1) Take-the-Best performs well on city size estimation (and other tasks); (2) Take-the-Best is fast; (3) Take-the-Best is frugal (i.e., it uses relatively little information from memory).

We suggest that none of these points provide strong grounds for presuming that Take-the-Best is more plausible than a range of other algorithms. Specifically, we argue: (1) that many standard algorithms perform comparably with Take-the-Best; (2) that these algorithms may be just as fast as Take-the-Best; and (3) that frugality in terms of informational retrieval may not confer any advantage in terms of cognitive plausibility. We begin by reviewing Take-the-Best and Gigerenzer and Goldstein's original competition.

4.1. Gigerenzer and Goldstein's competition

Gigerenzer and Goldstein (1996, 1999) consider a range of algorithms for comparing the populations of pairs of cities, based on a list of features of each city. They show that a very simple algorithm, Take-the-Best, performs as well as "various 'rational' decision proce-

[15] The argument is complicated here by the fact that Gigerenzer and colleagues argue that 'rationality' should be used in an ecological sense, rather than in a normative sense—i.e., a mental process is ecologically rational if it 'just works' even if there may, putatively, be no normative explanation for why it works. Hence, in common with the present view, they embrace the conclusion that fast and frugal algorithms can also be 'rational'—e.g., "models of reasoning need not forsake rationality for psychological plausibility..." (Todd & Gigerenzer, 1999, p. 365). But clearly this conclusion has very different implications from ours, precisely because 'rationality' has been decoupled from normative explanation.

dures" (p. 650). As we noted in the introduction, Take-the-Best discriminates the size of cities by sequentially considering individual features (until the two cities differ on a feature). The features are considered in descending order of the validity of each feature for city size–the first feature in the order on which the cities differ determines the judgment concerning which city has the larger population.

Gigerenzer and Goldstein's five comparison algorithms are linear regression and various approximations to linear multiple regression ("tallying," "weighted tallying," the "unit weighted linear model," the "weighted linear model"). The most important comparison is between Take-the-Best and multiple regression, for two reasons. First, the other algorithms are approximations to multiple regression, and hence would be expected to show comparable (or poorer) levels of performance—if Take-the-Best can match multiple regression, it is likely to match or exceed the performance of the other algorithms. Second, in a vast range of tasks, from clinicians predicting the outcomes of their patients to bankers predicting the viability of companies, linear regression performs as well as, and often better than, human experts (e.g., Einhorn, 1972; Libby, 1976; Meehl, 1954; Sawyer, 1966).

It is crucial, though, to recognise that multiple regression has a very different status from the normative laws of logic, expected utility, or probability. Whereas these laws are candidates for universal rational principles, which can be given a priori justification, this is definitely not the case for multiple regression. Multiple regression requires that different pieces of information combine linearly, and this is known to be appropriate only in a very restricted set of cases. The restricted character of linear regression is often recognized in the psychological literature. Indeed, much research has been concerned with exploring when cues are integrated linearly, and when they are integrated non-linearly (e.g., Hammond & Summers, 1965), and with clarifying the conditions under which linear regression works well in practice, even if its underlying linearity assumptions are violated (e.g., Dawes & Corrigan, 1974). The limitations of linear regression have also motivated the vast statistical literature on various non-linear regression methods, including projection-pursuit regression (Intrator, 1993), regression using multilayer connectionist networks (Neal, 1993)[16] and exemplar-based regression methods (Duda & Hart, 1973). Multiple regression is a pragmatically useful tool, which works well under restricted conditions. The results of multiple regression cannot, therefore, be viewed as embodying principles of rationality defining "correct" reasoning, despite this impression sometimes being given in research in the literature.[17]

Gigerenzer and Goldstein present computer simulations that compare these algorithms. Take-the-Best and the other methods learn on a subset of cities (which are therefore "known"), and it is "tested" on the entire set of cities. Specifically, Take-the-Best learns cue validities from the initial subset, from which the ordering and significance of cues is derived. The generalization performance of Take-the-Best and three of the alternative algorithms, linear regression, tallying and weighted tallying are almost identical. These simulations provide an interesting and useful set of comparisons in a psychologically important yet tractable domain. Gigerenzer and Goldstein argue that the good performance of Take-the-Best in their competition, and its speed, is evidence for its cognitive plausibility. But we will argue that the more general and psychologically familiar algorithms that we now consider are at least as cognitively plausible as models of city size estimation.[18]

4.2. A new competition

The algorithms we consider are drawn from the range of standard methods used in cognitive psychology and artificial intelligence research, rather than originating in statistics. We suggest that these kinds of models may usefully be viewed as 'null hypotheses' against which more specialized algorithms (such as Take-the-Best) can be compared.

The first algorithms are exemplar-based, and assume that people store previous examples, and judge new examples in relation to their similarity to stored examples. Exemplar-based methods are general with respect to domain: They have been widely used in psychological models of categorization (e.g., Nosofsky, 1986) and memory (e.g., Hintzman, 1986), and are related to non-parametric statistical methods which have been extensively applied in statistics, pattern recognition and artificial intelligence concerned with both classification

[16] For those familiar with connectionism, it may be useful to note that multiple regression is mathematically closely related to the single layer perceptron; connectionist research was largely abandoned in the late 1960s partly because this device could learn such as limited range of problems (Minsky & Papert, 1969; Rumelhart & McClelland, 1986).

[17] We thank an anonymous reviewer for raising this issue. Gigerenzer and Goldstein (1996) signal this distinction by using explicitly putting "rationality" in inverted commas when referring to the linear regression and its variants, which integrate information across all features, but using the term without inverted commas when referring to putatively universal rational norms such as probability theory or logic.

[18] Although these algorithms are general, and psychologically familiar, these algorithms have no transparent relationship to rational analysis—although the project of explaining the performance of neural network and exemplar models in 'rational' terms is now quite well developed, as touched on briefly above (e.g., Chater, 1995; McClelland, 1998; Mackay, 1992; Neal, 1993; Nosofsky, 1990).

and interpolation problems (Duda & Hart, 1973; Parzen, 1962; see Ashby & Alfonso-Reese, 1995, for discussion). They are also extremely general with respect to the structure of the data they can handle (Cover & Hart, 1967).[19] Interestingly, moreover, an exemplar-based model of memory and generalization (Hintzman's (1984, 1988), MINERVA2 model) has recently been used as the basis for an influential recent model of the processes underlying probability judgments (Dougherty et al., 1999), as well as the PROBEX model of probabilistic inference (Juslin & Persson, 2001).

The second type of algorithm is the multilayered, feedforward neural network, trained by back-propagation (Rumelhart, Hinton, & Williams, 1986). Like exemplar models, these models are very general with respect to domain. This is shown by the vast amount of psychological and applied research using these methods across a range of problem types (Christiansen, Chater, & Seidenberg, 1999; Rumelhart & McClelland, 1986). They are also very general with respect to the structure of the data that they can handle, being able to deal, for example, with highly non-linear mappings, and mappings that mix rules and exceptions (e.g., Seidenberg & McClelland, 1989), although the limits of this generality are not clear.

The third type of algorithm is a standard classification learning algorithm from machine learning research, the C4.5 decision tree model, which has been used across a wide range of problem domains. It has also been used to a limited degree in psychological modelling (e.g., Ling & Marinov, 1993, 1994).[20]

4.2.1. Representation of data

Gigerenzer and Goldstein represent each city as a vector of nine binary (0 or 1) cue values. To facilitate comparison between algorithms, we represented each pair of cities by nine features representing the difference between the nine cue values for each city. For example, for the cities with features (1, 1, 0, 0, 1, 0, 0, 1, 1) and (1, 0, 1, 1, 1, 0, 1, 1, 0), the corresponding difference pattern would be (0, 1, −1, −1, 0, 0, −1, 0, 1). Each pattern is associated with a label indicating whether the population of the first city was smaller than, equal to, or larger than, the population of the second city. This change of representation has no effect on Take-the-Best. Taking all pairs of distinct cities in both orders yielded a possible $83 \times 82 = 6806$ training patterns.

In order to capture the effects of limited knowledge, we trained each of the algorithms on a subset of the 6806 comparisons. In Fig. 1, the percentage of training examples refers to the percentage of these comparisons presented during the training of each algorithm. The values shown in Fig. 1 are for generalization performance, for predicting the outcome of all 6806 comparisons. This approach allowed the algorithms to be assessed on an equal footing.

4.2.2. Take-the-Best

As described by Gigerenzer and Goldstein, the cue validity for each feature was calculated as the fraction of training examples in which the feature correctly picked out the larger city divided by the total number of training pairs on which that feature differed between cities. The cue validities determined the order in which features are considered by Take-the-Best, and the first feature in this order that discriminated between the two cities was taken to be the model's answer.

4.2.3. Exemplar-based models

We used two exemplar-based models: Nearest Neighbor (e.g., Cover & Hart, 1967) and the Generalized Context Model (Nosofsky, 1990). In both models, exemplars are the difference patterns, representing the difference between the features for pairs of cities. As noted above, each difference pattern is associated with a label indicating whether the first or second city in the pair is larger.

In the nearest neighbor algorithm, the difference pattern associated with each test pair is constructed. Thus, each pattern corresponds to a point inside in a nine dimensional hypercube, with values on each dimension taking values −1, 0, or 1. The nearest difference pattern associated with a training pair of cities is then selected. If this difference pattern is associated with the first city of the training pair being larger, then the algorithm responds that the first city of the test pair is larger; and similarly if the second city of the training pair is larger, then the algorithm responds that the second city of the test pair is larger. Distance between difference patterns is measured by the Euclidean distance metric (i.e., the square root of the sum of the squared distance along each dimension).

[19] Of course, the cognitive plausibility of these algorithms can also be challenged, e.g., on the grounds that the requirement that past cases are retrieved imposes an unreasonable memory load on the cognitive system. Advocates of exemplar models might respond by suggesting that retrieval of quite a small subset of items from memory would produce very similar results. But, from the present point of view, we note simply that this class of algorithms is widely used in cognitive modeling in categorization and memory, and hence is presumably viewed, at least by its advocates, as cognitively plausible. The same point applies to the other classes of algorithms in our competition.

[20] One difference between Gigerenzer and Goldstein's simulations and those reported here is that we have assumed that people only have to compare familiar cities. Gigerenzer and Goldstein considered cases in which some cities might not be recognized, and used their "recognition principle" to deal with this case. The recognition principle is that cities which are recognized are assumed to be larger than cities which are not recognized. Gigerenzer and Goldstein allow all the algorithms in their competition to use this principle. In these simulations, we consider only the case where all cities and all features of each city are known, and hence the recognition principle is not relevant (because all cities are recognized).

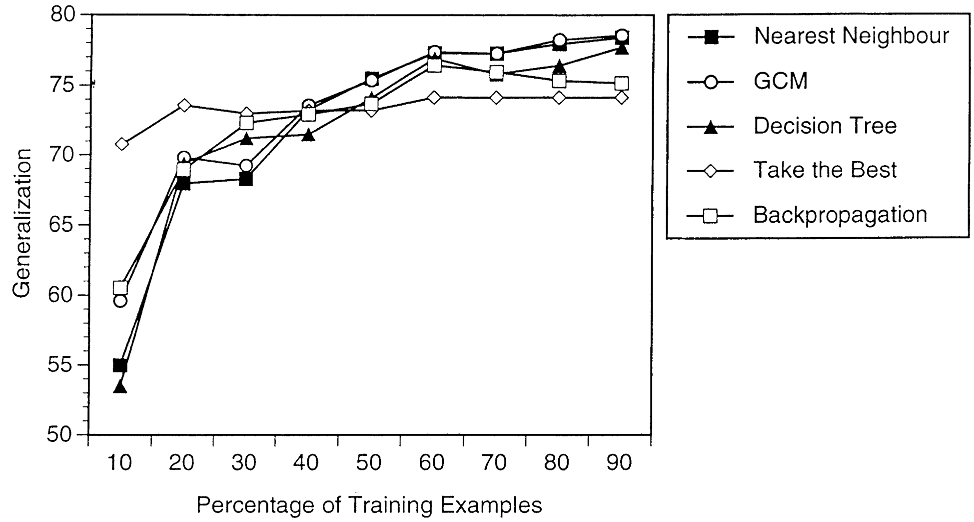

Fig. 1. Results of the competition. Percentage of correct inferences about the populations of German cities as a function of percentage of cities seen in training.

The Generalized Context Model (Nosofsky, 1990) is similar to Nearest Neighbor, but the response is determined by a weighted sum of all training pairs, rather than just the nearest training pair. The influence of a training pair is determined by its similarity to the test example. Specifically, the influence of each training pair is a Gaussian function of its distance from the test example, again using a Euclidean distance metric. In practice, in this model, which city is judged to be the larger depends on the influence of possibly several nearby difference patterns, rather than just the nearest difference pattern, as in the Nearest Neighbor algorithm. The Generalized Context Model has adjustable parameters concerning the relative weighting of each feature, and also allows a parameter for response bias. For simplicity we did not include such parameters, and thus each feature was weighted equally, and there was no response bias. We included just one adjustable parameter, the standard deviation of the Gaussian. The standard deviation of the Gaussian determines the 'narrowness' of the search among the differences patterns. If the standard deviation is small, then typically only the nearest item will have any substantial influence on the judgment, and hence the results of the algorithm become identical with the Nearest Neighbor algorithm. If the standard deviation is large, then many difference patterns have some influence (although that influence still diminishes with increasing distance from the difference pattern representing the pair of cities about whom the judgment is being made). This standard deviation was optimized straightforwardly by measuring the generalization score for many different values and choosing the best.

It might be thought that, by not allowing free choice for the other parameters in the Generalized Context Model, we unreasonably disadvantaged the Generalized Context Model. Certainly, if these additional parameters are adjusted freely, post hoc, in order to obtain the highest level of accuracy, a modest improvement in performance is obtained. On the other hand, however, this improvement is wiped out if we 'train' the parameters on only some of the city comparisons, and then generalize to the remaining comparisons. That is, it appears that the slight advantage of using these parameters is due to 'overfitting.' Hence, in the simulations shown, we decided to avoid such complexities and set the parameter values to be equal (the specific fixed value chosen for all of these parameters is arbitrary).

It turns out, moreover, that performance is remarkably insensitive to the specific parameter values that are chosen. This is appears to be an analog of the 'flat maximum' phenomenon that is found for linear regression for similar problems: that quite large variations of regression weights in a linear model lead to remarkably similar levels of performance.

4.2.4. Feedforward connectionist network

We used a three-layer feedforward network with nine input units, two hidden units, and one output unit, trained using the backpropagation algorithm (Rumelhart & McClelland, 1986). The inputs were the difference patterns, and the output corresponded to the decision about which city is larger. The target values for the output were 0, .5, and 1, for smaller, equal to, and bigger, respectively. Weights were initialized to random values within the range (−.5, .5). The net was trained for 100 epochs (passes through each training sample), with a learning rate of .01, and a momentum of .9 (these parameters were not adjusted to obtain good performance—the parameter values used in the simulations reported here were the first values that we used). The order of the training examples was randomized within each epoch. During test, output values less than .5 were classified as "smaller," and values greater than .5 were classified as larger. The results that we obtained appear to be insensitive to the specific choices of parameters (e.g., numbers of hidden units, learning rate, momentum, etc.).

4.2.5. Decision trees: C4.5

The decision tree algorithm, C4.5 (Quinlan, 1993), was used to construct a decision tree on the basis of the nine features of the difference patterns. At the top level of a decision tree the feature that best distinguishes smaller from larger cities is used to divide the difference patterns into two groups. One group putatively contains the pairs of cities for which the first city is the larger, the other group putatively contains the pairs of cities for which the second city is the larger. Classification on the basis of this single difference feature will be wrong for some city pairs. Therefore at the next level in the tree, another feature can be used to subdivide these cases. The leaves of the decision tree are associated with a "decision"—that the first or second city is the larger. The number of levels and mode of construction of the decision tree is determined by an information-theoretic measure. See Quinlan (1993) for a detailed description and source code for the C4.5 algorithm.

Like Take-the-Best, decision trees use a small number of features, rather than integrating all the features. However, whereas Take-the-Best relies on a single cue, decision trees may make reference to several cues in navigating through the tree to reach a decision. The precise way in which this is done is somewhat complex (Quinlan, 1993); nonetheless, the approach is of interest, partly because it has been successfully applied in other psychological contexts (Ling & Marinov, 1993, 1994).

5. Results and discussion

The performance of Take-the-Best is again impressive: it outperforms the other algorithms where limited data is available, and performs almost as well, when most or all relevant data is available.[21]

Notice that, that the overall levels of performance across many simulations algorithms are very similar. Combined with Gigerenzer and Goldstein's observation that Take-the-Best had almost exactly the same performance profile as tallying, weighted tallying and multiple regression, this suggests that this population estimation task is a poor discriminator between algorithms.[22] The familiar cognitive algorithms used in this competition are widely used to model data from other domains, and match Take-the-Best's performance in Gigerenzer and Goldstein's cognitive estimation task. It would therefore seem that these widely used cognitive models have a comparable level of prima facie plausibility as Take-the-Best as potential cognitive models in this city size estimation task. There results are also broadly in line with recent results from Gigerenzer's laboratory, in which Take-the-Best is found to perform nearly as well as two Bayesian statistical methods: so-called naïve Bayes, which assumes (typically incorrectly) that all cues are conditionally independent given the outcomes values (i.e., which city is the larger); and, by contrast, a cutting-edge 'Bayesian network' learning algorithm (specifically, an algorithm developed by Cooper & Herskovits, 1992; Friedman & Goldszmit, 1996; see Frey, 1998; Pearl, 1988). Successful performance by Take-the-Best was, moreover, found to hold across the wide range of data sets used by Czerlinski et al. (1999).

Gigerenzer and Goldstein, however, argue that Take-the-Best is particularly attractive because it is fast (it involves a small number of serial processing steps) and frugal (it draws on very limited information, because it is non-integrative). We now argue that neither consideration straightforwardly gives Take-the-Best an advantage over the available alternatives.

5.1. Is Take-the-Best especially fast?

Gigerenzer and Goldstein argue that Take-the-Best is faster than the other algorithms in their competition, in terms of the amount of information searched in memory. The possibility therefore arises that Take-the-Best may be preferable to the general purpose algorithms we have considered on grounds of speed. There are two points to consider.

First, very rapid integration of vast amount of information is believed to occur in language processing, perception, motor control, and commonsense reasoning, as we shall discuss in the next subsection. Therefore there seems no reason to suppose that integrative processing cannot be fast enough to account for presumably relatively slow human responses in cognitive estimation tasks such as city size estimation. Without empirical evidence concerning human performance on the cognitive estimation task, and in particular without information about how rapidly people might perform it, the emphasis on speed as a deciding factor between algorithms may be inappropriate.

Second, Gigerenzer and Goldstein's measure of speed, which favors Take-the-Best, depends on specific assumptions about the architecture of the cognitive system (Chater & Oaksford, 1990; Oaksford & Chater, 1991, 1993, 1995). On a serial architecture, in which it may be presumed that information is searched in

[21] We thank Gerd Gigerenzer for stressing the importance of Take-the-Best's performance with limited data.

[22] Persson and Juslin (1999) provide important additional simulations on this task, showing very similar performance levels for PROBEX, an exemplar-based algorithm, ridge regression, and Take-the-Best. Both PROBEX and ridge regression outperform Take-the-Best for small amounts of data in these simulations.

memory at a constant rate, Take-the-Best would be more rapid than, for example, multiple regression, or the neural network model and exemplar accounts we have considered here. But in a parallel architecture, speed of processing will not generally be related to the amount of information searched in memory, because large amounts of information can be searched in memory simultaneously. So, for example, both the learning and application of multiple regression can be implemented in parallel using a connectionist network with a single layer of connections. This implementation could operate very rapidly—in the time it takes to propagate activity across one layer of connections (e.g., Hinton, 1989). Similarly the back-propagation account could also be rapidly implemented in parallel, in connectionist hardware. In the same way, in an instance-based architecture, where instances can be retrieved in parallel, the Nearest Neighbor and General Context Model algorithms would be the quickest.

Taken literally, the sequential character of Take-the-Best might even make it a rather slow algorithm, if implemented in these architectures, although fast parallel implementations of Take-the-Best may be possible (e.g., using exponentially distributed weights on a single layer neural network, so that each weight dominates the sum of all smaller weights—see Martignon & Hoffrage, 1999). In any case, Take-the-Best only appears to have a clear advantage over other algorithms if we assume that cognitive processes are serial. Given that there are extensive research programs aimed at establishing the viability of instance-based and connectionist architectures as general accounts of cognitive architecture (e.g., Kolodner, 1993; Rumelhart & McClelland, 1986), it seems that considerable caution must be used in applying a measure of speed which presupposes a serial architecture.

5.2. *Is frugality an advantage?*

Take-the-Best is undoubtedly a very frugal algorithm. Rather than integrating all the information that it is given (all the features of the cities), it draws on only enough feature values to 'break the tie' between the two cities. For example, across Gigerenzer and Goldstein's (1996) simulations, only about 1/3 of features were retrieved. But does the frugality of Take-the-Best make it more cognitively plausible? Comparison with other domains suggests that it may not.

In other cognitive domains, there is a considerable evidence for the integration of multiple sources of information. For example, in speech perception, there is a wealth of experimental work showing rapid integration of different cues, including cues from different modalities (e.g., Massaro, 1987). This integration even appears to obey law-like regularities (e.g., Morton, 1969), which follow from a Bayesian approach to cue integration (Movellan & Chadderdon, 1996), and can be modelled by neural network learning models (Movellan & McClelland, 1995). Recent work on sentence processing has also shown evidence for the rapid integration of multiple "soft" constraints of many different kinds (MacDonald, Pearlmutter, & Seidenberg, 1994; Taraban & McClelland, 1988). Motor coordination is a very different domain in which a vast number of constraints must be rapidly and simultaneously respected in order to plan a successful action (Jeannerod, 1988). Finally, Brunswick (1934) provided a wide range of examples where different sources of information appear to be integrated in judging, for example, the weight or value of a collection of coins (see Gigerenzer & Murray, 1987: p. 66 for discussion).

Two points from these examples are relevant to the cognitive plausibility of Take-the-Best. First, the ability to integrate large amount of information may be cognitively quite natural—and hence it is at least not to be taken for granted that the non-frugality of connectionist or exemplar-based models should count against their cognitive plausibility. Second, the processes above appear to require rich and rapid information integration, which cannot be handled by a non-integrative type of algorithm such as Take-the-Best. Thus a non-integrative algorithm such as Take-the-Best may be at a prima facie disadvantage with respect to the generality of its cognitive performance.

A possible objection to this viewpoint may be that evidence for rapid integration of large amounts of information in perceptual and motor domains does not necessarily carry over to the kind of reasoning involved in a judgment task, such as deciding which of two cities is the larger.[23] Perhaps, in such a task, retrieval from memory is slow and sequential, and hence rapid information integration cannot occur. This is an important possibility, which would need to be supported by detailed empirical work on how people make city size estimates, and related judgments. Moreover, there is evidence from related choice tasks that people may focus on a restricted amount of information, rather than attempting to integrate many pieces of information.

On the other hand, though, note that in most areas of everyday reasoning (e.g., reasoning about each other's behavior, reasoning about the physical world) it seems that very large amounts of knowledge are rapidly recruited (e.g., Oaksford & Chater, 1991, 1998a)—indeed, the amount of knowledge recruited appears to be indefinitely large, as the 'frame problem' in artificial intelligence and cognitive science appears to show (Pylyshyn, 1987). It seems reasonable to view estimating, say, the approximate size of a city as a specific example of everyday inference, and hence to assume that large

[23] We thank Gerd Gigerenzer for pointing out this objection.

amounts of information from memory can be retrieved and applied rapidly in this case also. If so, there may be no strong reason to favor cognitive algorithms that are 'frugal.'

Building on the ideas of Payne et al. (1993) and Gigerenzer and Todd (1999a), one might postulate that whether or not frugality is a relevant constraint for a cognitive algorithm depends on the nature of the task. Perhaps, for example, the cognitive system is able to bring to bear very large amount of information in parallel when that information is already known (and perhaps even heavily overlearned); but perhaps there are strong limitations on the deployment of such information when information must be 'loaded in' to memory in the experimental task. This kind of division would explain why, on the one hand, it appears possible to deploy and integrate vast amounts of information in perception, motor control, and common sense inference, whereas the amount of information that can be integrated and deployed in some explicit judgment tasks is severely limited (e.g., if an experimental participant is given a list of properties or cues of an object, loading these into memory, let alone integrating them together, may be slow and difficult). This may reconcile the apparent severe limits of the cognitive system (e.g., Payne et al., 1993; Shepard, 1967) with its apparently impressive information integration performance (e.g., Anderson, 1981). This viewpoint is consistent with the idea that "Higher order cognitive mechanisms can often be modeled by simpler algorithms than can lower order mechanisms" (Gigerenzer & Todd, 1999a, p. 31), with the proviso that this applies only to higher order mechanisms for which there is severe bottleneck of information uptake.

If this is right, this opens up the intriguing possibility that apparently minor variations of the judgment tasks to which Take-the-Best has been applied (Czerlinski et al., 1999; Gigerenzer & Goldstein, 1996) may have different psychological characteristics. Where the background 'cues' that are being drawn upon are part of the prior general knowledge of the experimental participant, one might not expect limitations of information integration to be severe. On the other hand, where the cues are not part of the participant's background knowledge, and are explicitly provided to the participant in the experiment, one might expect information uptake would be limited, and hence that the cognitive algorithm would be limited. Gigerenzer and Goldstein (1996) are not explicit about which scenario they would take to be most appropriate for testing Take-the-Best, or whether it should apply in both cases. As we shall discuss in the next subsection, subsequent experimental work has focused on cases where cues are not part of the background knowledge of the participant in experimental work testing the approach (perhaps the most favorable conditions for Take-the-Best). This is essentially because it allows the experimenter to control the cues that participants can use, and the validity of these cues.

At minimum, then, it seems that the conditions under which frugality is a crucial constraint on judgment are not straightforward, and hence that there is no automatic advantage of a simple non-integrative and 'frugal' algorithm such as Take-the-Best over alternative cognitive algorithms (such as we considered in our new competition above) that do allow information integration.

Overall, we may conclude that issues both of speed and frugality are difficult to assess outside the context of specific assumptions about cognitive architecture. For a serial processor, with very limited memory capacity, searching one item at a time may optimise speed, and be appropriately frugal. But for a connectionist network, parallel information integration may be very rapid; and integration all available information may be computationally just as easy, or even easier, than 'frugally' choosing only the most important information.

5.2.1. Empirical evidence for Take-the-Best

We have considered computational evidence that Take-the-Best performs well in relation to a range of cognitive algorithms, in judgments in the city size estimation task. We now turn to consider empirical evidence. This evidence can be both direct and indirect. Indirect evidence concerns the degree to which there is independent empirical motivation for the underlying principles upon which Take-the-Best is founded, in relation to rival cognitive models. Here, the picture is unclear.

On the one hand, the tradition of empirically successful models of choice using non-integrative methods (e.g., Payne et al., 1993; Tversky, 1969, 1972) and the fact that, under some circumstances at least, people appear to be able to focus only on a limited amount of information (e.g., Shepard, 1967), provides some independent motivation for the underlying principles in Take-the-Best and related models. Moreover, as a rule-based system, Take-the-Best can draw on a long tradition of rule-based proposals concerning cognitive architectures (e.g., Anderson, 1983; Newell, 1991—although the specifics of these architectures are quite distantly related to Take-the-Best).

On the other hand, as we have discussed the other models in our competition also have been used across other cognitive domains, and are here applied with little modification (whereas Take-the-Best is specifically constructed in order to perform this kind of judgment task). Most notably, the now vast tradition of detailed cognitive psychological modeling using connectionist methods suggests that connectionist principles arguably have a strong empirical basis (e.g., Christiansen & Chater, 2001; MacLeod, Plunkett, & Rolls, 1998; Rumelhart & McClelland, 1986).

In the absence of clear evidence concerning the empirical plausibility of the principles underlying Take-the-Best vs. rival algorithms, we turn to the direct empirical evidence.[24]

Working in Gigerenzer's laboratory, Rieskamp and Hoffrage (1999) attempted to gather relevant empirical evidence, in the context of a task in which people had to judge the most profitable of four companies, on the basis of six cues. Unlike the city size case, described above, the cues here are real valued, rather than binary: share price, recognition rate (i.e., what proportion of people have heard of the company), dividend, share capital, investment and number of employees. The sample in the experiments was taken from information about 70 German companies. Cue validities (which are here defined slightly differently from the cue validities in the city size case) were given to the participants (and were visible throughout the experiment). The experimental task was to sample information from a matrix of companies and their properties, by clicking on a cell in the matrix with a mouse—only one piece of information was visible at a time. Rieskamp and Hoffrage found that people sampled information quite frugally (i.e., they did not exhaustively sample the entire matrix, even when they had time to do so). Moreover, they tended to sample information 'cue-wise'—i.e., they chose a particular cue, and then sampled the values of each company on that cue. Rieskamp and Hoffrage note, though, that their data does not provide rich enough evidence to back up Take-the-Best, at least when considered on its own. Perhaps most fundamentally, integrative algorithms appear to be severely disadvantaged from the outset, because the information can only be sampled sequentially in this experimental paradigm. Moreover, cue-wise sampling seems to be appropriate on almost any algorithm, assuming that participants have poor base rate estimates for the various cues, so that they have no way of interpreting the likely significance of, say, the absolute value of the share capitalization of one company, without assessing this in relation to other companies. More generally, the complexity of the relationship between strategies for information integration, and optimal data selection methods (Berger, 1995; Lindley, 1956; Oaksford & Chater, 1994) geared to those strategies is sufficiently great, that it seems unlikely that this source of data will be decisive. Nonetheless, Rieskamp and Hoffrage's results certainly appear compatible with Take-the-Best, as well as a broad range of other algorithms.

A more direct test of Take-the-Best was performed by Bröder (2000). His experiment involved three phases: a period of training, during which participants learned cue validities from experience; the core phase of experimental judgments; and a final phase, in which participants were explicitly asked about cue validities, to check that they had learned them successfully. Bröder used carefully crafted sets of cues. The critical variable in the experiment was the whether 'dominated' cues (i.e., cues that should never be assessed, if cues are sampled in order of validity) agreed or disagree with the 'dominant' cue. According to Take-the-Best the values of these cues should be irrelevant, because people will not sample them. Nonetheless, Bröder found substantial effects of these cues, aggregated across his participant population. Bröder also conducted careful statistical analysis of individual participants behavior, and argued that a substantial number of participants (28% in one study; 53% in another) do roughly conform to Take-the-Best (or, more generally, some kind of non-compensatory decision strategy—these might include decision trees, as used in the competition above). These findings are consistent with the idea that non-compensatory strategies, such as Take-the-Best may be among the heuristics available to the adaptive decision maker (Payne et al., 1993) or present in the adaptive toolbox (Gigerenzer & Selten, 2001).

Similar conclusions were reached in a study of a close variant of Take-the-Best (the Matching Heuristic) in the context of judgments concerning whether defendants should, or should not, be given bail, where participants were lay magistrates in the English legal system (Dhami & Ayton, 2001). The Matching Heuristic was found to provide the best fit to data from a substantial minority of magistrates, when compared to two integrative algorithms. Again, this suggests that non-compensatory algorithms may be among the cognitive strategies that people can use, although it also suggests that people are able to integrate information.[25] In a complementary study, in a medical context, the Matching Heuristic was found to do as well as logistic regression in describing English doctors decisions about prescriptions (Dhami & Harries, 2002).

[24] Gigerenzer and Goldstein (1996) and Goldstein and Gigerenzer (1999) note that there is empirical evidence for an aspect of the Take-the-Best strategy, the recognition principle. Although important, this evidence for the recognition principle does not bear on their competition, because they allow that the recognition principle is combined with all the algorithms in their competition. Similarly, the recognition principle is not relevant to present competition, because we consider the case where all cities (and their features) are known, so that the recognition principle is not triggered. So empirical evidence concerning the recognition principle does not provide empirical evidence that can help decide between rival 'competitors,' with similar levels of 'ecological' success.

[25] Note, though, that this latter conclusion is not straightforward to infer once one allows the possibility of noise in the process of ordering the cues to be assessed. This would also lead to the apparently non-compensatory patterns in Bröder's study, because 'dominated' cues would be sampled on some occasions, and hence could influence participants' decisions.

Finally, Newell and Shanks (2001) built on Bröder's (2000) studies. Specifically, participants had to buy cue values that they could use to help them decide which of two shares would have the best payout (the experimental set-up and cover story was borrowed from Bröder's experiments 3 and 4). According to Take-the-Best, one might expect that people would buy just enough information to allow them to discriminate the two shares (i.e., to break the tie between the two), and then would buy no further information. However, people typically bought a great deal more information than they needed, although some participants went to the opposite extreme of undersampling information, and hence 'guessing' on many trials. Although the conditions used by Newell and Shanks appeared well suited to engaging the Take-the-Best heuristic (e.g., information could only be sampled sequentially) there appeared to be large individual variation in performance, with few participants conforming at all closely to the predictions of Take-the-Best.

Overall, the state of the empirical evidence is mixed. There seems some tentative reason to believe that Take-the-Best, or some similar non-compensatory heuristic, may be among the strategies that people can adopt (Bröder, 2000; Dhami & Ayton, 2001; Rieskamp & Hoffrage, 1999); but there seems to be large individual variation, and, in at least one study (Newell & Shanks, 2001), a poor fit between detailed patterns of behavior with the predictions of Take-the-Best. Note, in particular, that the other algorithms that we considered in the competition above (e.g., connectionist, exemplar-based and decision-tree models) have not been explicitly compared with empirical data. Given the closely related performance of all these algorithms (together with traditional statistical algorithms related to linear regression), discriminating between these empirically represents an important challenge for future work.

5.2.2. Summary

In this section, we have argued that familiar and widely applicable cognitive algorithms give comparable results to Take-the-Best on the city size estimation task, and are at least as plausible on grounds on speed and 'frugality.' Moreover, we suggest that there is not presently sufficient empirical evidence to tip the balance in favor of Take-the-Best—and indeed, much of the empirical evidence appears to face Take-the-Best algorithm with some challenges. Overall, then, the algorithms in our competition, which are well established in psychology and artificial intelligence, appear at least as cognitively plausible as Take-the-Best. Moreover, the current empirical evidence does little to help resolve these issues. So far, it appears that Take-the-Best may capture the data for a substantial minority of participants—but how alternative models might fair in relation to the same empirical data is by no means clear. At present, it seems reasonable to conclude that Take-the-Best is not a universal cognitive algorithm for judgment tasks; but the data is consistent with it being an element of the 'adaptive toolbox' that participants have available to them (Gigerenzer, 2001).

A more general implication of this discussion is that there may be scope for a broader interchange between research on judgment and decision making, and general cognitive science research. It has been persuasively argued that rich experimental techniques may be usefully imported into judgment and decision making research from cognitive psychology (e.g., Dougherty et al., 1999; Payne et al., 1993; Weber et al., 1995); we suggest that it may also be useful to import the general purpose cognitive architectures and algorithms that have been developed within cognitive science and artificial intelligence.

6. Conclusions

In this paper, we have argued for two claims. First, we have argued that standard notions of rational explanation in psychology and the biological and social sciences are not undermined by recent challenges. This pattern of explanation involves providing rational descriptions which explain why behavior is successful, rather than viewing the mind as a probabilistic or statistical calculating machine. The standard notion of rational explanation: (A) stresses rather than ignores the environment; (B) takes cognitive limitations into account; and (C) allows that algorithmic theories may run ahead of rational explanation, as Take-the-Best illustrates. Second, we have argued that familiar cognitive algorithms equal the performance of Take-the-Best on the city size estimation task, and may be equally plausible in terms of speed, memory requirements, and, currently, empirical evidence. We also emphasise the importance of the project of providing rational descriptive explanations that can explain when and why the cognitive system is adaptively successful.

Acknowledgments

We would like to thank Gerd Gigerenzer, Alex Kacelnik, Michael Bacharach and three anonymous reviewers for helping to substantially strengthen this paper. Nick Chater was partially supported by a small grant from the British Academy, by a Senior Research Fellowship from BT, by European Commission Grant RTN-HPRN-CT-1999-00065, by the Human Frontiers Science Program, the ESRC, the Leverhulme Trust, and by Oliver, Wyman & Company.

References

Akerlof, G., & Yellen, J. (1985). Can small deviations from rationality make significant differences to economic equilibria? *American Economic Review, 75*, 708–720.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine. *Econometrica, 21*, 503–546.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and Brain Sciences, 14*, 471–517.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review, 96*, 703–719.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 1*, 396–408.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Anscombe, F., & Aumann, R. (1963). A definition of subjective probability. *Annals of Mathematical Statistics, 34*, 199–205.

Arrow, K. J., Colombatto, E., Perlman, M., & Schmidt, C. (Eds.). (1996). *The rational foundations of economic behavior*. London: Macmillan.

Ashby, F. G., & Alfonso-Reese, L. A. (1995). Categorization as probability density estimation. *Journal of Mathematical Psychology, 39*, 216–233.

Ayton, P. (2000). Do the birds and bees need cognitive reform? *Behavioral and Brian Sciences, 23*, 666–667.

Ayton, P., & Hardman, D. (1997). Are two rationalities better than one? *Current Psychology of Cognition, 16*, 39–51.

Barkow, J., Cosmides, L., & Tooby, J. (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

Berger, J. O. (1995). *Statistical decision theory and Bayesian analysis*. New York: Springer-Verlag.

Bernado, J. M., & Smith, A. F. M. (1995). *Bayesian theory*. Chichester, Sussex: Wiley.

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review, 85*, 1–21.

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.

Brown, G. D. A. (1998). The endpoint of reading instruction: The ROAR model. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 121–138). Mahwah, NJ: Erlbaum.

Brunner, D., Kacelnik, A., & Gibbons, J. (1992). Optimal foraging and timing processes in the starling, *Sturnus vulgaris*: Effect of inter-capture interval. *Animal Behavior, 44*, 597–613.

Brunswick, E. (1934). *Wahrnehmung und Gegenstandswelt: Grundlegung einer Psychologie vom Gegenstand her*. Leipzig: Deuticke.

Bullinaria, J. A. (1994). Internal representations of a connectionist model of reading aloud. In *Proceedings of the sixteenth annual conference of the cognitive science society* (pp. 84–89). Hillsdale, NUJ: Erlbaum.

Chater, N. (1995). Neural networks: The new statistical models of mind. In J. P. Levy, D. Bairaktaris, J. A. Bullinaria, & P. Cairns (Eds.), *Connectionist models of memory and language* (pp. 207–228). London: UCL Press.

Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review, 103*, 566–581.

Chater, N., & Oaksford, M. (1990). Autonomy, implementation and cognitive architecture: A reply to Fodor and Pylyshyn. *Cognition, 34*, 93–107.

Chater, N., & Oaksford, M. (2000). The rational analysis of mind and behavior. *Synthese, 122*, 93–131.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367–405.

Cherniak, C. (1986). *Minimal rationality*. Cambridge, MA: MIT Press.

Chew, S. H. (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica, 51*, 1065–1092.

Chomsky, N. (1980). *Rules and representations*. Cambridge, MA: MIT Press.

Christiansen, M. H., & Chater, N. (Eds.). (2001). *Connectionist psycholinguistics*. Westport, CT: Ablex.

Christiansen, M., Chater, N., & Seidenberg, M. (Eds.) (1999). Special issue on connectionist natural language processing. *Cognitive Science, 23*(4).

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences, 4*, 317–370.

Colman, A. M. (1995). *Game theory and its applications in the social and biological sciences* (2nd ed.). Oxford: Butterworth-Heinemann.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*, 589–608.

Cooper, G., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning, 9*, 309–347.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13*, 21–27.

Crawford, C., Smith, M., & Krebs, D. (1987). *Sociobiology and psychology*. Hillsdale, NJ: Erlbaum.

Cyert, R., & de Groot, M. (1974). Rational expectations and Bayesian analysis. *Journal of Political Economy, 82*, 521–536.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P. Todd, & The ABC Group (Eds.), *Simple heuristics that make us smart* (pp. 97–118). Oxford: Oxford University Press.

Davidson, D. (1984). *Inquiries into truth and interpretation*. Oxford: Clarendon Press.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81*, 95–106.

de Canio, S. (1979). Rational expectations and learning from experience. *Quarterly Journal of Economics, 93*, 47–57.

Dhami, M. K., & Ayton, P. (2001). Bailing and jailing the fast and frugal way. *Journal of Behavioral Decision Making, 14*, 141–168.

Dhami, M. K., & Harries, C. (2002). Fast and frugal versus regression models of human judgment. *Thinking and Reasoning, 7*, 5–27.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory process model for judgments of likelihood. *Psychological Review, 106*, 180–209.

Duda, R. O., & Hart, P. E. (1973). *Patterns classification and scene analysis*. New York: Wiley.

Einhorn, H. J. (1970). The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin, 73*, 221–230.

Einhorn, H. J. (1971). Use of nonlinear, noncompensatory models as a function of task and amount of information. *Organizational Behavior and Human Performance, 6*, 1–27.

Einhorn, H. J. (1972). Expert measurement and mechanical combination. *Organizational Behavior and Human Performance, 13*, 171–192.

Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics, 75*, 643–669.

Elster, J. (Ed.). (1986). *Rational choice*. Oxford: Basil Blackwell.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning*. London: Routledge & Kegan Paul.

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences*. Hillsdale, NJ: Erlbaum.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove, Sussex: Psychology Press.

Evans, J. St. B. T., & Over, D. (1997). Rationality in reasoning: The problem of deductive competence. *Cahiers de Psychologie Cognitive, 16*, 1–35.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning*. Hillsdale, NJ: Erlbaum.

Fishburn, P. C. (1983). Transitive measurable utility. *Journal of Economic Theory, 31*, 293–317.

Flood, M. M. (1958). Some experimental games. *Management Science, 5*, 5–26.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3–71.

Frey, B. (1998). *Graphical models for machine learning and digital communication*. Cambridge, MA: MIT Press.

Friedman, M. (1953). *Methodology of positive economics*. Chicago: Chicago University Press.

Friedman, N., & Goldszmit, M. (1996). Learning Bayesian networks with local structure. In *Proceedings of the 12th conference on uncertainty in artificial intelligence* (pp. 252–262). San Mateo, CA: Morgan Kaufmann.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Ganzach, Y. (1995). Nonlinear models of clinical judgment: Meehl's data revisited. *Psychological Bulletin, 118*, 422–429.

Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-completeness*. San Francisco: W.H. Freeman.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world*. Oxford: Oxford University Press.

Gigerenzer, G. (2001). The adaptive toolbox. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 37–50). Cambridge, MA: MIT Press.

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

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason. In G. Gigerenzer, P. Todd, & The ABC Group (Eds.), *Simple heuristics that make us smart* (pp. 75–95). Oxford: Oxford University Press.

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base-rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 513–525.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684–704.

Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.

Gigerenzer, G., & Selten, R. (2001). Rethinking rationality. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 1–13). Cambridge, MA: MIT Press.

Gigerenzer, G., Todd, P. & The ABC Group (1999a). *Simple heuristics that make us smart*. Oxford: Oxford University Press.

Gigerenzer, G., & Todd, P. (1999b). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. Todd, & The ABC Group (Eds.), *Simple heuristics that make us smart* (pp. 3–34). Oxford: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P. Todd, & The ABC Group (Eds.), *Simple heuristics that make us smart* (pp. 37–58). Oxford: Oxford University Press.

Good, I. J. (1971). Twenty seven principles of rationality. In V. P. Godambe & D. A. Sprott (Eds.), *Foundations of statistical inference*. Toronto: Holt, Rhinehart & Winston.

Hahn, U., & Chater, N. (1998). Similarity and rules: Distinct? Exhaustive? Empirically distinguishable? *Cognition, 65*, 197–230.

Hammond, K. R., & Summers, D. A. (1965). Cognitive dependence on linear and non-linear cues. *Psychological Review, 72*, 215–224.

Harsanyi, J. C., & Selten, R. (1988). *A general theory of equilibrium selection in games*. Cambridge, MA: MIT Press.

Hertwig, R. (2000). The questionable utility of "cognitive ability" in explaining cognitive illusions. *Behavioral and Brian Sciences, 23*, 678–679.

Hinton, G. E. (1989). Connectionist learning procedures. *Artificial Intelligence, 40*, 185–234.

Hintzman, D. L. (1984). MINERVA2: A simulation model of human memory. *Behavior Research Methods, Instruments and Computers, 16*, 96–101.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple trace memory model. *Psychological Review, 93*, 411–428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*, 528–551.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.

Intrator, N. (1993). On the use of projection pursuit constraints for training neural networks. In S. J. Hanson, J. D. Cowan, & C. L. Giles (Eds.), *Advances in neural information processing systems* (Vol. 5, pp. 3–10). San Mateo, CA: Morgan Kaufman.

Jeannerod, M. (1988). *The neural and behavioral organisation of goal-directed movements*. Oxford: Oxford University Press.

Juslin, P. & Persson, M. (2001). PROBabilities from EXemplars (PROBEX): A "Lazy" algorithm for probabilistic inference from generic knowledge. Unpublished manuscript, Department of Psychology, Uppsala University, Sweden.

Kacelnik, A. (1998). Normative and descriptive models of decision making: Time discounting and risk sensitivity. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 54–70). Oxford: Oxford University Press.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

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

Kolodner, J. L. (1993). *Case-based reasoning*. San Mateo, CA: Morgan Kaufman.

Kreps, D. M. (1990). *A course in microeconomic theory*. New York: Harvester Wheatsheaf.

Kripke, S. (1982). *Wittgenstein on rules and private languages: An elementary exposition*. Oxford: Blackwell.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. E. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 111–194). Princeton: Princeton University Press.

Leeuwenberg, E., & Boselie, F. (1988). Against the likelihood principle in visual form perception. *Psychological Review, 95*, 485–491.

Libby, R. (1976). Man versus model of man: Some conflicting evidence. *Organizational Behavior and Human Performance, 16*, 1–12.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *Annals of Mathematical Statistics, 27*, 986–1005.

Ling, C. X., & Marinov, M. (1993). Answering the connectionist challenge: A symbolic model of learning the past tense of English verbs. *Cognition, 49*, 235–290.

Ling, C. X., & Marinov, M. (1994). A symbolic model of the nonconscious acquisition of information. *Cognitive Science, 18*, 595–621.

Loewenstein, G. (1992). The fall and rise of psychological explanations in the economics of intertemporal choice. In G. Loewenstein & J. Elster (Eds.), *Choice over time* (pp. 3–34). New York: Russell Sage Foundation.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal, 92*, 805–824.

McClelland, J. L. (1998). Connectionist models of Bayesian inference. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 21–53). Oxford: Oxford University Press.

MacLeod, P., Plunkett, K., & Rolls, E. T. (1998). *Introduction to connectionist modelling of cognitive processes*. Oxford: Oxford University Press.

McCloskey, D. N. (1985). *The rhetoric of economics*. Madison: University of Wisconsin Press.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.

McDermott, D. (1987). A critique of pure reason. *Computational Intelligence, 3*, 151–160.

McFarland, D., & Houston, A. (1981). *Quantitative ethology: The state space approach*. London: Pitman.

Mackay, D. J. C. (1992). A practical Bayesian framework for backpropagation networks. *Neural computation, 4*, 448–462.

Machina, M. J. (1982). "Expected utility" analysis without the independence axiom. *Econometrica, 39*, 277–323.

Marr, D. (1982). *Vision*. San Francisco: W.H. Freeman.

Martignon, L., & Hoffrage, U. (1999). Why does one reason decision making work? A case study in ecological rationality. In G. Gigerenzer, P. Todd, & The ABC Group (Eds.), *Simple heuristics that make us smart* (pp. 119–140). Oxford: Oxford University Press.

Martignon, L., & Laskey, K. (1999). Bayesian benchmarks for fast and frugal heuristics. In G. Gigerenzer, P. Todd, & The ABC Group (Eds.), *Simple heuristics that make us smart* (pp. 169–188). Oxford: Oxford University Press.

Massaro, D. W. (1987). *Speech perception by ear and eye*. Hillsdale, NJ: Erlbaum.

May, K. O. (1954). Intransitivity, utility, and the aggregation of preference patterns. *Econometrica, 22*, 1–13.

Maynard-Smith, J., & Price, G. R. (1973). The logic of animal conflict. *Nature, 246*, 15–18.

Meehl, P. E. (1954). *Clinical versus statistical predictions: A theoretical analysis and revision of the literature*. Minneapolis: University of Minnesota Press.

Messick, D. M. (1991). On the evolution of group-based altruism. In R. Selten (Ed.), *Game equilibrium models I: Evolution and game dynamics* (pp. 304–328). Berlin: Springer-Verlag.

Minsky, M., & Papert, S. (1969). *Perceptrons: An introduction to computational geometry*. Cambridge, MA: MIT Press.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review, 76*, 165–178.

Movellan, J. R., & Chadderdon, G. (1996). Cognition and the statistics of natural signals. In G. Cottrell (Ed.), *Proceedings of the eighteenth annual conference of the cognitive science society* (pp. 381–394). Hillsdale, NJ: Erlbaum.

Movellan, J. R., & McClelland, J. L. (1995). Stochastic interactive processing, channel separability and optimal perceptual inference: An examination of Morton's Law. Technical Report PDP.CNS.95.4, Carnegie Mellon University, Pittsburgh, PA.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica, 29*, 315–335.

Nash, J. (1950). The bargaining problem. *Econometrica, 28*, 155–162.

Neal, R. (1993). Bayesian learning via stochastic dynamics. In S. J. Hanson, J. D. Cowan, & C. L. Giles (Eds.), *Advances in neural information processing systems* (Vol. 5, pp. 475–482). San Mateo, CA: Morgan Kaufman.

Nelson, R., & Winter, S. (1982). *An evolutionary theory of economic capabilities and behavior*. Cambridge, MA: Harvard University Press.

Newell, A. (1982). The knowledge level. *Artificial Intelligence, 18*, 87–127.

Newell, A. (1991). *Unified theories of cognition*. Cambridge, UK: Cambridge University Press.

Newell, B. R., & Shanks, D. R. (2001). Take the best or look at the rest? Factors influencing 'one-reason' decision-making. Unpublished manuscript, Department of Psychology, University College London.

Nosofsky, R. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115*, 39–57.

Nosofsky, R. M. (1990). Relations between exemplar similarity and likelihood models of classification. *Journal of Mathematical Psychology, 34*, 393–418.

Oaksford, M., & Chater, N. (1991). Against logicist cognitive science. *Mind and Language, 6*, 1–38.

Oaksford, M., & Chater, N. (1993). Reasoning theories and bounded rationality. In K. I. Manktelow & D. E. Over (Eds.), *Rationality* (pp. 31–60). London: Routledge.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608–631.

Oaksford, M., & Chater, N. (1995). Theories of reasoning and the computational explanation of everyday inference. *Thinking and Reasoning, 1*, 121–152.

Oaksford, M., & Chater, N. (1998a). *Rationality in an uncertain world: Essays in the cognitive science of human reasoning*. Hove, Sussex: Psychology Press.

Oaksford, M., & Chater, N. (Eds.). (1998b). *Rational models of cognition*. Oxford: Oxford University Press.

Oaksford, M., Chater, N., & Stenning, K. (1990). Connectionism, classical cognitive science and experimental psychology. *AI & Society, 4*, 73–90.

Paris, J. (1992). *The uncertain reasoner's companion*. Cambridge: Cambridge University Press.

Parzen, E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics, 35*, 1065–1076.

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

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

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1990). The adaptive decision maker: Effort and accuracy in choice. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 129–153). Chicago: Chicago University Press.

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

Payne, J. W., Bettman, J. R., & Luce, M. F. (1996). When time is money: Decision behavior under opportunity-cost time pressure. *Organizational Behavior and Human Decision Processes, 66*, 131–152.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Mateo, CA: Morgan Kaufman.

Persson, M., & Juslin, P. (1999). The ecological rationality of PROBEX. Unpublished manuscript, Department of Psychology, Uppsala University, Sweden.

Pinker, S. (1998). *How the mind works*. Harmondsworth, UK: Penguin.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56–115.

Pomerantz, J. R., & Kubovy, M. (1987). Theoretical approaches to perceptual organization. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance, Volume II: Cognitive processes and performance* (pp. 36.1–36.46). New York: Wiley.

Pylyshyn, Z. W. (1987). *The robot's dilemma: The frame problem in artificial intelligence*. Norwood, NJ: Ablex.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. Los Altos: Morgan Kaufmann.
Reiner, R. (1995). Arguments against the possibility of perfect rationality. *Minds and Machines, 5*, 373–389.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness in the variation of reinforcement and non reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditional II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics and how can we tell? In G. Gigerenzer, P. Todd, & The ABC Group (Eds.), *Simple heuristics that make us smart* (pp. 141–168). Oxford: Oxford University Press.
Rips, L. J. (1994). *The psychology of proof*. Cambridge, MA: MIT Press.
Roth, A. (1996). Individual rationality as a useful approximation: Comments on Tversky's "Rational theory of constructive choice". In K. J. Arrow, E. Colombatto, M. Perlman, & C. Schmidt (Eds.), *The rational foundations of economic behavior* (pp. 198–202). London: Macmillan.
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*, 533–536.
Rumelhart, D. E., McClelland, J. L., & The PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructures of cognition* (Vols. 1 & 2). Cambridge, MA: MIT Press.
Samuelson, P. (1937). A note on the measurement of utility. *Review of Economic Studies, 4*, 155–161.
Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.
Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin, 66*, 178–200.
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.
Shanks, D. R. (1995a). Is human learning rational? *Quarterly Journal of Experimental Psychology, 48A*, 257–279.
Shanks, D. (1995b). *The psychology of associative learning*. Cambridge: Cambridge Univerisity Press.
Shepard, R. N. (1967). On subjectively optimum selections among multi-attribute alternatives. In W. Edwards & A. Tversky (Eds.), *Decision making* (pp. 257–283). Harmondsworth, UK: Penguin.
Simon, H. A. (1959). Theories of decision-making in economics and behavioral science. *American Economic Review, 49*, 253–283.
Simon, H. A. (1992). *Economics, bounded rationality and the cognitive revolution*. Aldershot: Elgar.
Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Hillsdale, NJ: Erlbaum.
Stanovich, K. E., & West, R. F. (1998a). Cognitive ability and variation in selection task performance. *Thinking and Reasoning, 4*, 193–230.
Stanovich, K. E., & West, R. F. (1998b). Individual differences in framing and conjunction effects. *Thinking and Reasoning, 4*, 289–317.
Stanovich, K. E., & West, R. F. (1998c). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*, 161–188.
Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*, 645–726.
Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive science*. Oxford: Oxford University Press.
Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.
Stich, S. (1990). *The fragmentation of reason*. Cambridge, MA: MIT Press.
Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language, 27*, 597–632.
Todd, I. A., & Kacelnik, A. (1993). Psychological mechanisms and the Marginal Value Theorem: Dynamics of scalar memory for travel time. *Animal Behavior, 46*, 765–775.
Todd, P., & Gigerenzer, G. (1999). What we have learned (so far). In G. Gigerenzer, P. Todd, & The ABC Group (Eds.), *Simple heuristics that make us smart* (pp. 357–365). Oxford: Oxford University Press.
Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76*, 31–48.
Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79*, 281–299.
Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 125*, 1124–1131.
Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59*, 251–278.
van Damme, E. (1991). *Stability and the perfection of Nash equilibria*. Berlin: Springer-Verlag.
van der Helm, P. A., & Leeuwenberg, E. L. J. (1996). Goodness of visual regularities: A non-transformational approach. *Psychological Review, 103*, 429–456.
von Helmholtz, H. (1910/1962). In J. P. Southall (Ed. and translation). *Treatise on physiological optics* (Vol. 3). New York: Dover.
von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.
Weber, E. U., Goldstein, W. M., & Barlas, S. (1995). And let us not forget memory: The role of memory processes and techniques in the study of judgment and choice. *The Psychology of Learning and Motivation, 32*, 33–81.

Received 16 January 2001