

Evolutionary Versus Instrumental Goals:  
 How Evolutionary Psychology Misconceives Human Rationality  
 Keith E. Stanovich  
 University of Toronto  
 Richard F. West  
 James Madison University

Stanovich, K. E., & West, R. F. (2003). Evolutionary versus instrumental goals: How evolutionary psychology misconceives human rationality. In D. E. Over (Ed.), Evolution and the psychology of thinking: The debate, Psychological Press. [Series on Current Issues in Thinking and Reasoning]

Evolutionary Versus Instrumental Goals:  
 How Evolutionary Psychology Misconceives Human Rationality

An important research tradition in the cognitive psychology of reasoning--called the heuristics and biases approach--has firmly established that people's responses often deviate from the performance considered normative on many reasoning tasks. For example, people assess probabilities incorrectly, they display confirmation bias, they test hypotheses inefficiently, they violate the axioms of utility theory, they do not properly calibrate degrees of belief, they overproject their own opinions onto others, they display illogical framing effects, they uneconomically honor sunk costs, they allow prior knowledge to become implicated in deductive reasoning, and they display numerous other information processing biases (for summaries of the large literature, see Baron, 1998, 2000; Dawes, 1998; Evans, 1989; Evans & Over, 1996; Kahneman & Tversky, 1972, 1984, 2000; Kahneman, Slovic, & Tversky, 1982; Nickerson, 1998; Shafir & Tversky, 1995; Stanovich, 1999; Tversky, 1996).

It has been common for these empirical demonstrations of a gap between descriptive and normative models of reasoning and decision making to be taken as indications that systematic irrationalities characterize human cognition. However, over the last decade, an alternative interpretation of these findings has been championed by various evolutionary psychologists, adaptationist modelers, and ecological theorists (Anderson, 1990, 1991; Chater & Oaksford, 2000; Cosmides & Tooby, 1992; 1994b, 1996; Gigerenzer, 1996a; Oaksford & Chater, 1998, 2001; Rode, Cosmides, Hell, & Tooby, 1999; Todd & Gigerenzer, 2000). They have reinterpreted the modal response in most of the classic heuristics and biases experiments as indicating an optimal information processing adaptation on the part of the subjects. It is argued by these investigators that the research in the heuristics and biases tradition has not demonstrated human irrationality at all and that a Panglossian position (see Stanovich & West, 2000) which assumes perfect human rationality is the proper default position to take.

It will be argued in this chapter that although the work of the evolutionary psychologists has uncovered some fundamentally important things about human cognition, these theorists have misconstrued the nature of human rationality and have conflated important distinctions in this domain. What these theorists have missed (or failed to sufficiently emphasize) is that definitions of rationality must coincide with the level of the entity whose optimization is at issue. This admonition plays out most directly in the distinction between evolutionary rationality and instrumental

rationality--necessitated by the fact that the optimization procedures for replicators and for vehicles (to use Dawkins', 1976, terms) need not always coincide. The distinction follows from the fact that the genes--as subpersonal replicators--can increase their fecundity and longevity in ways that do not necessarily serve the instrumental goals of the vehicles built by the genome (Skyrms, 1996; Stanovich, 1999). Despite their frequent acknowledgements that the conditions in the environment of evolutionary adaptedness [EEA] do not match those of modern society, evolutionary psychologists have a tendency to background potential mismatches between genetic interests and personal interests.

We will argue below that dual process models of cognitive functioning provide a way of reconciling the positions of the evolutionary psychologists and researchers in the heuristics and biases tradition. Such models acknowledge the domain specificity of certain modular processes emphasized by the evolutionary psychologists. But importantly, they also posit general, interactive, nonautonomous, and central serial-processing operations of executive control and problem solving that serve to guarantee instrumental rationality by overriding the responses generated by autonomous modules when the latter threaten optimal outcomes at the personal level.

#### Debates About the Normative Response in Heuristics and Biases Tasks: Some Examples

The empirical data pattern that provoked our attempted reconciliation of the positions of the evolutionary psychologists and researchers in the heuristics and biases tradition is the repeated finding in our research (Stanovich & West, 1998a, 1998b, 1998c, 1998d, 1999, 2000) that the modal response was different from the response given by the more cognitively able subjects. We have related this finding to the disputes about which response is normative in various heuristics and biases tasks. An example is provided by the most investigated task in the entire reasoning and problem solving literature--Wason's (1966) selection task. The participant is shown four cards lying on a table showing two letters and two numbers (A, D, 3, 8). They are told that each card has a number on one side and a letter on the other and that the experimenter has the following rule (of the if P, then Q type) in mind with respect to the four cards: "If there is a vowel on one side of the card, then there is an even number on the other side". The participant is then told that he/she must turn over whichever cards are necessary to determine whether the experimenter's rule is true or false. Performance on such abstract versions of the selection task is extremely low (Evans, Newstead, & Byrne, 1993; Manktelow, 1999; Newstead & Evans, 1995). Typically, less than 10% of participants make the correct selections of the A card (P) and 3 card (not-Q)--the only two cards that could falsify the rule. The most common incorrect choices made by participants are the A card and the 8 card (P and Q) or the selection of the A card only (P).

Numerous alternative explanations for the preponderance of incorrect PQ and P responses have been given (see Evans, Newstead, & Byrne, 1993; Hardman, 1998; Johnson-Laird, 1999; Liberman & Klar, 1996; Margolis, 1987; Newstead & Evans, 1995; Oaksford & Chater, 1994; Sperber, Cara, & Girotto, 1995; Stanovich & West, 1998a). What is important in the present context is that several of these alternative theories posit that the incorrect PQ response results from the operation of efficient and optimal cognitive mechanisms. For example, Oaksford and Chater (1994, 1996; see also Nickerson, 1996) argue that rather than interpreting the task as one of deductive reasoning (as the experimenter intends), many people interpret it as an inductive problem of probabilistic hypothesis testing (see Evans & Over, 1996). They show that the P and Q response is actually the expected one if an inductive interpretation of the problem is assumed along with optimal data selection (which they modeled with a Bayesian analysis). Although their model is different, Sperber et al. (1995) stress that selection task performance is driven by optimized cognitive mechanisms.

They explain selection task performance in terms of inferential comprehension mechanisms that are "geared towards the processing of optimally relevant communicative behaviors" (p. 90).

Our second example of theorists defending as rational the response that heuristics and biases researchers have long considered incorrect is provided by the much-investigated Linda Problem (Tversky & Kahneman, 1983):

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable.

- a. Linda is a teacher in an elementary school
- b. Linda works in a bookstore and takes Yoga classes
- c. Linda is active in the feminist movement
- d. Linda is a psychiatric social worker
- e. Linda is a member of the League of Women Voters
- f. Linda is a bank teller
- g. Linda is an insurance salesperson
- h. Linda is a bank teller and is active in the feminist movement

Because alternative h (Linda is a bank teller and is active in the feminist movement) is the conjunction of alternatives c and f, the probability of h cannot be higher than that of either c (Linda is active in the feminist movement) or f (Linda is a bank teller), yet 85% of the participants in Tversky and Kahneman's (1983) study rated alternative h as more probable than f, thus displaying the so-called conjunction fallacy. Those investigators argued that logical reasoning on the problem (all feminist bank tellers are also bank tellers, so h cannot be more probable than f) was trumped by a heuristic based on so-called representativeness that primes answers to problems based on an assessment of similarity (a feminist bank teller seems to overlap more with the description of Linda than does the alternative "bank teller"). Of course, logic dictates that the subset (feminist bank teller)--superset (bank teller) relationship should trump assessments of representativeness when judgments of probability are at issue.

However, several investigators have suggested that rather than illogical cognition, it is rational pragmatic inferences that lead to the violation of the logic of probability theory in the Linda Problem (see Adler, 1991; Dulany & Hilton, 1991; Politzer & Noveck, 1991; Slugoski & Wilson, 1998). Hilton (1995) summarizes the view articulated in these critiques by arguing that "the inductive nature of conversational inference suggests that many of the experimental results that have been attributed to faulty reasoning may be reinterpreted as being due to rational interpretations of experimenter-given information" (p. 264).

In short, these critiques imply that displaying the conjunction fallacy is a rational response triggered by the adaptive use of social cues, linguistic cues, and background knowledge (see Hilton, 1995). For example, Macdonald and Gilhooly (1990) argue that it is possible that subjects will "usually assume the questioner is asking the question because there is some reason to suppose that Linda might be a bank teller and the questioner is interested to find out if she is....If Linda were chosen at random from the electoral register and 'bank teller' was chosen at random from some list of occupations, the probability of them corresponding would be very small, certainly less than 1 in 100....the question itself has suggested to the subjects that Linda could be a feminist bank teller. Subjects are therefore being asked to judge how likely it is that Linda is a feminist bank teller when there is some unknown reason to suppose she is, which reason has prompted the question itself" (p. 59).

Hilton (1995; see Dulany & Hilton, 1991) provides a similar explanation of subjects' behavior on the Linda Problem. Under the assumption that the detailed information given about the target means that the experimenter knows a considerable amount about Linda, then it is reasonable to think that the phrase "Linda is a bank teller" does not contain the phrase "and is not active in the feminist movement" because the experimenter already knows this to be the case. If "Linda is a bank teller" is interpreted in this way, then rating h as more probable than f no longer represents a conjunction fallacy.

Several investigators have suggested that pragmatic inferences lead to seeming violations of the logic of probability theory in the Linda Problem (see Adler, 1984, 1991; Hertwig & Gigerenzer, 1999; Politzer & Noveck, 1991; Slugoski & Wilson, 1998). Most of these can be analyzed in terms of Grice's (1975) norms of rational communication (see Hilton & Slugoski, 2000; Sperber & Wilson, 1986; Sperber et al., 1995) which require that the speaker be cooperative with the listener--and one of the primary ways that speakers attempt to be cooperative is by not being redundant. The key to understanding the so-called Gricean maxims of communication is to realize that to understand a speaker's meaning the listener must comprehend not only the meaning of what is spoken but also what is implicated in a given context assuming that the speaker intends to be cooperative. And Hilton (1995) is at pains to remind us that these are rational aspects of communicative cognition. They are rational heuristics as opposed to the suboptimal shortcuts as emphasized in the heuristics and biases literature. Thus, they are not to be seen as processing modes that are likely to be given up for more efficient processing modes when the stakes become high:

"However, it is not clear why increasing the financial stakes in an experiment should cause respondents to abandon an interpretation that is pragmatically correct and rational...Incentives are not going to make respondents drop a conversationally rational interpretation in favor of a less plausible one in the context...the conversational inference approach does not predict that increased incentives lead respondents to change an interpretation that seems rational in the context" (Hilton, 1995, pp. 265-266)

Clearly, in the view of these theorists, committing the conjunction fallacy in such contexts does not represent a cognitive error.

Many theorists have linked their explanation of Linda-problem performance to the automatic linguistic socialization of information. These theorists commonly posit that the socialization tendency reflects evolutionary adaptations in the domain of social intelligence. This linkage stems from many theories that, although varied in their details, all posit that much of human intelligence has foundations in social interaction (Baldwin, 2000; Barton & Dunbar, 1997; Brothers, 1990; Byrne & Whiten, 1988; Bugental, 2000; Caporael, 1997; Cosmides, 1989; Cummins, 1996; Dunbar, 1998; Humphrey, 1976; Jolly, 1966; Kummer, Daston, Gigerenzer, & Silk, 1997; Mithen, 1996; Tomasello, 1999; Whiten & Byrne, 1997).

In a seminal essay that set the stage for this hypothesis, Nicholas Humphrey (1976) argued that the impetus for the development of primate intelligence was the need to master the social world. Based on his observation of nonhuman primates, Humphrey (1976) concluded that the knowledge and information processing necessary to engage efficiently with the physical world seemed modest compared to the rapidly changing demands of the social world with its everchanging. Humphrey (1976) posited that the latter was the key aspect of the environment that began to bootstrap higher intelligence in all primates.

This social, or interactional intelligence, forms that substrate upon which all future evolutionary and cultural developments in modes of thought are overlaid. That such social intelligence forms the basic substrate upon which all higher forms of intelligence must build leads to the important assumption that a social orientation toward problems is always available as a default

processing mode when computational demands become onerous. The cognitive illusions demonstrated by three decades of work in problem solving, reasoning, and decision making (Evans, 1989; Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 1996, 2000; Stanovich, 1999) seem to bear this out. As in the Linda Problem and four-card selection task discussed above, the literature is full of problems where an abstract, decontextualized--but computationally expensive--approach is required for the normatively appropriate answer. However, often, alongside such a solution, resides a tempting social approach ("oh, yeah, the author of this knows a lot about Linda") that with little computational effort will prime a response.

Since our theme has now been established with the selection task and Linda problem examples, our final two examples of theorists defending as rational the response that heuristics and biases researchers have long considered incorrect will be described only briefly.

#### Covariation Detection

The 2 x 2 covariation detection task is run in a variety of different formats (Levin, Wasserman, & Kao, 1993; Stanovich & West, 1998d; Wasserman, Dornier, & Kao, 1990). In one, for example, subjects are asked to evaluate the efficacy of a drug based on a hypothetical well-designed scientific experiment. They are told that:

- 150 people received the drug and were cured
- 150 people received the drug and were not cured
- 300 people did not receive the drug and were cured
- 75 people did not receive the drug and were not cured

These data correspond to four cells of the 2 x 2 contingency table traditionally labeled A, B, C, and D (see Levin et al., 1993). Subjects are asked to evaluate the effectiveness of the drug on a scale. In this case, they have to detect that the drug is ineffective. In fact, not only is it ineffective, it is positively harmful. Only 50% of the people who received the drug were cured (150 out of 300), but 80% of those who did not receive the drug were cured (300 out of 375).

Much previous experimentation has produced results indicating that subjects weight the cell information in the order cell A > cell B > cell C > cell D--cell D receiving the least weight and/or attention (see Arkes & Harkness, 1983; Kao & Wasserman, 1993; Schustack & Sternberg, 1981). The tendency to ignore cell D is nonnormative, as indeed is any tendency to differentially weight the four cells. The normatively appropriate strategy (see Allan, 1980; Kao & Wasserman, 1993; Shanks, 1995) is to use the conditional probability rule--subtracting from the probability of the target hypothesis when the indicator is present the probability of the target hypothesis when the indicator is absent. Numerically, the rule amounts to calculating the  $\phi$  statistic:  $[A/(A+B)] - [C/(C+D)]$  (see Allan, 1980). For example, the  $\phi$  value for the problem presented above is -.300, indicating a fairly negative association.

Despite the fact that it is a nonnormative strategy, the modal subject in such experiments underweights (sometimes markedly underweights, see Stanovich & West, 1998d) cell D. However, Anderson (1990) has modeled the 2 x 2 contingency assessment experiment using a model of optimally adapted information processing and come to a startling conclusion. He demonstrates that an adaptive model can predict the much-replicated finding that the D cell (cause absent and effect absent) is vastly underweighted (but see Over & Green, in press) and concludes that "this result makes the point that there need be no discrepancy between a rational analysis and differential weighting of the cells in a 2 x 2 contingency table" (p. 160). Thus, here again in another task is the pattern where the modal response is nonnormative--but that response has been defended from the standpoint of an adaptationist analysis.

#### Probability Matching

The probabilistic contingency experiment has many versions in psychology (Gal & Baron, 1996; Tversky & Edwards, 1966). In one, the subject sits in front of two lights (one red and one blue) and is told that she or he is to predict which of the lights will be flashed on each trial and that there will be several dozen of such trials (subjects are often paid money for correct predictions). The experimenter has actually programmed the lights to flash randomly, with the provision that the red light will flash 70 percent of the time and the blue light 30 percent of the time. Subjects do quickly pick up the fact that the red light is flashing more, and they predict that it will flash on more trials than they predict that the blue light will flash. Most often, they switch back and forth, predicting the red light roughly 70 percent of the time and the blue light roughly 30 percent of the time.

This strategy of probability matching is suboptimal because it insures that, in this example, the subject will correctly predict only 58% of the time ( $.7 \times .7 + .3 \times .3$ ) compared to the 70% hit rate achieved by predicting the more likely color on each trial. In fact, much experimentation has indicated that animals and humans often fail to maximize expected utility in the probabilistic contingency experiment<sup>1</sup> (Estes, 1964, 1976; Gallistel, 1990; Tversky & Edwards, 1966). Nevertheless, Gigerenzer (1996b; see also, Cooper, 1989) shows how probability matching could, under some conditions, actually be an evolutionarily stable strategy (see Skyrms, 1996, for many such examples). Thus, we have in probability matching our final example of how a nonnormative response tendency is defended on an evolutionary or adaptationist account.

#### Dissociations Between Cognitive Ability and the Modal Response in Heuristics and Biases Tasks

We will argue in this chapter that, in each of these examples evolutionary rationality has dissociated from normative rationality--where the latter is viewed as utility maximization for the individual organism (instrumental rationality) and the former is defined as survival probability at the level of the gene (Dawkins, 1976, 1982). Our conceptualization of these findings explicitly acknowledges the impressive record of descriptive accuracy enjoyed by a variety of adaptationist and evolutionary models in predicting the modal response (Anderson, 1990, 1991; Gigerenzer, 1996b; Oaksford & Chater, 1994, 1996; Rode et al., 1999), but our account attempts to make sense of another important empirical fact--that cognitive ability often dissociates from the response deemed adaptive on an evolutionary analysis (Stanovich, 1999; Stanovich & West, 2000). Specifically, we have repeatedly found that in cases where the normative response is not the modal response, the subjects in the sample who were the highest in cognitive ability gave the normative response rather than the modal response. This is true for each of the four tasks described above.

For example, Table 1 presents the results from an investigation of ours (Stanovich & West, 1998a) using a selection task with a nondeontic rule, the so-called Destination rule (in this instance: If 'Baltimore' is on one side of the ticket, then 'plane' is on the other side of the ticket). The Table presents the mean SAT scores for several of the dominant choices on this selection rule (the SAT test is a test used for university admissions in the United States that is highly loaded on psychometric g). From the Table, it is clear that respondents giving the deductively correct P and not-Q response had the highest SAT scores--followed by the subjects choosing the P card only. All other responses, including the modal P and Q response (chosen by 49% of the sample), were given by subjects having SAT scores almost 100 points lower than those giving the correct response under a deductive construal. It is to the credit of models of optimal data selection (Oaksford & Chater, 1994) that they predict the modal response. But we are left with the seemingly puzzling finding that the response deemed optimal under such an analysis (PQ) is given by subjects of substantially lower general intelligence than the minority giving the response deemed correct under a strictly deductive interpretation of the problem (PNQ).

A similar puzzle surrounds findings on the Linda conjunction problem. Gricean analyses assume that those subjects committing the conjunction fallacy in such a contrived problem are reflecting the evolved use of sociolinguistic cues. Because this group is in fact the vast majority in most studies--and because the use of such pragmatic cues and background knowledge is often interpreted as reflecting adaptive information processing (e.g., Hilton, 1995)--it might be expected that these individuals would be the subjects of higher cognitive ability. We found the contrary. In our study (Stanovich & West, 1998b), we examined the performance of 150 subjects on the Linda Problem. Consistent with the results of previous experiments on this problem (Tversky & Kahneman, 1983), 80.7% of our sample committed the conjunction effect--they rated the feminist bank teller alternative as more probable than the bank teller alternative. However, the mean SAT score of the 121 subjects who committed the conjunction fallacy was 82 points lower than the mean score of the 29 who avoided the fallacy. This difference was highly significant and it translated into an effect size of .746 (which Rosenthal & Rosnow, 1991, classify as large). Thus, the pragmatic interpretations of why the conjunction effect is the modal response on this task might well be correct--but the modal response happens not to be the one given by the most intelligent subjects in the sample.

Likewise, in the 2 x 2 covariation detection experiment, we have found (Stanovich & West, 1998d) that it is those subjects weighting cell D more equally (not those underweighting the cell in the way that the adaptationist model dictates) who are higher in cognitive ability and who tend to respond normatively on other tasks. Again, Anderson (1990, 1991) might well be correct that a rational model of information processing in the task predicts underweighting of cell D by most subjects, but more severe underweighting is in fact associated with lower cognitive ability in our individual differences analyses.

Finally, in several recently completed experiments on probability matching using a variety of different paradigms (West & Stanovich, 2002) we have found a similar pattern. For example, in one experiment involving choices among general strategies for approaching the probabilistic prediction task, subjects were given the following task description:

A die with 4 red faces and 2 green faces will be rolled 60 times. Before each roll you will be asked to predict which color (red or green) will show up once the die is rolled. You will be given one dollar for each correct prediction. Assume that you want to make as much money as possible. What strategy would you use in order to make as much money as possible by making the most correct predictions?

They were asked to choose from among the following five strategies:

Strategy A: Go by intuition, switching when there has been too many of one color or the other.

Strategy B: Predict the more likely color (red) on most of the rolls but occasionally, after a long run of reds, predict a green.

Strategy C: Make predictions according to the frequency of occurrence (4 of 6 for red and 2 of 6 for green). That is, predict twice as many reds as greens.

Strategy D: Predict the more likely color (red) on all of the 60 rolls.

Strategy E: Predict more red than green, but switching back and forth depending upon "runs" of one color or the other.

The probability matching strategy corresponds to Strategy C here, and the normatively optimal strategy is Strategy D which maximizes expected utility. Table 2 presents the number of subjects choosing each of the five strategies and their mean SAT scores. The probability matching and maximizing strategies were both preferred over the three foil strategies, with the former being the modal choice. Again, the choice defensible on evolutionary grounds (probability matching, see

Gigerenzer, 1996b), is the modal choice. But again, as before, it is the maximizing, normatively-dictated choice that is the choice of the subjects with the highest intellectual ability. The mean SAT scores of those choosing the maximizing choice was 55 points higher than those who preferred probability matching ( $p < .001$ ).

#### Reconciling the Two Data Patterns Within a Two-Process View

We see in the results just reviewed two basic patterns that must be reconciled. The evolutionary psychologists and optimal data selection theorists correctly predict the modal response in a host of heuristics and biases tasks. Yet in all of these cases--despite the fact that the adaptationist models predict the modal response quite well--individual differences analyses demonstrate associations that also must be accounted for. Correct responders on the nondeontic selection task (P and not-Q choosers--not those choosing P and Q) are higher in cognitive ability. Despite conversational implicatures cuing the opposite response, individuals of higher cognitive ability disproportionately tend to adhere to the conjunction rule. In the 2 x 2 covariation detection experiment, it is those subjects weighting cell D more equally who are higher in cognitive ability. Finally, subjects of higher intelligence disproportionately avoid the evolutionarily justified probability matching tendency.

We believe that a useful framework for incorporating both of these data patterns is provided by two-process theories of reasoning (Epstein, 1994; Evans, 1984, 1996; Evans & Over, 1996; Sloman, 1996; Stanovich, 1999). Such a framework can encompass both the impressive record of descriptive accuracy enjoyed by a variety of evolutionary/adaptationist models as well as the fact that cognitive ability sometimes dissociates from the response deemed optimal on an adaptationist analysis.

A summary of terms used by several two-process theorists and the generic properties distinguished by several two-process views is presented in Table 3. Although the details and technical properties of these dual-process theories do not always match exactly, nevertheless there are clear family resemblances (for discussions, see Evans & Over, 1996; Gigerenzer & Regier, 1996; Sloman, 1996). In order to emphasize that his concept of these two process involved the synthesis of a prototype of the different models in the literature (rather than an attempt to defend the specific and unique properties of any one view), Stanovich (1999) adopted the generic labels System 1 and System 2.

The key differences in the properties of the two systems are listed in Table 3. System 1 processes are characterized as automatic, heuristic-based, and relatively undemanding of computational capacity. Thus, System 1 processes conjoin properties of automaticity, modularity, and heuristic processing as these constructs have been variously discussed in the literature. There is a sense in which the term System 1 is a misnomer in that it implies that it is referring to a single system. In fact, we intend the term System 1 to refer to a (probably large) set of systems in the brain (partially encapsulated modules in some views) that operate autonomously--in response to their own triggering stimuli and not under the control of a central processing structure (System 2).

System 2 conjoins the various characteristics that have been viewed as typifying controlled processing--serial, rule-based, language-biased, computationally expensive cognition. System 2 encompasses the processes of analytic intelligence that have traditionally been studied by information processing theorists trying to uncover the computational components underlying intelligence. Evans and Over (1999) argue that the function of the explicit processes of System 2 is to support hypothetical thinking. In their view, hypothetical thinking involves representing possible states of the world rather than actual states of affairs; for example, "deductive reasoning is hypothetical when its premises are not actual beliefs, but rather assumptions or suppositions....Consequential decision making consists of forecasting a number of possible future world states and representing the possible actions available....Scientific thinking is itself hypothetical



when entertaining hypotheses about the way the world might be and deducing their consequences for making predictions" (Evans & Over, 1999, p. 764). Evans and Over (1999) posit that hypothetical thought involves representing assumptions--and these necessarily must be represented as such; otherwise content would be confounded with belief. Linguistic forms such as conditionals provide a medium for such representations--and the serial manipulation of this type of representation seems to be largely a System 2 function. Language provides the discrete representational tools that fully exploit the computational power of the serial manipulations of which System 2 is capable (following Dennett, 1999, we think that System 2 processing is computationally demanding because the serial processes must be simulated by a largely parallel network).

The two systems tend to lead to different types of task construals. Construals triggered by System 1 are highly contextualized, personalized, and socialized. They are driven by considerations of relevance and are aimed at inferring intentionality by the use of conversational implicature even in situations that are devoid of conversational features (see Hilton, 1995). These properties characterize what Levinson (1995) has termed interactional intelligence--a system composed of the mechanisms that support a Gricean theory of communication that relies on intention-attribution. The primacy of these mechanisms leads to what has been termed the fundamental computational bias<sup>2</sup> in human cognition (Stanovich, 1999, in press)--the tendency or predilection toward automatic contextualization of problems. In contrast, System 2's more controlled processes serve to decontextualize and depersonalize problems. This system is more adept at representing in terms of rules and underlying principles. It can deal with problems without social content and is not dominated by the goal of attributing intentionality nor by the search for conversational relevance.

Using the distinction between System 1 and System 2 processing, Stanovich and West (2000) argued that in order to observe large cognitive ability differences in a problem situation, the two systems must strongly cue different responses. One reason that this outcome is predicted is that it is assumed that individual differences in System 1 processes (interactional intelligence) are smaller and bear little relation to individual differences in System 2 processes (analytic intelligence--see Reber, 1993, and McGeorge, Crawford, & Kelly, 1997; Reber, Walkenfeld, & Hernstadt, 1991). If the two systems cue opposite responses, rule-based System 2 will tend to differentially cue those of high analytic intelligence and this tendency will not be diluted by System 1 nondifferentially drawing subjects to the same response. For example, in nondeontic selection tasks there is ample opportunity for the two systems to cue different responses. A deductive interpretation conjoined with an exhaustive search for falsifying instances yields the response P and not-Q. This interpretation and processing style is likely associated with the rule-based System 2. In contrast, within the heuristic-analytic framework of Evans (1984, 1989, 1996), the matching response of P and Q reflects the heuristic processing of System 1 (in Evans' theory, a linguistically-cued relevance response).

The sampling of experimental results reviewed here (see Stanovich, 1999, for further examples) indicates that the alternative responses favored by the critics of the heuristics and biases literature were the choices of the subjects of lower analytic intelligence. We will explore the possibility that these alternative construals may have been triggered by heuristics that make evolutionary sense--as the evolutionary psychologists argue-- but that subjects higher in a more flexible type of analytic intelligence (and those more cognitively engaged, see Smith & Levin, 1996; Stanovich & West, 1999) are more prone to follow normative rules that maximize personal utility.

#### Evolutionary Rationality is Not Instrumental Rationality

The argument depends on the distinction between evolutionary adaptation and instrumental rationality (utility maximization given goals and beliefs). The key point is that for the latter (variously termed practical, pragmatic, or means/ends rationality), maximization is at the level of the

individual person. Adaptive optimization in the former case is at the level of the genes. In Dawkins' (1976, 1982) terms, evolutionary adaptation concerns optimization processes relevant to the so-called replicators (the genes), whereas instrumental rationality concerns utility maximization for the so-called vehicle (or interactor, to use Hull's, 1982, term), which houses the genes. Anderson (1990, 1991) emphasizes this distinction in his treatment of adaptationist models in psychology. Anderson (1990) accepts Stich's (1990; see also Skyrms, 1996) argument that evolutionary adaptation (hereafter termed evolutionary rationality) does not guarantee perfect human rationality in the instrumental sense which is focused on goals of the whole organism. As a result, a descriptive model of processing that is adaptively optimal could well deviate substantially from a normative model of instrumental rationality (Skyrms, 1996, spends an entire book demonstrating just this) because there may be different models characterizing optimization at the subpersonal and personal levels, respectively<sup>3</sup>.

A key aspect of our framework is the assumption that the goal structures that are keyed to primarily the genes' interests and the goal structures keyed primarily to the organism's interests are differentially represented in System 1 and 2 (see Reber, 1992, 1993, for a theoretical and empirical basis for this claim). It is hypothesized that the features of System 1 are designed to very closely track increases in the reproduction probability of genes. System 2, while also clearly an evolutionary product, is primarily a control system focused on the interests of the whole person. It is the primary maximizer of an individual's personal utility. Maximizing the latter will occasionally result in sacrificing genetic fitness (Barkow, 1989; Cooper, 1989; Skyrms, 1996). Because System 2 is more attuned to instrumental rationality than is System 1, System 2 will seek to fulfill the individual's goals in the minority of cases where those goals conflict with the responses triggered by System 1.

Thus, the last difference between System 1 and 2 listed in Table 3 is that System 1 instantiates short-leashed genetic goals, whereas System 2 instantiates a flexible goal hierarchy that is oriented toward maximizing goal satisfaction at the level of the whole organism. We borrow the short-/long-leash terminology by way of another metaphor used by Dawkins, (1976), Dennett (1984), and Plotkin (1988)--the "Mars Rover" analogy. Dennett (1984) describes how, when controlling a device such as a model airplane, one's sphere of control is only limited by the power of the equipment, but when the distances become large, the speed of light becomes a non-negligible factor. NASA engineers responsible for the Mars explorer vehicle knew that direct control was impossible because "the time required for a round trip signal was greater than the time available for appropriate action...Since controllers on Earth could no longer reach out and control them, they had to control themselves" (italics in original, p. 55). The NASA engineers had to move from the "short-leash" direct control, as in the model airplane case, to the "long-leash" control of the Mars explorer case where the vehicle is not given moment-by-moment instructions on how to act, but instead is given a more flexible type of intelligence plus some generic goals.

As Dawkins (1976) in his similar discussion of the Mars explorer logic in the science fiction story A for Andromeda notes, there is an analogy here to the type of control exerted by the genes when they build a brain: "The genes can only do their best in advance by building a fast executive computer for themselves....Like the chess programmer, the genes have to 'instruct' their survival machines not in specifics, but in the general strategies and tricks of the living trade....The advantage of this sort of programming is that it greatly cuts down on the number of detailed rules that have to be built into the original program." (p. 55, 57). Human brains represent, according to Dawkins (1976) "the culmination of an evolutionary trend towards the emancipation of survival machines as executive decision-makers from their ultimate masters, the genes....By dictating the way survival machines and their nervous systems are built, genes exert ultimate power over behavior. But the

moment-to-moment decisions about what to do next are taken by the nervous system. Genes are the primary policy-makers; brains are the executives....The logical conclusion to this trend, not yet reached in any species, would be for the genes to give the survival machine a single overall policy instruction: do whatever you think best to keep us alive" (p. 59-60).

This type of long-leash control that Dawkins is referring to is built in addition to (rather than as a replacement for) the short-leash genetic control mechanisms that earlier evolutionary adaptation has installed in the brain. That is, the different types of brain control that evolve do not replace earlier ones but are layered on top of them (and of course perhaps alter the earlier structures as well, see Badcock, 2000, pp. 27-29). Dennett (1996), in his short but provocative book Kinds of Minds (see also, Dennett, 1975), describes the overlapping short-leashed and long-leashed strategies embodied in our brains by labeling them as different "minds"--all lodged within the same brain in the case of humans--and all simultaneously operating to solve problems.

One key distinction between Dennett's kinds of minds is how directly the various systems code for the goals of the genes. Dennett (1996) distinguishes four different kinds of minds, the Darwinian mind, the Skinnerian mind, the Popperian mind, and Gregorian mind (see Figure 1). The minds reflect increasingly powerful mechanisms for predicting the future world. As Dennett (1991) notes, brains are anticipation machines. The four minds he proposes reflect increasing sophisticated modes of anticipation. It will be argued here that the minds, in the order listed above, also reflect decreasing degrees of direct genetic control.

The different minds control in different ways how the vehicle will react to stimuli in the environment. The Darwinian mind uses prewired reflexes and thus produces hardwired phenotypic behavioral patterns (the genes have "said" metaphorically "do this when x happens because it is best"). The Skinnerian mind uses operant conditioning to shape itself to an unpredictable environment (the genes have "said" metaphorically "learn what is best as you go along"). The Popperian mind (after the philosopher Karl Popper) can represent possibilities and test them internally before responding (the genes have "said" metaphorically "think about what is best before you do it"). The Gregorian mind (after the psychologist Richard Gregory) exploits the mental tools (see Clark, 1997) discovered by others (the genes have "said" metaphorically "imitate and use the mental tools used by others to solve problems"). In humans, all four "minds" are simultaneously operative (see Figure 1). The Darwinian and Skinnerian minds have short-leash goals installed ("when this stimulus appears, do this"). In contrast, the Popperian and Gregorian minds are characterized by long-leash goals ("operate with other agents in your environment so as to increase your longevity").

When confronted with a problem all these parts of the brain contribute potential solutions. It is variable which one will dominate. We have argued (see Stanovich, 1999; Stanovich & West, 2000) that measures of psychometric intelligence are measures of current computational capacity instantiated at the algorithmic level of System 2. This computational capacity is available to be deployed in a System 1 override function if the intentional-level goals of System 2 dictate that this will achieve goal maximization (see Figure 2). This override of System 1-triggered responses will not always be successful and thus it is predicted that on tasks where System 1 and 2 are triggering different responses, the instrumentally optimal response will be made by individuals with higher psychometric intelligence. It is precisely this that is accounting for the pattern of results we have previously reviewed. In short, we argue that high analytic intelligence may lead to task construals that track instrumental rationality; whereas the alternative construals of subjects low in analytic intelligence (and hence more dominated by System 1 processing) might be more likely to track evolutionary rationality in situations that put the two types of rationality in conflict--as is conjectured to be the case with the problems discussed previously. It is the failure to recognize the possibility of

goal conflict between the two systems that we feel plagues the treatment of human rationality in the evolutionary psychology literature.

#### Where Evolutionary Psychology Goes Wrong

Consider the bee. As a Darwinian creature, it has a goal structure as indicated in Figure 3. The area labeled A indicates the majority of cases where the replicator and vehicle goals coincide. Not flying into a brick wall serves both the interests of the replicators (the bee has a function in the hive that will facilitate replication) and of the bee itself--as a coherent organism. Of course the exact area represented by A is nothing more than a guess. The important point is that there exists a nonzero area B--a set of goals that serve only the interests of the replicators and that are antithetical to the interests of the vehicle itself<sup>4</sup>. A given bee will sacrifice itself as a vehicle if there is greater benefit to the same genes by helping other individuals (for instance, causing its own death when it loses its stinger while protecting its genetically-related hive-Queen). There are no conflicting goals in a Darwinian creature. Its goals are the genes' goals pure and simple. It is just immaterial as far as evolutionary rationality is concerned how much genetic goals overlap with vehicle goals. Perfect rationality for the bee means local fitness optimization for its genes--because for the bee the only relevant rationality is evolutionary rationality.

The error that evolutionary psychologists tend to make is that they stop right there--with an implicit assumption that evolutionary rationality is all there is; that there is no instrumental rationality (no maximization issue at the level of the whole organism--the vehicle). Evolutionary psychologists, in effect, treat humans as if they were bees. This error comes about for two reasons. First, despite emphasizing in their writings that the EEA was different from the modern environment, evolutionary psychologists have been reluctant to play out the implications of this fact. Secondly, because of their advocacy of a strictly modular view of mind and their tendency to eschew domain-general mechanisms (Cosmides & Tooby, 1992, 1994b; Tooby & Cosmides, 1992), evolutionary psychologists deemphasize the utility of the flexible goal structures of System 2 and the functions of the serial, systematically analytic processes carried out by that system. In short, evolutionary psychologists take issue with the characterization of the algorithmic level of System 2 (that it can instantiate domain-general procedures), but in doing so they miss the important function of the flexible goal structure that rides on top of the algorithmic level of System 2 (at the intentional level of analysis, see Stanovich, 1999). They are so focused on denying domain generality in algorithmic-level mechanisms (in part because they mistakenly believe that it is meant by theorists to displace the modular mind, see below) that they miss the functionality (and implications for rationality) of the goal structure at the intentional level of System 2.

With the advent of the higher-level System 2 minds (of the Popperian and Gregorian type), evolution has inserted into the architecture of the brain a flexible system that is somewhat like the ultimate long-leash goal suggested by Dawkins: "Do whatever you think best". But "best for whom?" is the critical question here. The key point is that for a creature with a flexible intelligence, long-leash goals, and a Popperian/Gregorian mind, we have the possibility of genetic optimization becoming dissociated from the vehicle's goals. For the first time, we have the possibility of a goal structure like that displayed in Figure 4. Here, although we have area A (where gene and vehicle goals coincide) and area B (goals serving the genes' interests but not the vehicle's) as before, we have a new area, C (again, the sizes of these areas in all diagrams in this chapter represent pure conjecture). In humans we have the possibility of goals that serve the vehicle's interests but not those of the genes.

Why does area C come to exist only in creatures with long-leash goals? When they started building Popperian and Gregorian minds, the genes were giving up on the strategy of coding moment-by-moment responses, and moving to a long-leash strategy that at some point was the

equivalent of saying “Things will be changing too fast out there, brain, for us to tell you exactly what to do--you just go ahead and do what you think is best given the general goals (survival, sexual reproduction) that we (the genes) have inserted”. And there is the rub. In long-leash brains, genetically coded goals can only be represented in the most general sense. There is no goal of “mate with person X at 6:57PM on Friday, June 13” but instead “have sex because it is pleasurable”. But once the goal has become this general, a potential gap has been created whereby behaviors that might serve the vehicle’s goal might not serve that of the genes. We need not go beyond the obvious example of sex with contraception--an act which serves the vehicle’s goal of pleasure without serving the genes’ goal of reproduction. What is happening here is that the flexible brain is coordinating multiple long-term goals--including its own survival and pleasure goals--and these multiple long-term goals come to overshadow its reproductive goal. From the standpoint of the genes, the human brain can sometimes be like a Mars explorer run amok. It is so busy coordinating its secondary goals (master your environment, engage in social relations with other agents, etc.) that it sometimes ignores the primary goal of replicating the genes that the secondary ones were supposed to serve.

Ironically, what from an evolutionary design point of view could be considered design defects actually make possible instrumental rationality--optimizing the utility of the person rather than the fitness of subpersonal units called genes. That is, inefficient design (from an evolutionary point of view) in effect creates the possibility of a divergence between organism-level goals and gene-level goals--which is an implication of Millikan's (1993) point that "there is no reason to suppose that the design of our desire-making systems is itself optimal. Even under optimal conditions these systems work inefficiently, directly aiming, as it were, at recognizable ends that are merely roughly correlated with the biological end that is reproduction. For example, mammals do not, in general, cease efforts at individual survival after their fertile life is over" (p. 67).

Our framework integrates the insight of the possibility of vehicle/replicator goal mismatch in the direction of the vehicle (although the possibility of area B has been acknowledged for some time, the implications of area C have been incompletely worked out) with some assumptions about the intentional-level properties of Systems 1 and 2 drawn largely from Reber (1992, 1993). The integrated framework is displayed in Figure 5 (of course, the exact size of the areas of overlap are mere guesses). Again, an assumption reflected in the Figure is that in the vast majority of real-life situations, evolutionary rationality also serves the goals of instrumental rationality. But the most important feature of the Figure is that it illustrates the asymmetries in the “interests” served by the goal distribution of the two systems. The remnants of the Darwinian creature structure (see Figure 3) are present in the System 1 brain structures of humans. Many of the goals instantiated in this system were acquired nonreflectively--they have not undergone an evaluation in terms of whether they served the person's interests. They have in fact been evaluated, but by a different set of criteria entirely: whether they enhanced the longevity and fecundity of the replicators. From the standpoint of the individual person (the vehicle) these are the dangerous goals, the ones that sacrifice the vehicle to the goals of replicators--the ones that lead the bee to sacrifice itself for its genetically related Queen. As Pinker (1997) notes, "the problem with emotions is not that they are untamed forces vestiges of our animal past; it is that they are designed to propagate copies of the genes that built them rather than to promote happiness, wisdom, or moral values. We often call an act 'emotional' when it is harmful to the social group, damaging to the actor's happiness in the long run, uncontrollable and impervious to persuasion, or a product of self-delusion. Sad to say, these outcomes are not malfunctions but precisely what we would expect from well-engineered emotions" (p. 370).

What the right side of Figure 5 (indicating the goal structure of System 2) indicates is that a bee with a Popperian/Gregorian intelligence might well decide that it would rather forgo the sacrifice! It is the reflective processes embodied in System 2 that derive the flexible long-leash goals that often have utility for the organism but thwart the goals of the genes (sex with contraception; resource use after the reproductive years have ended; etc.). These are the goals at the top of the right side of Figure 5 that overlap with vehicle interests but not genetic interests.

Failure to acknowledge the divergence of "interests" (see Footnote 4) between replicators and their vehicles is an oversight that sociobiologists were certainly guilty of (see Symons, 1992, on the "genetic fallacy") and that evolutionary psychologists are sometimes guilty of. For example, evolutionary psychologists are fond of pointing to the optimality of cognitive functioning--of showing that certain reasoning errors that cognitive psychologists have portrayed as a characteristic and problematic aspect of human reasoning (Kahneman & Tversky, 1984, 1996, 2000) have in fact a logical evolutionary explanation (Brase, Cosmides, & Tooby, 1998; Cosmides & Tooby, 1996; Gigerenzer, 1996b; Rode et al., 1999). The connotation, or unspoken assumption, is that therefore there is nothing to worry about--that since human behavior is optimal from an evolutionary standpoint, the concern for cognitive reform that has been characteristic of many cognitive psychologists (termed Meliorists by Stanovich, 1999) has been misplaced. But this sanguine attitude too readily conflates genetic optimization with goal optimization for the vehicle. To avoid the error, the different "interests" of the replicators and vehicles must be recognized--and we must keep evaluations of efficiency consistent with the entity whose optimization is at issue. The bee, as a Darwinian creature, needs no cognitive reform because it has no "interests" other than its genes' interests. Humans, with Gregorian minds, have interests as vehicles and thus might benefit from cognitive reform in situations where vehicle interests conflict with genetic interests and their Darwinian minds are siding with the latter. In such a case, it is imperative that System 2 carry out its override function (as depicted in Figure 2), and suppress the System 1 response, and substitute one more congruent with vehicle well-being.

Situations where evolutionary and instrumental rationality dissociate might well be rare, but the few occasions on which they occur might be important ones. This is because knowledge-based, technological societies often put a premium on abstraction and decontextualization, and they sometimes require that the fundamental computational bias of human cognition toward contextualization of problems (see Stanovich, 1999, in press; Stanovich & West, 2000) be overridden by System 2 processes.

Evolutionary psychologists are prone to emphasize situations where genetic goals and personal goals coincide. They are not wrong to do so, because this is most often the case. Accurately navigating around objects in the natural world was adaptive during the EEA, and it similarly serves our personal goals as we carry out our lives in the modern world. Likewise, with other evolutionary adaptations: It is a marvel that humans are exquisite frequency detectors (Hasher & Zacks, 1979), that they infer intentionality with almost supernatural ease (Levinson, 1995), and that they acquire a complex language code from impoverished input (Pinker, 1994). All of these mechanisms several personal goal fulfillment in the modern world. But none of this means that the overlap is necessarily one-hundred percent.

Unfortunately, the modern world tends to create situations where some of the default values of evolutionarily adapted cognitive systems are not optimal. Modern technological societies continually spawn situations where humans must decontextualize information--where they must deal abstractly (Adler, 1984) and in a depersonalized manner with information rather than in the context-specific way assumed by proponents of the massive modularity thesis (Samuels, 1998). Such situations require the active suppression of the personalizing and contextualizing styles that

characterize the fundamental computational biases (Stanovich, 1999, in press). Such biases directly conflict with the demands for decontextualization that a highly bureaucratized society puts on its citizens. Indeed, this is often why schools have to explicitly teach such skills of cognitive decontextualization. Increasingly, modern society is demanding such skills (Dickens & Flynn, 2001; Frank & Cook, 1995; Gottfredson, 1997; Hunt, 1995, 1999)--and in some cases it is rendering economically superfluous anyone who does not have them (Bronfenbrenner, McClelland, Wethington, Moen, & Ceci, 1996; Frank & Cook, 1995). For example, many aspects of the contemporary legal system put a premium on detaching prior belief and world knowledge from the process of evidence evaluation. There has been understandable vexation at odd jury verdicts rendered because of jury theories and narratives concocted during deliberations that had nothing to do with the evidence but instead that were based on background knowledge and personal experience.

The need to decontextualize also characterizes many work settings in contemporary society. Consider the common admonition in the retail service sector of "the customer is always right". This admonition is often interpreted to include even instances where customers unleash unwarranted verbal assaults which are astonishingly vitriolic. The service worker is supposed to remain polite and helpful under this onslaught, despite the fact that such emotional social stimuli are no doubt triggering evolutionarily instantiated modules of self defense and emotional reaction. All of this emotion, all of these personalized attributions--all fundamental computational biases--must be set aside by the service worker and instead an abstract rule that "the customer is always right" must be invoked in this special, socially-constructed domain of the market-based transaction. The worker must realize that he/she is not in an actual social interaction with this person (which if true, might call for socking them in the nose!), but in a special, indeed "unnatural" realm where different rules apply.

Concerns about the real-world implications of the failure to engage in necessary cognitive abstraction (see Adler, 1984) were what led Luria (1976) to warn against minimizing the importance of decontextualizing thinking styles. In discussing the syllogism, he notes that "a considerable proportion of our intellectual operations involve such verbal and logical systems; they comprise the basic network of codes along which the connections in discursive human thought are channeled" (p. 101). Einhorn and Hogarth (1981) highlight the importance of decontextualized environments in their discussion of the optimistic and pessimistic views of the cognitive biases revealed in laboratory experimentation. Einhorn and Hogarth (1981) note that "the most optimistic asserts that biases are limited to laboratory situations which are unrepresentative of the natural ecology" (p. 82), but they go on to caution that "in a rapidly changing world it is unclear what the relevant natural ecology will be. Thus, although the laboratory may be an unfamiliar environment, lack of ability to perform well in unfamiliar situations takes on added importance" (p. 82).

Critics of the abstract content of most laboratory tasks and standardized tests have been misguided on this very point. Evolutionary psychologists have singularly failed to understand the implications of Einhorn and Hogarth's (1981) warning. They regularly bemoan the "abstract" problems and tasks in the heuristics and biases literature and imply that since these tasks are not like "real life" we need not worry that people do poorly on them. The issue is that, ironically, the argument that the laboratory tasks and tests are not like "real life" is becoming less and less true. "Life," in fact, is becoming more like the tests! Try using an international ATM machine with which you are unfamiliar; or try arguing with your HMO about a disallowed medical procedure. In such circumstances, we invariably find out that our personal experience, our emotional responses, our stimulus-triggered intuitions about social justice--all are worthless. All are for naught when talking over the phone to the representative looking at a computer screen displaying a spreadsheet with a hierarchy of branching choices and conditions to be fulfilled. The social context, the idiosyncrasies

of individual experience, the personal narrative--all are abstracted away as the representatives of modernist technological-based services attempt to "apply the rules".

Modern mass communication technicians have become quite skilled at implying certain conclusions without actually stating those conclusions (for fear of lawsuits, bad publicity, etc.). Advertisements rely on the fundamental computational bias (particularly its enthymematic processing feature) to fill in the missing information. Margolis (1987; see Margolis, 1996) warns of the ubiquitousness of this situation in modern society: "We can encounter cases where the issue is both out-of-scale with everyday life experience and contains important novelties, so that habitual responses can be highly inappropriate responses. The opportunity for unrecognized contextual effects akin to the scenario effects...[demonstrated in the laboratory] can be something much more than an odd quirk that shows up in some contrived situation" (p. 168).

Evolutionary psychologists have argued that some problems can be more efficiently solved if represented to coincide with how various brain modules represent information<sup>5</sup>. Nevertheless, they often seem to ignore the fact that the world will not always let us deal with representations that are optimally suited to our evolutionarily designed cognitive mechanisms. For example, in a series of elegant experiments, Gigerenzer, Hoffrage, and Kleinbolting (1991) have shown how at least part of the overconfidence effect in knowledge calibration studies is due to the unrepresentative stimuli used in such experiments--stimuli that do not match the participants' stored cue validities which are optimally tuned to the environment. But there are many instances in real-life when we are suddenly placed in environments where the cue validities have changed. Metacognitive awareness of such situations and strategies for suppressing incorrect confidence judgments generated by automatic responses to cues will be crucial here. Every high school musician who aspires to a career in music has to recalibrate when they arrive at university and see large numbers of talented musicians for the first time. If they persist in their old confidence judgments they may not change majors when they should. Many real-life situations where accomplishment yields a new environment with even more stringent performance requirements share this logic. Each time we "ratchet up" in the competitive environment of a capitalist economy (Frank & Cook, 1995) we are in a situation just like the overconfidence knowledge calibration experiments with their unrepresentative materials. It is important to have learned strategies that will temper one's overconfidence in such situations (Koriat, Lichtenstein, & Fischhoff, 1980).

#### How Evolutionary Psychology Goes Wrong

Dawkins (1976) notes that there is an "uneasy tension....between gene and individual body as fundamental agent of life" (p. 234). Many evolutionary psychologists have missed this essential tension by focusing on parallels between the evolutionary optimization of humans and other animals. But humans are vehicles with interests beyond those of their genes' replication. Humans aspire to be more than mere survival machines serving the "ends" of their genes (which are replication pure and simple). Only humans really turn the tables (or at least have the potential to) by occasionally ignoring the interests of the genes in order to further the interests of the vehicle. Evolutionary psychology--for all its important insights about human behavior--has failed to develop this profound insight. By failing to highlight the interests of the vehicle in discussions of optimal cognitive functioning, evolutionary psychology has colluded with the genes in delivering their most sophisticated vehicle (human beings) over to them, as if this vehicle--like the bee--had no interests other than replication.

As argued above, evolutionary psychologists background the evolutionary/instrumental rationality distinction because many are: 1. wedded to a cognitive architecture that displays massive modularity; 2. as a result, they eschew domain general System 2 mechanisms; 3. they conjoin these



two theoretical assumptions with a tendency to ignore the implications of mismatches between the EEA and the cognitive requirements of technological societies.

To the extent that modern society increasingly requires the fundamental computational biases to be overridden, then dissociations between evolutionary and individual rationality will become more common--and System 2 overrides will be more essential to personal well being. Cosmides and Tooby (1996) argue that "in the modern world, we are awash in numerically expressed statistical information. But our hominid ancestors did not have access to the modern accumulation which has produced, for the first time in human history, reliable, numerically expressed statistical information about the world beyond individual experience. Reliable numerical statements about single event probabilities were rare or nonexistent in the Pleistocene" (p. 15). "It is easy to forget that our hominid ancestors did not have access to the modern system of socially organized data collection, error checking, and information accumulation....In ancestral environments, the only external database available from which to reason inductively was one's own observations" (Brase, Cosmides, & Tooby, 1998, p. 5).

Although this may be entirely correct (but see Footnote 5), let us carry through with the implications of this point. We are living in a technological society where we must: decide which health maintenance organization to join based on just such statistics; figure out whether to invest in an individual retirement account; decide what type of mortgage to purchase; figure out what type of deductible to get on our auto insurance; decide whether to trade in a car or sell it ourselves; decide whether to lease or to buy; think about how to apportion our retirement funds; and decide whether we would save money by joining a book club--to simply list a random set of the plethora of modern-day decisions and choices. And we must make all of these decisions based on information represented in a manner for which our brains may not be adapted (in none of these cases have we coded individual frequency information from our own personal experience). In order to reason normatively in all of these domains (in order to maximize our personal utility) we are going to have to deal with probabilistic information represented in nonfrequentistic terms--in representations that the evolutionary psychologists have argued are different from our adapted algorithms for dealing with frequency information (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995).

Consider the work of Brase et al. (1998), who improved performance on a difficult probability problem (Bar-Hillel & Falk, 1982; Falk, 1992; Granberg, 1995) by presenting the information as frequencies and in terms of whole objects--both alterations designed to better fit the posited frequency-computation systems of the brain. In response to a query about why the adequate performance observed was not even higher given that our brains contain such well-designed frequency-computation systems, Brase et al. (1998) replied that "in our view it is remarkable that they work on paper-and pencil problems at all. A natural sampling system is designed to operate on actual events" (p. 13). The problem is that in a symbol-oriented postindustrial society, we are presented with paper-and pencil problems all the time, and much of what we know about the world comes not from the perception of actual events but from abstract information preprocessed, prepackaged, and condensed into symbolic codes such as probabilities, percentages, tables, and graphs (the voluminous statistical information routinely presented in USA Today comes to mind).

What we are attempting to combat here is a connotation implicit in some discussions of findings in evolutionary psychology and indeed in the situated cognition literature as well (see Anderson et al., 1996) that there is nothing to be gained from being able to understand a formal rule at an abstract level (the conjunction rule of probability, etc.)--and no advantage in flexibly overriding the fundamental computational biases. We can see the tendency of evolutionary psychologists to fall into this trap in the following statement:

“In actuality, adaptationist approaches offer the explanation for why the psychic unity of humankind is genuine and not just an ideological fiction; for why it applies in a privileged way to the most significant, global, functional, and complexly organized dimensions of our architecture; and for why the differences among humans that are caused by genetic variability that geneticists have found are so overwhelmingly peripheralized into architecturally minor and functionally superficial properties” (Tooby & Cosmides, 1992, p. 79)

This statement provides an example of how and why evolutionary psychology goes off the rails. Let us see what is in some of that "genetic variability that geneticists have found" and let us ask ourselves, seriously, whether it does reflect "functionally superficial properties".

Well, for starters, some of that "genetic variability that geneticists have found" is in general intelligence (g)--which virtually everyone who has looked at the evidence agrees is at least 40-50% heritable (Deary, 2000; Grigorenko, 1999; Neisser et al., 1996; Plomin & Petrill, 1997). Is g a "functionally superficial" individual difference property of human cognition? No responsible psychologist thinks so. It is, indeed, the single most potent psychological predictor of human behavior in both laboratory and real-life contexts that has ever been identified (Lubinski, 2000; Lubinski & Humphreys, 1997). It is a predictor of real-world outcomes that are critically important to the maximization of personal utility (to instrumental rationality) in a modern technological society. Objective measures of the requirements for cognitive abstraction have been increasing across most job categories in technological societies throughout the past several decades (Gottfredson, 1997). This is why measures of the ability to deal with abstraction such as g remain the best employment predictor and the best earnings predictor in postindustrial societies (Brody, 1997; Gottfredson, 1997; Hunt, 1995). The psychometric literature contains numerous indications that cognitive ability is correlated with the avoidance of harmful behaviors and with success in employment settings, as well as social status attainment (MacDonald & Geary, 2000), independent of level of education (Brody, 1997; Gottfredson, 1997; Hunt, 1995; Lubinski & Humphreys, 1997).

We view individual differences in g as indicating differences in the current computational capacity of the algorithmic level of System 2. It is critically related to the override function of System 2 discussed above--the override function necessary to trump the fundamental computational biases of System 1 when they lead to a response that is antithetical to the interests of the vehicle.

Our algorithmic-level understanding of System 2 borrows from Dennett (1991), who conceives of System 2 as a serial von Neumann computer simulated by the massively parallel computational network of the brain. It is language-based, rule-based, and at least more logic-based than System 1 (Evans & Over, 1996, 1997)--and is the focus of our awareness (it is the system we use to construct a model of the self). As mentioned previously, Evans and Over (1999) discuss the fundamental importance of System 2 as the mechanism that supports hypothetical thinking. In contrast to the holistic/associative nature of System 1, System 2 is analytic in operation, and it is demanding in terms of computational capacity. We view general intelligence to encompass two fundamental classes of property (that perhaps map into the fluid/crystallized distinction from the Horn/Cattell model, Horn, 1982; Horn & Cattell, 1967). First, there is the computational power of the parallel network to sustain the serial simulation<sup>6</sup> (this is probably closer to fluid intelligence in the Horn/Cattell model of intelligence). The second major factor is the power of the cultural tools used during serial simulation--the Gregorian mind in Dennett's (1991) Tower of Intellect model (individual differences in this factor might relate to variance in crystallized intelligence in the Horn/Cattell model).

Intelligence is not the only type of "genetic variability that geneticists have found" that is manifestly not "functionally superficial". Similar stories could be told about many personality

variables (reflective of intentional-level cognitive variability, see Stanovich, 1999) that have been shown to be heritable but also important predictors of behavioral outcomes (see p. 394 of Buss, 1999; Matthews & Deary, 1998). Indeed, this stance by some evolutionary psychologists against heritable cognitive traits with demonstrable linkages to important real-world behaviors has become an embarrassment even to some evolutionary theorists. Buss (1999) characterizes the view of Tooby and Cosmides as the notion that "heritable individual differences are to species-typical adaptations, in this view, as differences in the colors of the wires in a car engine to the engine's functional working components" (p. 394), and points to some of the same embarrassing empirical facts noted above. For example, heritable personality traits such as conscientiousness and impulsivity have been related to important life goals such as work, status attainment, mortality, and faithfulness in partnerships. Buss' (1999) alternative interpretation is in terms of genetic concepts such as frequency-dependent selection. But whether or not one accepts such explanations, the point is that many evolutionary theorists have mistakenly downplayed cognitive constructs that are heritable (intelligence, personality dimensions, thinking styles) and that have demonstrated empirical relationships to behaviors that relate to utility maximization for the individual (job success, personal injury, success in relationships, substance abuse).

Despite Buss' (1999) more nuanced position on individual differences, other influential evolutionary psychologists repeat like a mantra the view that any psychological processes with genetic variation lack any importance (and presumably lack any relevance for rationality, since this is obviously important to the vehicle):

"Human genetic variation...is overwhelmingly sequestered into functionally superficial biochemical differences, leaving our complex functional design universal and species typical" (Tooby & Cosmides, 1992, p. 25)

"humans share a complex, species typical and species-specific architecture of adaptations, however much variation there might be in minor, superficial, nonfunctional traits" (Tooby & Cosmides, 1992, p. 38)

One boggles at general intelligence--one of the most potent psychological predictors of life outcomes--being termed "nonfunctional". But then one realizes what is motivating these statements--a focus on the gene. Even if one buys the massive-modularity-of-adaptations line of the evolutionary psychologist and views general intelligence as some kind of spandrel or byproduct<sup>7</sup>, from the standpoint of the vehicle's interests, it is certainly not nonfunctional. Only a focus on the subpersonal replicators would spawn such a statement--one which backgrounds important cognitive traits such as intelligence and conscientiousness (Lubinski, 2000; Matthews & Deary, 1998). As soon as one focuses on the organismic level of optimization rather than genetic optimization, the "nonfunctional" traits spring to the foreground as the System 2 algorithmic-level (intelligence) and intentional-level (conscientiousness, openness) constructs that explain individual differences in attaining one's goals (Baron, 1993, 1994; Stanovich, 1999).

The downplaying of the importance of a heritable cognitive indicator such as general intelligence by evolutionary psychologists often results from their tendency to caricature cognitive theories that stress a domain-general mechanism (like the type of analytic processing hypothesized for System 2 by many dual-process theorists). The evolutionary theorists purport to dispute theories which view the evolutionary history of human cognition as the replacement of context-dependent modules with context-independent general intelligence mechanisms. For example, in attacking the so-called Standard Social Science Model (SSSM), Tooby and Cosmides (1992) argue that this default social science model "views an absence of content-specific structure as a precondition for richly flexible behavior" (p. 113). Their view of the standard cognitive model in psychology is that general processing mechanisms replace domain-specific ones. Actually, as the long history of dual-

process models attests (see Table 3), the standard view in psychology is much more similar to the evolutionary psychology view than Tooby and Cosmides want to admit. As in Dennett's (1996) "Tower of Intellect" model, all of the two-process views listed in Table 3 conceive of analytic processes developing in conjunction with domain-specific mechanisms (see Mithen, 1996). Analytic processing mechanisms develop in addition to System 1 modules--they do not replace them.

Evolutionary psychologists also tend to misleadingly minimize the consequences of mismatches between the EEA and the modern environment. Tooby and Cosmides (1992, p. 72) approvingly paraphrase Shepard's (1987) point that evolution insures a mesh between the principles of the mind and the regularities of the world. But this "mesh" concerns regularities in the EEA, not in the modern world--with its unnatural requirements for decontextualization (requirements that do not "mesh" with the fundamental computational biases toward comprehensive contextualization of situations). One page later in their chapter, Tooby and Cosmides (1992) reveal the characteristic bias of evolutionary psychologists--the belief that "often, but not always, the ancestral world will be similar to the modern world (e.g., the properties of light and the laws of optics have not changed)" (p. 73). We largely agree. However, although the laws of optics haven't changed, the type of one-shot, abstract, probabilistic, and symbolically represented decision situations a modern human being must deal with are certainly unprecedented in human history. Think of insurance decisions, retirement decisions, investment decisions, home buying decisions, relocation decisions, and school choices for children. These are not the highly practiced, frequency coded, time pressured, recognition-based situations (Klein, 1998) where evolutionary heuristics work best. Instead, these are all the type of situations that invoke just the type of representativeness, availability, sunk cost, confirmation bias, overconfidence, and other effects that the heuristics and biases researchers have studied (see the many real-life examples in Kahneman & Tversky, 2000). We can walk and navigate among objects as well as we ever did, but no evolutionary mechanism has sculpted my brain to estimate the deductible I need on my insurance or how I should evaluate the cost of a disability policy to cover salary loss.

Tooby and Cosmides (1992) seem to take a completely one-sided message from the potential mismatch between the EEA and modern conditions--when in fact the mismatch has more than one implication. Using the example of how our color constancy mechanisms fail under modern sodium vapor lamps, they warn that "attempting to understand color constancy mechanisms under such unnatural illumination would have been a major impediment to progress" (p. 73)--a fair enough point. But our purpose here is to stress a different corollary point that one might have drawn. The point is that if the modern world were structured such that making color judgments under sodium lights was critical to our well-being, then this would be troublesome for us because our evolutionary mechanisms have not naturally equipped us for this. One might be given impetus to search for a cultural invention that would circumvent this defect (relative to the modern world, not the EEA) in our cognitive apparatus.

We argue that humans in the modern world are in just this situation vis-à-vis the mechanisms needed for fully rational action in industrial and bureaucratized societies. The processing of probabilistic information provides a case in point. We argued above that it is critical to many tasks faced by a full participant in a First World society. Of course, the heuristics and biases literature is full of demonstrations of the problems that people have in dealing with probabilistic information. Evolutionary psychologists have done important work that suggests that the human cognitive apparatus may be more adapted to dealing with frequencies than with probabilities (Brase, Cosmides, & Tooby, 1998; Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995; but see Footnote 5 and Over, this volume). For example, it has been found that when tasks such as the Linda problem, knowledge calibration tasks, and base rate tasks are revised in terms of estimating the

frequency of categories rather than judging probabilities that performance is improved (see Cosmides & Tooby, 1996; Fiedler, 1988; Gigerenzer, 1991, 1993; Gigerenzer & Hoffrage, 1995; Tversky & Kahneman, 1983; but see Mellers, Hertwig, & Kahneman, 2001). As useful as this research has been (and indeed it can usefully be adapted to tell us how to more understandably present probabilistic information in real-life settings, see Gigerenzer, Hoffrage, & Ebert, 1998), it will not remove the necessity of being able to process probabilistic information when it is presented in the real world.

The evolutionary psychologists and ecological rationality theorists are sometimes guilty of implying just this--that if the human cognitive apparatus can be shown to have been adapted during evolution to some other representation (other than that required for a problem in modern society) then somehow it has been shown that there really is no cognitive problem. For example, in the titles and subheadings of several papers on frequency representations, Gigerenzer (1991, 1993, Gigerenzer et al., 1991) has used the phrasing "how to make cognitive illusions disappear." This is a strange way to phrase things, because the original illusion has of course not "disappeared." As Kahneman and Tversky (1996) note, the Muller-Lyer illusion is removed when the two figures are embedded in a rectangular frame, but this does not mean that the original illusion has "disappeared" in this demonstration (see also Samuels, Stich, & Tremoulet, 1999). The cognitive illusions in their original form still remain (although their explanation has perhaps been clarified by the different performance obtained in the frequency version), and the situations (real-life or otherwise) in which these illusions occur have not been eliminated. Banks, insurance companies, medical personal, and many other institutions of modern society are still exchanging information using linguistic terms like probability and applying that term to singular events. My physician has on occasion given me a migraine prescription (Imitrex, for instance) with the assurance that he is 90% certain it will work in my case. As many Bayesian investigators in the calibration literature have pointed out, it is likely that I would be quite upset if I found out that for 50% of his patients so advised the medication did not work.

Drawing on Sperber's (1994) distinction between the actual domain and the proper domain (the modern environment versus the EEA), Samuels, Stich, and Tremoulet (1999) argue that "we suspect that those Panglossian-inclined theorists who describe Darwinian modules as 'elegant machines' are tacitly assuming that normative evaluation should be relativized to the proper domain, while those who offer a bleaker assessment of human rationality are tacitly relativizing their evaluations to the actual domain, which, in the modern world, contains a vast array of information-processing challenges that are quite different from anything our Pleistocene ancestors had to confront" (p. 114; see also Davies, 1996; Looren de Jong & van der Steen, 1998). Perhaps both groups are guilty of some disproportionate emphasis here. Evolutionary theorists err by emphasizing the proper domain so much that they seem to forget about the actual domain, and the Meliorists in the heuristics and biases camp are so prone to emphasize the errors occurring in the actual domain that they fail to acknowledge that humans really are optimally designed for a proper domain.

Buss (1999) shows the former tendency when he asks the question: "If humans are so riddled with cognitive mechanisms that commonly cause errors and biases, how can they routinely solve complex problems that surpass any system that can be developed artificially?" (p. 378)--and answers it by quoting an unpublished paper by Tooby and Cosmides where the argument is made that our criteria for recognizing sophisticated performance "have been parochial" (p. 378). Buss seems to be calling our natural privileging of the present environment--the one we actually have to operate in--unnecessarily parochial. The devaluing of the actual decontextualized environment in which we must operate in modern technological society continues as Buss (1999) repeatedly minimizes rational thinking errors by pointing out that they occur in "artificial or novel" (p. 378) situations.

The latter of course seems damning to his own argument (that these errors are trivial) because novel symbolic situations are exactly what bureaucratically immersed workers and citizens in technological societies must constantly deal with.

With respect to the "artificial situations" criticism, Buss (1999) trots out the old sodium vapor lamps example, saying that the experiments have used "artificial, evolutionarily unprecedented experimental stimuli analogous to sodium vapor lamps" (p. 379). Like Tooby and Cosmides (1992), Buss (1999) takes exactly the wrong message from the potential mismatch between EEA and modern conditions. It is a very serious worry that we are essentially in situations where we must work under sodium vapor lamps! The cognitive equivalent of the sodium vapor lamps are: the probabilities we must deal with; the causation we must infer from knowledge of what might have happened; the vivid advertising examples we must ignore; the unrepresentative sample we must disregard; the favored hypothesis we must not privilege; the rule we must follow that dictates we ignore a personal relationship; the narrative we must set aside because it does not square with the facts; the pattern that we must infer is not there because we know a randomizing device is involved; the sunk cost that must not affect our judgment; the judge's instructions we must follow despite their conflict with common sense; the contract we must honor despite its negative affects on a relative; the professional decision we must make because we know it is beneficial in the aggregate even if unclear in this case. These are all the "sodium vapor lamps" that modern society presents to our cognitive apparatus--and if evolution has not prepared us to deal with them so much the worse for our rational behavior in the modern world (Stanovich, 1999, in press). Luckily, the Gregorian tools of rational thought, running as virtual machines on our System 2 serial simulator are there to help us in situations such as this.

#### The Slippery Notion of Ecological Rationality

Many of the foregoing arguments about matching conceptions of rationality to the level of the entity being optimized apply to the concept of ecological rationality as well, and the work of those who have championed this concept. But the concept itself is not straightforward. A textual analysis of its usage reveals that it is a slippery concept indeed.

Typical of these confusions is a statement at the end of a volume summarizing the work of one of the laboratories responsible for popularizing the term: "Ultimately, ecological rationality depends on decision making that furthers an organism's adaptive goals in the physical or social environment" (Gigerenzer & Todd, 1999, p. 364). In statements such as this, we see a double ambiguity which makes the ecological rationality term devilishly difficult to pin down and hence to evaluate. First, the phrasing "organism's adaptive goals" makes it unclear what level of analysis we are talking about. The word adaptive suggests we are talking, in the technical sense, about evolutionary (hence genetic) goals--that ecological rational is about how organisms are optimized to achieve the goals of their genes. On the other hand, one looks at the same phrase and wonders whether the word "organism" is not key here--that the word adaptive is actually being used more colloquially--and that we are to put a stress on it (as in the "organism's adaptive goals") and view the ecological rational concept as akin to instrumental rationality (as maximizing the vehicle's utility). This ambiguity in the "organism's adaptive goals" phrase introduces a second ambiguity into the second part of the quote. Because we are unsure whether adaptive goals refer to the genes' goals or the vehicle's goals it becomes unclear whether "the physical or social environment" is meant to refer to the current (modern) environment or to the EEA.

These two interpretations of the phrase "organism's adaptive goals" slip in and out throughout the 400 pages of the book by Gigerenzer and Todd (1999) and the Adaptive Behavior and Cognition Group at the Max Planck Institute in Berlin (see also, Todd & Gigerenzer, 2000). On page 335 of the book, they draw tight links between work in behavioral ecology and ecological rationality

("ecological rationality is what behavioral ecology is all about") which suggest that genetic fitness maximization is, likewise, what ecological rationality is all about. This view is reinforced in other parts of the book where we are told that "the collection of specialized cognitive mechanisms that evolution has built into the human mind for specific domains of inference and reasoning" (p. 30) include the fast and frugal heuristics that are the focus of a dozen chapters in the Gigerenzer and Todd (1999) book. Elsewhere, we are told that "evolution would seize upon informative environmental dependencies such as this one and exploit them with specific heuristics if they would give a decision-making organism an adaptive edge" (Gigerenzer & Todd, 1999, p. 19).

Buss (1999), discussing Tooby and Cosmides' use of the term ecological rationality in an unpublished manuscript, likewise links the term to fitness in the evolutionary sense: "Over evolutionary time, the human environment has had certain statistical regularities....These statistical regularities are called ecological structure. Ecological rationality consists of evolved mechanisms containing design features that utilize this ecological structure to facilitate adaptive problem solving" (p. 378). Thus, Buss (1999), as do Gigerenzer and Todd (1999), seems to clearly imply that what ecological rationality is designed to optimize is genetic fitness in the EEA.

After all this emphasis on evolution being the superstructure on which ecological rationality sits, it is a surprise to hear Todd, Fiddick, and Krauss (2000) reply to a critique of Over (2000) by saying that although evolutionary psychology is grounded in ecological rationality the converse is not true. Unlike Buss' (1999) emphasis on the human environment "over evolutionary time," Todd et al. (2000), in a complete theoretical reversal from the quotes above, instead assert that ecological rationality "encompasses decision making in present environments without privileging problems with fitness consequences" (p. 379).

In short, there is considerable inconsistency in the writings of the ecological rationality theorists about whether ecological rationality is optimization to the EEA or to the modern environment--in short, about whether ecological rationality is maximizing for the genes or for the vehicle. The term needs to be explicitly identified with what I have termed evolutionary rationality (Stanovich, 1999) or with what has traditionally been viewed as instrumental rationality--rationality for the whole organism (see Over, 2000). Instead, these theorists seem to slip back and forth in their usage--adopting whichever stance is most convenient for the argument being made. Ecological rationality theorists seem to want the imprimatur of evolution (and the biological plausibility that evolutionary adaptation provides) without accepting other inconvenient implications of evolutionary explanations. One implication is that we cannot assume that System 1 heuristics (adapted for the EEA) are optimal for achieving rationality in the modern world (Stanovich, 1999, in press). Many important decisions in life are nearly "one shot" affairs (job offers, pension decisions, investing decisions, housing decisions, marriage decisions, reproductive decisions, etc.). Some of these decisions were not present at all in the EEA, and we have had no time nor learning trials to acquire extensive personal frequency information about them. Instead we need to make certain logical and probabilistic inferences using various rules of inference, and most importantly, we must decouple myriad sources of information that our autonomously functioning modules might be detecting and feeding into the decision ("no, the likability of this salesperson should not be a factor in my deciding on this \$25,000 car").

In fact, some of the System 1 heuristics that are in place might seriously subvert instrumental goals in a modern technological society. For example, one chapter in the Gigerenzer and Todd (1999) book (see Goldstein & Gigerenzer, 1999) is devoted to the so-called recognition heuristic--the chapter subheading being "How Ignorance Makes us Smart". The idea behind such "ignorance-based decision making" as they call it, is that the fact that some items of a subset are unknown can be exploited to aid decision making. The yes/no recognition response can be used as a frequency

estimation cue. With ingenious simulations, Goldstein and Gigerenzer (1999) demonstrate how certain information environments can lead to such things as less-is-more effects: where those who know less about an environment can display more inferential accuracy in it.

One is certainly convinced after reading material like this that the recognition heuristic is certainly efficacious in some situations. But then one immediately begins to worry when we ponder how it relates to a market environment specifically designed to exploit it. If the senior author of this chapter left his home--located in the middle of the financial and industrial capital of a First World country--and relied solely on the recognition heuristic, he could easily be led to:

1. buy a \$3 coffee when in fact a \$1.25 one would satisfy him perfectly
2. eat in a single snack the number of fat grams he should have in an entire day
3. pay the highest bank fees (because the highest fees are charged by the most recognized banks in Canada)
4. incur credit card debt rather than pay cash
5. buy a mutual fund with a 6% sales charge rather than a no-load fund

None of these behaviors serves his long-term instrumental goals at all--none of them help get him toward his reflectively acquired aspirations. Yet the recognition heuristic triggers these and dozens more that will trip him up while trying to make his way through the maze of modern society.

The proponents of ecological rationality refuse to acknowledge this downside of the ecological approach. For example, Borges, Goldstein, Ortman, and Gigerenzer (1999) take curious pride in the finding that a portfolio of stocks recognized by a group of Munich pedestrians beat two benchmark mutual funds of a six-month period during the mid-1990s. This finding is of course a pure artifact of an extraordinary short period in the 1990s when large capitalization stocks outperformed small capitalization stocks (Over, 2000). The adaptive heuristics investigated by Borges et al. (1999) haven't repealed the basic laws of investing. Risk is still related to reward, and over longer time periods small capitalization stocks outperformed their less-risky large capitalization counterparts. Obviously, the Munich pedestrians had better recognition for the large companies--precisely those enjoying a good run in that particular six-month period (which is of course too short for various risk/reward relationships to show themselves).

Borges et al. (1999) might have alternatively focused on another well-known finding in the domain of personal finance discussed by Bazerman (2001)--that consumers of financial services overwhelmingly purchase high-cost products that underperform in terms of investment return the low-cost strategies recommended by true experts (e.g., dollar-cost averaging into no-load index mutual funds). The reason is, of course, that the high-cost fee-based products and services are the ones with high immediate recognizability in the marketplace, whereas the low-cost strategies must be sought out in financial and consumer publications.

One leaves the writings of the ecological rationality theorists--whatever they take the term to mean--thinking that they, like the evolutionary psychologists, are being much too sanguine about the ability of System 1 processes to achieve instrumental rationality--to optimize the broad and reflective goals of System 2 rather than the short-leash evolutionary goals of System 1.

#### The Unacknowledged Importance of the Meme

Why do the evolutionary psychologists and ecological theorists show such a tendency to misconstrue human rationality--to fail to distinguish the interests of the replicators from the interests of the vehicle? Our conjecture is that the error follows from a particular overgeneralization that is encouraged by the evolutionary psychologists' attack on what Tooby and Cosmides (1992) call the Standard Social Science Model (SSSM). These evolutionary psychologists believe that the SSSM stands in the way of a fully articulated evolutionary psychology based on the assumption that the human brain is composed of content-specific information-processing mechanisms that have evolved



as adaptations. Instead, the SSSM has become the default model of most social scientists in their view, and the SSSM contains many misleading assumptions about human cognition, including the assumption that the human mind is structured as a general, unbiased learning mechanism. Among the many assumptions of the SSSM that are wrong according to Tooby and Cosmides (1992), are its assumptions about culture. Tooby and Cosmides (1992) feel that the default assumptions of the SSSM of most social scientists are that "the individual is the more or less passive recipient of her culture and is the product of that culture" (p. 32) and that "human nature is an empty vessel, waiting to be filled by social processes" (p. 29).

The idea of free-floating cultural products--those totally unconditioned by and unadapted to evolved mental mechanisms (what Tooby & Cosmides, 1992 call epidemiological culture<sup>8</sup>)--is an anathema to many evolutionary psychologists. In fact, Tooby and Cosmides (1992) labor hard to convince the reader that even epidemiological culture "is also shaped by the details of our evolved psychological organization" (p. 119). To use Dawkins' (1976) term, evolutionary psychologists are hostile to the concept of the meme. But because human rationality is in large part a memetic product--a set of cultural tools for the Gregorian mind--evolutionary psychologists are prone to miss or denigrate its importance.

Dawkins' (1976) introduced the term meme to refer to a unit of cultural information that is meant to be understood in rough (rather than one-to-one) analogy to a gene. Blackmore (1999) defines the meme as the instructions for behaviors and communications that can be learned by imitation broadly defined (in the sense of copying by the use of language, memory, or any other mechanism) and that can be stored in brains (or other storage devices). Collectively, genes contain the instructions for building the bodies that carry them. Collectively, memes build the culture that transmits them. Like the gene, the meme is a true replicator in the sense of the distinction made in theoretical biology between replicators and interactors (Dawkins, 1976; Hull, 1988; Sterelny, 2001; Sterelny & Griffiths, 1999; Williams, 1985, 1992). Replicators are entities that pass on their structure relatively intact after copying and interactors or vehicles are "those entities that interact as cohesive wholes with their environments in such a way as to make replication differential" (Hull, 1988 p. 27).

The key idea in memetic theory is that the meme is a true selfish replicator in the same sense that a gene is--it acts only in its own "interests". The anthropomorphic language about genes and memes having interests is shorthand for the complicated description of what is actually the case: that genes/memes that perform function X make more copies of themselves, copy with greater fidelity, or have greater longevity--and hence will leave more copies in future generations. Or, as Blackmore (1999) states it, "the shorthand 'genes want X' can always be spelled out as 'genes that do X are more likely to be passed on.' This is the only power they have--replicator power. And it is in this sense that they are selfish" (p. 5).

Memes are independent replicators. They do not necessarily exist in order to help the vehicle (those who hold the belief) they exist because through memetic evolution they have displayed the best fecundity, longevity, and copying fidelity--the defining characteristics of successful replicators. The fundamental insight triggered by memetic theory is that a meme may display fecundity and longevity without necessarily being true or helping the vehicle (the human being holding the belief) in any way. Memetic theorists often use the example of a chain letter. Here is a meme: "If you do not pass on this message to five people you will experience misfortune". This is an instruction for a behavior that can be copied and stored in brains. It survives because of its own self-replicating properties (it is neither good for the genes or for the vehicle). Dawkins (1976) argues that "what we have not previously considered is that a cultural trait may have evolved in the way it has, simply because it is advantageous to itself" (p. 27). Memetic theory asks instead what is it about certain

memes that leads them to collect many "hosts" for themselves. Indeed, this type of language was suggested by Dawkins (1976) himself who, paraphrasing Nick Humphreys, said that "when you plant a fertile meme in my mind you literally parasitize my brain, turning it into a vehicle for the meme's propagation in just the way that a virus may parasitize the genetic mechanism of a host cell" (p. 192).

With Dawkins' point in mind we are now in a position to extract from the writings of the memetic theorists (e.g., Aunger, 2000; Blackmore, 1999; Dawkins, 1993; Dennett, 1991, 1995; Lynch, 1996) a taxonomy of reasons for meme survival:

1. Memes survive and spread because they are helpful to the interactors that store them (most memes that reflect true information in the world would be in this category).
2. Memes become frequent because they fit genetic predispositions, or domain-specific evolutionary modules (this is the evoked culture that is emphasized by evolutionary psychologists, see Cosmides & Tooby, 1992; Sperber, 1996).
3. Memes spread because they facilitate the spread of the genes that make good hosts for these particular memes (religious beliefs that urge people to have more children would be in this category, see Lynch, 1996).
4. Memes survive and spread because of the self-perpetuating properties of the memes themselves.

We must consider these categories in the context of the fact that many of the intentional-level goals that humans have are meme-installed: they are the products of our culture, rather than installed by the genes that built the vehicle. A schematic that helps to understand our conception of the intentional-level goal structure of System 1 and System 2 in terms of which replicator is a source of the goal is portrayed in Figure 6 (again, absolute areas are guesses--for illustrative purposes only). The goal structure of System 1 is dominated by gene-installed goals. These are the short-leash goals discussed earlier--nearly universal in the sense that they are shared by most humans and not the result of the environmental history of the organism. They are not flexible or generic goals, but instead are content specific, situation specific, and hard-wired to trigger (disgust and repulsion to noxious smells and substances, and fear responses to animals like snakes, would be examples, see Buss, 1999; Rozin, 1996; Rozin & Fallon, 1987).

System 2, with its more general, flexible goals is more evenly balanced with genetic goals shared by most humans (e.g., rise in the dominance hierarchy of your conspecifics) and with meme-installed goals that are the result of the specific environmental experience (and culture) of the individual. In Figure 6 we distinguish between memetically-acquired goals that are "caught" like viruses (as in the Dawkins quote above)--what we call nonreflectively acquired memetic goals--and memetic goals that an individual takes on reflectively, with full awareness of their effects on the organism. The nonreflectively acquired goals are perhaps the equivalent of the parasites that Dawkins refers to. They may not actually be good for the individual, but just like the vehicle-sacrificing genes discussed previously, these memes use the vehicle merely to propagate themselves.

The diagram also indicates that meme-acquired goals need not be barred from becoming System 1 goals (automatic, autonomous, and rapidly triggering). Through practice, memetically installed goals can become lodged in the goal hierarchy of System 1. "Branding" and other advertising gimmicks aspire to do just this--to have a logo for X trigger the "must have X" response without much thought. These then become especially pernicious memes--parasites that, because they are not part of the reflective mind, become difficult to dislodge.

Of course, meme-derived goals that become part of System 1 can be good for the vehicle too. A reflectively acquired meme--one that was reflectively acquired because it served vehicle ends (perhaps even vehicle ends that thwart the genes' interests)--can become part of System 1 as well. This fact explains a part of Figure 5 that might have seemed perplexing when that Figure was first

presented. Why is there a small section of area in System 1 representing goals that serve the vehicle's interests only? One might have thought that all of the goals instantiated in System 1 would reflect the genes' interests whether or not they were serving the interests of the vehicle--rather like that of the Darwinian creature represented in Figure 3. However the possibility of the higher-level goal states of System 2 becoming installed in the more rigid and inflexible System 1 through practice opens up a new possibility. Reflectively acquired goal-states might be memes that were taken on for their unique advantages to the vehicle (advantages that might accrue because they trump contrary gene-installed goals--"don't flirt with your boss's wife"). Those particular memes becoming instantiated in System 1 through practice would create the area depicted in Figure 5--System 1 goal states serving the vehicle's interests only. We might say that in situations such as this, System 1 in humans reflects the outcome of residing in a brain along with a reflective System 2. This is why the goal-structure of System 1 in humans does not simply recapitulate the structure of a Darwinian creature depicted in Figure 3.

What evolutionary psychologists do not like about the previous conceptualization is the notion of memes becoming completely "unglued" from genetic control. Instead, evolutionary psychologists prefer Lumsden and Wilson's (1981) notion that the genes hold culture on a leash (see Sperber, 1996). What they do not like is the idea--which we are advancing here--that at a certain level of recursiveness a Gregorian mind populated with cultural tools in the form of memeplexes designed for the evaluation of other memeplexes (science; logic; some notions from decision science such as consistency, transitivity, etc.) acquires some autonomy from genetic control. But it is just such autonomy that we are arguing for. We are in fact arguing that the cultural tools of logic and decision science--when reflectively used in conjunction with the potent cultural insight that there can be a conflict of interest between replicators and vehicle--have the potential to create a creature with a uniquely critical and discerning type of self reflection. In short, we are arguing that understanding the full implications of the replicator/vehicle distinction may be a cultural tool that could foster even greater levels of self reflection than humans have heretofore achieved.

Combined with the tools of decision science, the vehicle/replicator distinction can spawn thoughts and new tools for the restructuring of human goals--new memes that further sever the connection between memeplexes resident in some brains and genetic goals installed by the replicators. Indeed, we propose that this cultural change is already underway. There are already memeplexes in the air (of which this book is one) that will, contrary to the emphasis in the writings of evolutionary psychologists, further background the role of the genes in human culture<sup>9</sup>.

Evolutionary psychologists resist this extrapolation--falling back on their "culture on a leash" notion. For example, Tooby and Cosmides (1992) insist that "epidemiological culture is also shaped by the details of our evolved psychological organization" (p. 119) and, even more strongly, that "our developmental and psychological programs evolved to invite the social and cultural worlds in, but only the parts that tended, on balance, to have adaptively useful effects" (p. 87). But the evolutionary psychologists seem to have underestimated the power of the memes to break this linkage. What they have neglected is the recursive power of evaluative memes in the context of an organism that has become aware of the replicator/vehicle distinction. Science writer Robert Wright (1994) paraphrases the Tooby and Cosmides statement above about "our developmental and psychological programs evolved..." into the more readable notion that ideas must "have a kind of harmony with the brains they settle into" (p. 366). However, unlike Tooby and Cosmides (1992), Wright (1994) realizes that there are implications that follow from becoming aware of this fact in the context of replicator/vehicle distinction. So after noting that ideas must have a kind of harmony with the brains they settle into, Wright (1994) warns that "that doesn't mean they're good for those brains in the long run" (p. 366).

A brain that realizes this startling (and still underappreciated) fact might begin the (admittedly difficult) process of slipping culture off the leash of the genes when the culture is dysfunctional for the person<sup>10</sup>. A brain aware of the replicator/vehicle distinction and in the possession of evaluative memplexes such as science, logic, and decision theory might begin a process of pruning vehicle-thwarting goals from intentional-level psychology and reinstalling memetic structures that serve the vehicle's interests more efficiently (of course, this is exactly what the canons of normative instrumental rationality--largely a product of the twentieth century--were designed to accomplish).

This might seem like a Promethean goal, but in fact, a rich tradition in cognitive science has emphasized how cultural changes and increased scientific knowledge results in changes in folk psychology. For example, Churchland (1989, 1995) has long emphasized how a mature neuroscience might change our folk language of the mental and of behavior. Other theorists have emphasized how moral notions change as general knowledge of the deterministic explanatory power of neuroscience becomes more widespread (Wright, 1994). Already, among educated citizens of the 21st Century, violations of transitivity and independence of irrelevant alternatives can be a cause of cognitive sanction in ways that are probably historically unprecedented. A full appreciation of the implications of the replicator/vehicle distinction--with its emphasis that differing optimization criteria apply to the personal and subpersonal levels of analysis (utility maximization versus genetic fitness)--could have equally profound cultural implications.

Of course, we do not mean to imply that all evolutionary psychologists are guilty of committing a sophisticated version of the genetic fallacy (inferring current function from ancestral function, see Dennett, 1995, p. 465). For example, Pinker (1997) does not endorse the culture-on-a-short-leash view and explicitly recognizes the implications of the differing interests of the replicators and the vehicle: "Genes are not puppetmasters; they acted as the recipes for making the brain and body and then they got out of the way. They live in a parallel universe, scattered among bodies, with their own agendas" (p. 401).

Clearly, not all evolutionary psychologists miss the implication of replicator/vehicle distinction for conceptions of rationality. But some evolutionary theorists do--quite egregiously. In an astonishing essay titled "How Evolutionary Biology Challenges the Classical Theory of Rational Choice", Cooper (1989) basically argues that when choosing between your own goals and those of your genes, you should opt for the latter! After a marvelous discussion of why a probability matching strategy (Estes, 1961, 1976) might be fitness optimizing rather than the utility maximizing strategy (picking the most frequent option each time), Cooper (1989) implies that this outcome undermines the prescriptive force of the utility maximizing strategy: "The upshot is that one is faced with a dilemma. Either rationality is not always the fittest policy, or else classical decision analysis is not as universally rational as is commonly claimed. If the latter horn of the dilemma is seized (and I shall argue that that is indeed the lesser of the evils)" (p. 459). Of course, early in the article one feels that this is a verbal slip. But ten pages on, we find out that the author does indeed wish to argue that we should follow goals that satisfy our genes rather than ourselves as individual organisms. The ordinary application of the logic of decision science is termed "naively applied" when interpreted "with the individual treated as an isolated locus of decision making and with the role of the genotype ignored" (p. 473). The instability in preference orderings that signal the failure of individual utility maximization (Dawes, 1998; Kahneman & Tversky, 2000; Slovic, 1995) are defended because "perhaps some of the observed instability is due to adaptive strategy mixing. If so, instability would have to be reevaluated; when one is acting as an agent of one's genotype, it could sometimes be a sound strategy" (p. 473). But who in the world would want to act as an agent of one's genotype rather than in the service of one's own life goals! This is precisely the choice Cooper

(1989) is posing when he pits the concerns of genetic fitness against those of instrumental rationality.

Lest the reader worry that we caricature, Cooper (1989) leaves us in no doubt because he concludes his paper with replies to possible criticisms of his view. In that section he makes it clear that his view is that "all adequate choice rules are seen as mere extensions of evolutionary principles" (p. 475) because "the maximization of fitness is a ubiquitous goal" (p. 475). A ubiquitous goal of the genes no doubt, but--completely ignoring the individual organism as a potential critical locus of utility maximization--Cooper (1989) proposes that we "revise the classical theory [utility theory] itself at its mathematical core, letting biologically motivated decision rules...replace or supplement the traditional ones as basic decision rules" (p. 475).

In his summary statement, Cooper (1989) makes it clear that the proposition he wishes to defend is that "the traditional theory of rationality is invalid as it stands, and in need of biological repair" (p. 479), and acknowledges that this is "a stance not likely to be popular with confirmed classical decision theorists, but perhaps understandable to evolutionists, psychologists, philosophers, and others that have been impressed by the pervasive explanatory power of the modern evolutionary perspective" (p. 479). The view explicitly championed is the notion that "behavioral rationality [be] interpreted in terms of fitness" (p. 480) and that any dissent from this policy be viewed as 'biologically naive' (p. 480). Like the sociobiologists before him, Cooper (1989) seems to have taken the defense of the genes as his brief!

Cooper's (1989) view may well seem extreme, and few evolutionary psychologists so explicitly throw out the vehicle with the bathwater. But many evolutionary psychologists and proponents of ecological rationality (Cosmides & Tooby, 1994a; Gigerenzer, 1996a; Gigerenzer & Todd, 1999) actually do do it implicitly in the way that they echo Cooper's (1989) contention that "the traditional theory of rationality is invalid as it stands, and in need of biological repair" (p. 479). For example, in a paper discussing economics and evolutionary psychology, Cosmides and Tooby (1994a) quite closely mimic Cooper's (1989) view when they argue that "evolutionary considerations suggest that traditional normative and descriptive approaches to rationality need to be reexamined" (p. 329). Throughout this essay they repeat the odd declaration that "despite widespread claims to the contrary, the human mind is not worse than rational (e.g., because of processing constraints)--but may often be better than rational" (p. 329).

It is in fact relatively common for the traditional normative rules of rational thought to be denigrated in the literature critical of the heuristics and biases approach. Gigerenzer and Goldstein (1996) adopt exactly Cooper (1989) extreme position in their argument that in their view "questions classical rationality as a universal norm and thereby questions the very definition of 'good' reasoning on which both the Enlightenment and the heuristics-and-biases views were built" (p. 651). The classical norms are referred to as just so much useless "baggage" in quotes such as the following: "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" (Gigerenzer & Todd, 1999, p. 365).

Likewise, as noted in our discussion above, Gigerenzer's repeated refrain that cognitive illusions (violations of the canons of normative rationality) "disappear" with more evolutionarily propitious problem representations implies that these normative violations are of no concern. Since we know that they (the normative violations) are of concern to the vehicle (a vehicle who does not follow them does not maximize utility), we can only conclude that what these authors want us to imply is that evolutionary rationality is the only rationality that need concern us--precisely Cooper's (1989) point (although the point is much more subtle and somewhat hidden in the writings of the ecological theorists).

Cosmides and Tooby (1994a), in an essay directed at economists, ignore completely the role of memetic evolution and culture in determining human preferences. In a series of points laid out like a series of axioms they argue that because "natural selection built the decision-making machinery in human minds" (p. 328) and because "this set of cognitive devices generates all economic behavior", "therefore...the design features of these devices define and constitute the human universal principles that guide economic decision making" (p. 328).

These postulates lead Cosmides and Tooby (1994a) to the grandiose claim that "evolutionary psychology should be able to supply a list of human universal preferences, and of the procedures by which additional preferences are acquired or reordered" (p. 331). But to the extent that the claim is true, it is only because the grain-size of the predictions will be all wrong. The economic literature is not full of studies debating whether humans who are dying of thirst prefer water or shelter--or whether men prefer 23-year-old females over 75-year-old ones. Instead, the literature is full of studies trying to determine the rationale for such fine-grained judgments as, for example, whether a poor briefcase produced by an athletic shoe company will adversely affect the family brand name (Ahluwalia & Gurhan-Canli, 2000). Economists and psychologists are not debating the reasons for preferences among basic biological needs. Instead, they are debating the reasons for fine-grained preferences among highly symbolic products embedded in a complex, information-saturated, "attention-based" (Davenport & Beck, 2001) economy. Even after we grant evolutionary assumptions like, for example, that people use clothes purchases for some type of modern dominance display or sexual display, we have not progressed very far in explaining how brand names wax and wane in the fashion world, or how price elastic such purchases will be, and/or what kind of substitutability there will be among these types of goods.

This essay by Cosmides and Tooby (1994a) directed to economists serves to reinforce all of the worst Panglossian tendencies in the latter discipline. For example, Kahneman, Wakker, and Sarin (1997) discuss why experienced utility is essentially ignored in modern economics despite psychological studies showing that experienced utility is not identical to expected utility. They argue that experienced utility is ignored by economists on the grounds that "choices provide all necessary information about the utility of outcomes because rational agents who wish to do so will optimize their hedonic experience" (p. 375). Two-process theories of cognition--in conjunction with the assumptions that we have made about goal structures--help to explain why this assumption might not hold. The choices triggered by the goal structures of System 1 might not always be oriented toward the optimization of hedonic experience for the individual agent. The hedonic experience is just a means to an end for most of the goals lodged in System 1 (largely genetic goals). This System will readily sacrifice the vehicle's hedonic pleasure if ultimate fitness goals are achievable without it.

#### Choosing the Vehicle Rather than the Replicators:

#### Evolutionary Psychology Without Greedy Reductionism

What the evolutionary psychologists and ecological rationality theorists have occasionally been guilty of in the domain of rationality is what Dennett (1995) has termed "greedy reductionism". According to Dennett (1995), in their "zeal to explain too much too fast" (p. 82), greedy reductionists tend to "skip whole layers or levels of theory in their rush to fasten everything securely and neatly to the foundation" (p. 82). Like Dennett (1995), we applaud reductionist efforts in the behavioral sciences. We are impressed with the seminal achievements of evolutionary psychology (see Table 1 of Buss, Haselton, Shackelford, Beske, & Wakefield, 1998, for a long list of important behavioral relationships that were in large part uncovered because of applications of the theoretical lens of evolutionary psychology) and consider its emergence as a dominant force in psychology during the 1990s (Barkow, Cosmides, & Tooby, 1992; Buss, 1999, 2000; Cartwright, 2000; Cosmides & Tooby, 1994b; Geary & Bjorklund, 2000; Pinker, 1997; Plotkin, 1998) to be a salutary

development. But in the area of rationality, the evolutionary psychologists have built a bridge too far. They too easily gloss over the important issue of replicator/vehicle goal mismatches and their implications. They too easily dismiss the role of general intelligence and/or general computational power in overriding deleterious System 1 responses (Stanovich, 1999; Stanovich & West, 2000). Because many of the tools of instrumental and epistemic rationality are cultural inventions (memes) and not biological modules, their usefulness in technological societies is too readily dismissed by evolutionary psychologists.

In the extreme, evolutionary theorists begin to sound as if they are siding with genetic interests against those of people. Cooper (1989), in the essay quoted extensively above, admits that "nonclassical behaviors such as betting against the probabilities are detrimental to the reasoner's own welfare" (p. 477), but argues that this is justified because "what if the individual identifies its own welfare with that of its genotype?" (p. 477). Well, what if? Then yes, maybe they should probability match. But who are these people with such loyalty to the random shuffle of genes that is their genotype? Which alleles, for example, do you have particularly emotional feelings for? Beyond a few scientists too narrowly focused on the promised explanatory power of evolutionary psychology, we doubt that there are such people<sup>11</sup>.

Gibbard (1990) offers the more reasoned view:

"it is crucial to distinguish human goals from the Darwinian surrogate of purpose in the 'design' of human beings....The Darwinian evolutionary surrogate for divine purpose is now seen to be the reproduction of one's genes. That has not, as far as I know, been anyone's goal, but the biological world looks as if someone quite resourceful had designed each living thing for that purpose....A person's evolutionary telos explains his having the propensities in virtue of which he develops the goals he does, but his goals are distinct from this surrogate purpose. My evolutionary telos, the reproduction of my genes, has no straightforward bearing on what it makes sense for me to want or act to attain....A like conclusion would hold if I knew that I was created by a deity for some purpose of his: his goal need not be mine...Likewise, if I know that my evolutionary telos is to reproduce my genes, that in itself gives me no reason for wanting many descendants" (pp. 28-29).

In short, "human moral propensities were shaped by something it would be foolish to value in itself, namely multiplying one's own genes" (p. 327)

Gibbard's (1990) view is shared by distinguished biologist George Williams (1988) who feels that "there is no conceivable justification for any personal concern with the interests (long-term average proliferation) of the genes we received in the lottery of meiosis and fertilization. As Huxley was the first to recognize, there is every reason to rebel against any tendency to serve such interest" (p. 403).

Dennett (1995) discusses this point in a different way--by making the astonishing observation that until quite recently, the genes were the only beneficiary of all of the selective forces on the planet. That is, "there were no forces whose principle beneficiary was anything else. There were accident and catastrophes (lightening bolts and tidal waves), but no steady forces acting systematically to favor anything but genes" (p. 328). But now we are here. There exist in the universe, for the first time, another set of interests because, unlike Darwinian creatures, our interests are not necessarily our genes'. Rationality is the meme that trumps genetic interests in cases such as this. The remarkable cultural project to advance human rationality concerns how to best advance human interests whether or not they coincide with genetic interests. Its emancipatory potential is lost if we fail to see the critical divergence of interests that creates the distinction between evolutionary and instrumental rationality.

## Notes

Preparation of this chapter was supported by a grant from the Social Sciences and Humanities Research Council of Canada to Keith E. Stanovich.

## References

- Adler, J. E. (1984). Abstraction is uncooperative. Journal for the Theory of Social Behaviour, 14, 165-181.
- Adler, J. E. (1991). An optimist's pessimism: Conversation and conjunctions. In E. Eells & T. Maruszewski (Eds.), Probability and rationality: Studies on L. Jonathan Cohen's philosophy of science (pp. 251-282). Amsterdam: Editions Rodopi.
- Ahluwalia, R., & Gurhan-Canli, Z. (2000). The effects of extensions on the family brand name: An accessibility-diagnostics perspective. Journal of Consumer Research, 27, 371-381.
- 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., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. Educational Researcher, 25(4), 5-11.
- Arkes, H. R., & Harkness, A. R. (1983). Estimates of contingency between two dichotomous variables. Journal of Experimental Psychology: General, 112, 117-135.
- Aunger, R. (2000). Darwinizing culture: the status of memetics as a science. New York: Oxford University Press.
- Badcock, C. (2000). Evolutionary psychology: A critical introduction. Cambridge, England: Polity Press.
- Baldwin, D. A. (2000). Interpersonal understanding fuels knowledge acquisition. Current Directions in Psychological Science, 9, 40-45.
- Bar-Hillel, M., & Falk, R. (1982). Some teasers concerning conditional probabilities. Cognition, 11, 109-122.
- Barkow, J. H. (1989). Darwin, sex, and status: Biological approaches to mind and culture. Toronto: University of Toronto Press.
- Barkow, J., Cosmides, L., & Tooby, J. (Eds.) (1992). The adapted mind. New York: Oxford University Press.
- Baron, J. (1993). Morality and rational choice. Dordrecht: Kluwer.
- Baron, J. (1994). Nonconsequentialist decisions. Behavioral and Brain Sciences, 17, 1-42.
- Baron, J. (1998). Judgment misguided: Intuition and error in public decision making. New York: Oxford University Press.
- Baron, J. (2000). Thinking and deciding (Third Edition). Cambridge, MA: Cambridge University Press.
- Barton, R. A., & Dunbar, R. (1997). Evolution of the social brain. In A. Whiten & R. W. Byrne (Eds.), Machiavellian intelligence II: Extensions and evaluations (pp. 240-263). Cambridge: Cambridge University Press.
- Bazerman, M. (2001). Consumer research for consumers. Journal of Consumer Research, 27, 499-504.
- Blackmore, S. (1999). The meme machine. New York: Oxford University Press.
- Brase, G. L., Cosmides, L., & Tooby, J. (1998). Individuation, counting, and statistical inference: The role of frequency and whole-object representations in judgment under uncertainty. Journal of Experimental Psychology: General, 127, 3-21.



- Brody, N. (1997). Intelligence, schooling, and society. American Psychologist, 52, 1046-1050.
- Bronfenbrenner, U., McClelland, P., Wethington, E., Moen, P., & Ceci, S. J. (1996). The state of Americans. New York: Free Press.
- Brothers, L. (1990). The social brain: A project for integrating primate behaviour and neuropsychology in a new domain. Concepts in Neuroscience, 1, 27-51.
- Bugental, D. B. (2000). Acquisitions of the algorithms of social life: A domain-based approach. Psychological Bulletin, 126, 187-219.
- Buss, D. M. (1999). Evolutionary psychology: The new science of the mind. Boston: Allyn and Bacon.
- Buss, D. M. (2000). The evolution of happiness. American Psychologist, 55, 15-23.
- Buss, D. M., Haselton, M. G., Shackelford, T., Beske, A., & Wakefield, J. (1998). Adaptations, exaptations, and spandrels. American Psychologist, 53, 533-548.
- Byrne, R. W., & Whiten, A. (Eds.). (1988). Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans. Oxford: Oxford University Press.
- Caporael, L. R. (1997). The evolution of truly social cognition: The core configurations model. Personality and Social Psychology Review, 1, 276-298.
- Cartwright, J. (2000). Evolution and human behavior. Cambridge, MA: MIT Press.
- Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), Unintended thought (pp. 212-252). New York: Guilford Press.
- Chater, N., & Oaksford, M. (2000). The rational analysis of mind and behaviour. Synthese, 122, 93-131.
- Churchland, P. M. (1989). A neurocomputational perspective: The nature of mind and the structure of science. Cambridge, MA: MIT Press.
- Churchland, P. M. (1995). The engine of reason, the seat of the soul. Cambridge, MA: MIT Press.
- Clark, A. (1997). Being there: Putting brain, body, and world together again. Cambridge, MA: MIT Press.
- Cooper, W. S. (1989). How evolutionary biology challenges the classical theory of rational choice. Biology and Philosophy, 4, 457-481.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), The adapted mind, (pp. 163-228). New York: Oxford University Press.
- Cosmides, L., & Tooby, J. (1994a). Better than rational: Evolutionary psychology and the invisible hand. American Economic Review, 84, 327-332.
- Cosmides, L., & Tooby, J. (1994b). Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. Cognition, 50, 41-77.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition, 58, 1-73.
- Cummins, D. D. (1996). Evidence for the innateness of deontic reasoning. Mind & Language, 11, 160-190.
- Davenport, T., & Beck, J. (2001). The attention economy. Cambridge, MA: Harvard Business School Press.
- Davies, P. S. (1996). Discovering the functional mesh: On the methods of evolutionary psychology. Minds and Machines, 6, 559-585.
- Dawes, R. M. (1998). Behavioral decision making and judgment. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), The handbook of social psychology (Vol. 1) (pp. 497-548). Boston: McGraw-Hill.

- Dawkins, R. (1976). The selfish gene (New edition, 1989). New York: Oxford University Press.
- Dawkins, R. (1982). The extended phenotype. New York: Oxford University Press.
- Dawkins, R. (1993). Viruses of the mind. In B. Dahlbom (Ed.), Dennett and his critics, (pp. 13-27). Cambridge, MA: Blackwell.
- Deary, I. J. (2000). Looking down on human intelligence: From psychometrics to the brain. Oxford: Oxford University Press.
- Dennett, D. C. (1975). Why the law of effect will not go away. Journal of the Theory of Social Behavior, 2, 169-187.
- Dennett, D. C. (1984). Elbow room: The varieties of free will worth wanting. Cambridge, MA: MIT Press.
- Dennett, D. C. (1987). The intentional stance. Cambridge, MA: MIT Press.
- Dennett, D. C. (1991). Consciousness explained. Boston: Little Brown.
- Dennett, D. C. (1995). Darwin's dangerous idea: Evolution and the meanings of life. New York: Simon & Schuster.
- Dennett, D. C. (1996). Kinds of minds: Toward an understanding of consciousness. New York: Basic Books.
- Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. Psychological Bulletin, 108, 346-369.
- Dulany, D. E., & Hilton, D. J. (1991). Conversational implicature, conscious representation, and the conjunction fallacy. Social Cognition, 9, 85-110.
- Dunbar, R. (1998). Theory of mind and the evolution of language. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), Approaches to the evolution of language (pp. 92-110). Cambridge: Cambridge University Press.
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. Annual Review of Psychology, 32, 53-88.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. American Psychologist, 49, 709-724.
- Estes, W. B. (1964). Probability learning. In A. W. Melton (Ed.), Categories of human learning (pp. 89-128). New York: Academic Press.
- Estes, W. K. (1976). The cognitive side of probability learning. Psychological Review, 83, 37-64.
- Estes, W. K. (1984). Global and local control of choice behavior by cyclically varying outcome probabilities. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 258-270.
- Evans, J. St. B. T. (1984). Heuristic and analytic processes in reasoning. British Journal of Psychology, 75, 451-468.
- Evans, J. St. B. T. (1989). Bias in human reasoning: Causes and consequences. London: Erlbaum Associates.
- Evans, J. St. B. T. (1996). Deciding before you think: Relevance and reasoning in the selection task. British Journal of Psychology, 87, 223-240.
- Evans, J. St. B. T., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. Memory & Cognition, 11, 295-306.
- Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). Human reasoning: The psychology of deduction. Hove, England: Erlbaum.
- Evans, J. St. B. T., & Over, D. E. (1996). Rationality and reasoning. Hove, England: Psychology Press.
- Evans, J. St. B. T., & Over, D. E. (1997). Rationality in reasoning: The problem of deductive competence. Cahiers de Psychologie Cognitive (Current Psychology of Cognition), 16, 3-38.

- Evans, J. St. B. T., & Over, D. E. (1999). Explicit representations in hypothetical thinking. Behavioral and Brain Sciences, 22, 763-764.
- Evans, J. St. B. T., Simon, J. H., Perham, N., Over, D. E., & Thompson, V. A. (2000). Frequency versus probability formats in statistical word problems. Cognition, 77, 197-213.
- Evans, J. St. B. T., & Wason, P. C. (1976). Rationalization in a reasoning task. British Journal of Psychology, 67, 479-486.
- Falk, R. (1992). A closer look at the probabilities of the notorious three prisoners. Cognition, 43, 197-223.
- Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. Psychological Research, 50, 123-129.
- Fodor, J. (1983). Modularity of mind. Cambridge: MIT Press.
- Foley, R. (1996). The adaptive legacy of human evolution: A search for the EEA. Evolutionary Anthropology, 4, 194-203.
- Frank, R. H., & Cook, P. J. (1995). The winner-take-all society. New York: Free Press.
- Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. Psychological Bulletin, 101, 75-90.
- Gal, I., & Baron, J. (1996). Understanding repeated simple choices. Thinking and Reasoning, 2(1), 81-98.
- Gallistel, C. R. (1990). The organization of learning. Cambridge, MA: MIT Press.
- Geary, D. C., & Bjorklund, D. F. (2000). Evolutionary developmental psychology. Child Development, 71, 57-65.
- Gibbard, A. (1990). Wise choices, apt feelings: A theory of normative judgment. Cambridge, MA: Harvard University Press.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". European Review of Social Psychology, 2, 83-115.
- Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. In K. Manktelow & D. Over (Eds.), Rationality: Psychological and philosophical perspectives (pp. 284-313). London: Routledge.
- Gigerenzer, G. (1996a). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). Psychological Review, 103, 592-596.
- Gigerenzer, G. (1996b). Rationality: Why social context matters. In P. B. Baltes & U. Staudinger (Eds.), Interactive minds: Life-span perspectives on the social foundation of cognition (pp. 319-346). Cambridge: Cambridge University Press.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. Psychological Review, 103, 650-669.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. Psychological Review, 102, 684-704.
- Gigerenzer, G., Hoffrage, U., & Ebert, A. (1998). AIDS counselling for low-risk clients. AIDS Care, 10, 197-211.
- Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. Psychological Review, 98, 506-528.
- Gigerenzer, G., & Regier, T. (1996). How do we tell an association from a rule? Comment on Sloman (1996). Psychological Bulletin, 119, 23-26.
- Gigerenzer, G., & Todd, P. M. (1999). Simple heuristics that make us smart. New York: Oxford University Press.
- Giroto, V., & Gonzalez, M. (2001). Solving probabilistic and statistical problems: A matter of information structure and question form. Cognition, 78, 247-276.

- Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer & P. M. Todd, Simple heuristics that make us smart (pp. 37-58). New York: Oxford University Press.
- Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. Intelligence, 24, 79-132.
- Graf, V., Bullock, D. H., & Bitterman, M. E. (1964). Further experiments on probability-matching in the pigeon. Journal of the Experimental Analysis of Behavior, 7, 151-157.
- Granberg, D. (1995). The Monte Hall dilemma. Personality and Social Psychology Bulletin, 31, 711-723.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), Syntax and semantics: Vol. 3, Speech acts (pp. 41-58). New York: Academic Press.
- Grigorenko, E. L. (1999). Heredity versus environment as the basis of cognitive ability. In R. J. Sternberg (Ed.), The nature of cognition (pp. 665-696). Cambridge, MA: MIT Press.
- Hardman, D. (1998). Does reasoning occur on the selection task? A comparison of relevance-based theories. Thinking and Reasoning, 4, 353-376.
- Harries, C., & Harvey, N. (2000). Are absolute frequencies, relative frequencies, or both effective in reducing cognitive biases? Journal of Behavioral Decision Making, 13, 431-444.
- Hasher, L., & Zacks, R. T. (1979). Automatic processing of fundamental information: The case of frequency of occurrence. Journal of Experimental Psychology: General, 39, 1372-1388.
- Hastie, R., & Rasinski, K. A. (1988). The concept of accuracy in social judgment. In D. Bar-Tal & A. Kruglanski (Eds.), The social psychology of knowledge (pp. 193-208). Cambridge: Cambridge University Press.
- Herrnstein, R. J., & Loveland, D. H. (1975). Maximizing and matching on concurrent ratio schedules. Journal of the Experimental Analysis of Behavior, 24, 107-116.
- Hertwig, R., & Gigerenzer, G. (1999). The conjunction fallacy revisited: How intelligent inferences look like reasoning errors. Journal of Behavioral Decision Making, 12, 275-305.
- Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. Psychological Bulletin, 118, 248-271.
- Hilton, D. J., & Slugoski, B. R. (2000). Judgment and decision making in social context: Discourse processes and rational inference. In T. Connolly, H. R. Arkes, & K. R. Hammond (Eds.), Judgment and decision making: An interdisciplinary reader (Second Edition) (pp. 651-676). Cambridge, MA: Cambridge University Press.
- Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F. I. M. Craik & S. Trehub (Eds.), Aging and cognitive processes (pp. 847-870). New York: Plenum Press.
- Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. Acta Psychologica, 26, 1-23.
- Hull, D. L. (1982). The naked meme. In H. C. Plotkin (Eds.), Learning, development, and culture: Essays in evolutionary epistemology (pp. 273-327). Chichester, England: John Wiley.
- Hull, D. L. (1988). Science as a process: An evolutionary account of the social and conceptual development of science. Chicago: University of Chicago Press.
- Humphrey, N. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), Growing points in ethology (pp. 303-317). London: Faber & Faber.
- Hunt, E. (1995). Will we be smart enough? A cognitive analysis of the coming workforce. New York: Russell Sage Foundation.
- Hunt, E. (1999). Intelligence and human resources: Past, present, and future. In P. Ackerman, P. Kyllonen, & R. Richards (Eds.), Learning and individual differences: Process, trait, and content determinants (pp. 3-28). Washington, DC: American Psychological Association.

- Johnson-Laird, P. N. (1983). Mental models. Cambridge, MA: Harvard University Press.
- Johnson-Laird, P. N. (1999). Deductive reasoning. Annual Review of Psychology, 50, 109-135.
- Jolly, A. (1966). Lemur social behaviour and primate intelligence. Science, 153, 501-506.
- Kahneman, D. (2000). A psychological point of view: Violations of rational rules as a diagnostic of mental processes. Behavioral and Brain Sciences, 23.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.) (1982). Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. Cognitive Psychology, 3, 430-454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. Psychological Review, 80, 237-251.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. American Psychologist, 39, 341-350.
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. Psychological Review, 103, 582-591.
- Kahneman, D., & Tversky, A. (Ed.). (2000). Choices, values, and frames. Cambridge: Cambridge University Press.
- Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. The Quarterly Journal of Economics, 112(2), 375-405.
- Kao, S. F., & Wasserman, E. A. (1993). Assessment of an information integration account of contingency judgment with examination of subjective cell importance and method of information presentation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 1363-1386.
- Klein, G. (1998). Sources of power: How people make decisions. Cambridge, MA: MIT Press.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory, 6, 107-118.
- Kruglanski, A. W., & Ajzen, I. (1983). Bias and error in human judgment. European Journal of Social Psychology, 13, 1-44. and "freezing". Psychological Review, 103, 263-283.
- Kummer, H., Daston, L., Gigerenzer, G., & Silk, J. B. (1997). The social intelligence hypothesis. In P. Weingart, S. D. Mitchell, P. J. Richerson, & S. Maasen (Eds.), Human by nature: Between biology and the social sciences (pp. 157-179). Mahwah, NJ: Lawrence Erlbaum Associates.
- LaCerra, P., & Bingham, R. (1998). The adaptive nature of the human neurocognitive architecture: An alternative model. Proceeds of the National Academy of Sciences, 95, 11290-11294.
- Levin, I. P., Wasserman, E. A., & Kao, S. F. (1993). Multiple methods of examining biased information use in contingency judgments. Organizational Behavior and Human Decision Processes, 55, 228-250.
- Levinson, S. C. (1995). Interactional biases in human thinking. In E. Goody (Eds.), Social intelligence and interaction (pp. 221-260). Cambridge: Cambridge University Press.
- Liberman, N., & Klar, Y. (1996). Hypothesis testing in Wason's selection task: Social exchange cheating detection or task understanding. Cognition, 58, 127-156.
- Looren de Jong, H., & van der Steen, W. J. (1998). Biological thinking in evolutionary psychology: Rockbottom or quicksand? Philosophical Psychology, 11, 183-205.
- Lubinski, D. (2000). Scientific and social significance of assessing individual differences: "Sinking shafts at a few critical points". Annual Review of Psychology, 51, 405-444.
- Lubinski, D., & Humphreys, L. G. (1997). Incorporating general intelligence into epidemiology and the social sciences. Intelligence, 24, 159-201.

- Lumsden, C. J., & Wilson, E. O. (1981). Genes, mind and culture. Cambridge, MA: Harvard University Press.
- Luria, A. R. (1976). Cognitive development: Its cultural and social foundations. Cambridge, MA: Harvard University Press.
- Lynch, A. (1996). Thought contagion. New York: Basic Books.
- Macchi, L. (1998). Computational features vs frequentist phrasing in the base-rate fallacy. Swiss Journal of Psychology, *57*, 79-85.
- MacDonald, K., & Geary, D. C. (2000). g and Darwinian algorithms. Behavioral and Brain Sciences, *23*, 685-686.
- MacDonall, J. S. (1988). Concurrent variable-ratio schedules: Implications for the generalized matching law. Journal of the Experimental Analysis of Behavior, *50*, 55-64.
- Macdonald, R. R., & Gilhooly, K. J. (1990). More about Linda or conjunctions in context. European Journal of Cognitive Psychology, *2*, 57-70.
- Manktelow, K. I. (1999). Reasoning & Thinking. Psychology Press.
- Margolis, H. (1987). Patterns, thinking, and cognition. Chicago: University of Chicago Press.
- Margolis, H. (1996). Dealing with risk. Chicago: University of Chicago Press.
- Matthews, G., & Deary, I. J. (1998). Personality traits. Cambridge: Cambridge University Press.
- McGeorge, P., Crawford, J., & Kelly, S. (1997). The relationships between psychometric intelligence and learning in an explicit and an implicit task. Journal of Experimental Psychology: Learning, Memory, and Cognition, *23*, 239-245.
- Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. Psychological Science, *12*, 269-275.
- Millikan, R. G. (1993). White Queen psychology and other essays for Alice. Cambridge, MA: The MIT Press.
- Mithen, S. (1996). The prehistory of mind: The cognitive origins of art and science. London: Thames and Hudson.
- Newstead, S. E., & Evans, J. St. B. T. (Eds.) (1995). Perspectives on thinking and reasoning. Hove, England: Erlbaum.
- Nickerson, R. S. (1996). Hempel's paradox and Wason's selection task: Logical and psychological puzzles of confirmation. Thinking and Reasoning, *2*, 1-31.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, *2*, 175-220.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), Consciousness and self-regulation, (pp. 1-18). New York: Plenum.
- Nozick, R. (1993). The nature of rationality. Princeton, NJ: Princeton University Press.
- 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. (1996). Rational explanation of the selection task. Psychological Review, *103*, 381-391.
- Oaksford, M., & Chater, N. (1998). Rationality in an uncertain world. Hove, England: Psychology Press.
- Oaksford, M., & Chater, N. (2001). The probabilistic approach to human reasoning. Trends in Cognitive Sciences, *5*, 349-357.

- Over, D. E. (2000). Ecological rationality and its heuristics. Thinking and Reasoning, 6, 182-192.
- Over, D. E., & Green, D. W. (in press). Contingency, causation, and adaptive inference. Psychological Review.
- Pinker, S. (1994). The language instinct. New York: William Morrow.
- Pinker, S. (1997). How the mind works. New York: Norton.
- Plomin, R., & Petrill, S. A. (1997). Genetics and intelligence: What's new? Intelligence, 24, 53-77.
- Plotkin, H. C. (1988). Behavior and evolution. In H. C. Plotkin (Ed.), The role of behavior in evolution (pp. 1-17). Cambridge, MA: MIT Press.
- Plotkin, H. (1998). Evolution in mind: An introduction to evolutionary psychology. Cambridge, MA: Harvard University Press.
- Politzer, G., & Noveck, I. A. (1991). Are conjunction rule violations the result of conversational rule violations? Journal of Psycholinguistic Research, 20, 83-103.
- Pollard, P., & Evans, J. St. B. T. (1987). Content and context effects in reasoning. American Journal of Psychology, 100, 41-60.
- Pollock, J. L. (1991). OSCAR: A general theory of rationality. In J. Cummins & J. L. Pollock (Eds.), Philosophy and AI: Essays at the interface (pp. 189-213). Cambridge, MA: MIT Press.
- Pollock, J. L. (1995). Cognitive carpentry: A blueprint for how to build a person. Cambridge, MA: MIT Press.
- Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Eds.), Information processing and cognition: The Loyola Symposium (pp. 55-85). New York: Wiley.
- Reber, A. S. (1992). An evolutionary context for the cognitive unconscious. Philosophical Psychology, 5, 33-51.
- Reber, A. S. (1993). Implicit learning and tacit knowledge. New York: Oxford University Press.
- Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and Explicit Learning: Individual Differences and IQ. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 888-896.
- Rode, C., Cosmides, L., Hell, W., & Tooby, J. (1999). When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory. Cognition, 72, 269-304.
- Rosenthal, R., & Rosnow, R. L. (1991). Essentials of behavioral research: Methods and data analysis (Second Edition). New York: McGraw-Hill.
- Rozin, P. (1996). Towards a psychology of food and eating: From motivation to module to model to marker, morality, meaning and metaphor. Current Directions in Psychological Science, 5(1), 18-24.
- Rozin, P., & Fallon, A. E. (1987). A perspective on disgust. Psychological Review, 94, 23-41.
- Samuels, R. (1998). Evolutionary psychology and the massive modularity hypothesis. British Journal for the Philosophy of Science, 49, 575-602.
- Samuels, R., Stich, S. P., & Tremoulet, P. D. (1999). Rethinking rationality: From bleak implications to Darwinian modules. In E. Lepore & Z. Pylyshyn (Eds.), What is cognitive science? (pp. 74-120). Oxford: Blackwell.
- Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. Journal of Experimental Psychology: General, 110, 101-120.
- Shafir, E., & Tversky, A. (1995). Decision making. In E. E. Smith & D. N. Osherson (Eds.), Thinking (Vol. 3) (pp. 77-100). Cambridge, MA: The MIT Press.
- Shanks, D. R. (1995). Is human learning rational? Quarterly Journal of Experimental Psychology, 48A, 257-279.

- Shepard, R. N. (1987). Evolution of a mesh between principles of the mind and regularities of the world. In J. Dupre (Eds.), The latest on the best: Essays on evolution and optimality (pp. 251-275). Cambridge, MA: MIT Press.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, *84*, 127-190.
- Skyrms, B. (1996). The evolution of the social contract. Cambridge: Cambridge University Press.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. Psychological Bulletin, *119*, 3-22.
- Sloman, S. A., & Stibel, J. M. (2001). Making the conjunction fallacy transparent: Explaining the frequency effect. Manuscript submitted for publication.
- Slovic, P. (1995). The construction of preference. American Psychologist, *50*, 364-371.
- Slugoski, B. R., & Wilson, A. E. (1998). Contribution of conversation skills to the production of judgmental errors. European Journal of Social Psychology, *28*, 575-601.
- Smith, S. M., & Levin, I. P. (1996). Need for cognition and choice framing effects. Journal of Behavioral Decision Making, *9*, 283-290.
- Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L. A. Hirschfeld & S. A. Gelman (Eds.), Mapping the mind: Domain specificity in cognition and culture (pp. 39-67). Cambridge: Cambridge University Press.
- Sperber, D. (1996). Explaining culture: A naturalistic approach. Oxford: Blackwell Publishers.
- Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. Cognition, *57*, 31-95.
- Sperber, D., & Wilson, D. (1986). Relevance: Communication and cognition. Cambridge, MA: Harvard University Press.
- Stanovich, K. E. (1999). Who is rational? Studies of individual differences in reasoning. Mahwah, NJ: Erlbaum.
- Stanovich, K. E. (2001). Reductionism in the study of intelligence: Review of "Looking Down on Human Intelligence" by Ian Deary. Trends in Cognitive Sciences, *5*(2), 91-92.
- Stanovich, K. E. (in press). The fundamental computational biases of human cognition: Heuristics that (sometimes) impair reasoning and decision making. In J. E. Davidson & R. J. Sternberg (Eds.), The psychology of problem solving. New York: Cambridge University Press.
- 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. (1998d). Who uses base rates and P(D/~H)? An analysis of individual differences. Memory & Cognition, *28*, 161-179.
- Stanovich, K. E., & West, R. F. (1999). Discrepancies between normative and descriptive models of decision making and the understanding/acceptance principle. Cognitive Psychology, *38*, 349-385.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? Behavioral and Brain Sciences, *23*, 645-726.
- Stent, G. S. (1978). Introduction. In G. S. Stent (Eds.), Morality as a biological phenomenon (pp. 1-18). Berkeley, CA: University of California Press.



- Sterelny, K. (2001). The evolution of agency and other essays. Cambridge: Cambridge University Press.
- Sterelny, K., & Griffiths, P. E. (1999). Sex and death: An introduction to philosophy of biology. Chicago: University of Chicago Press.
- Stich, S. P. (1990). The fragmentation of reason. Cambridge: MIT Press.
- Sutherland, N. S., & Mackintosh, N. J. (1971). Mechanisms of animal discrimination learning. New York: Academic Press.
- Symons, D. (1992). On the use and misuse of Darwinism in the study of human behavior. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), The adapted mind (pp. 137-159). New York: Oxford University Press.
- Todd, P. M., Fiddick, L., & Krauss, S. (2000). Ecological rationality and its contents. Thinking and Reasoning, 6, 375-384.
- Todd, P. M., & Gigerenzer, G. (2000). Precipitous of Simple Heuristics that Make Us Smart. Behavioral and Brain Sciences, 23, 727-780.
- Tomasello, M. (1999). The cultural origins of human cognition. Cambridge, MA: Harvard University Press.
- Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), The adapted mind, (pp. 19-136). New York: Oxford University Press.
- Tversky, A. (1996). Contrasting rational and psychological principles of choice. In R. Zeckhauser, R. Keeney, & J. Sebenius (Eds.), Wise choices (pp. 5-21). Boston, MA: Harvard Business School Press.
- Tversky, A., & Edwards, W. (1966). Information versus reward in binary choice. Journal of Experimental Psychology, 71, 680-683.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review, 90, 293-315.
- Wason, P. C. (1966). Reasoning. In B. Foss (Eds.), New horizons in psychology (pp. 135-151). Harmondsworth, England: Penguin.
- Wasserman, E. A., Dorner, W. W., & Kao, S. F. (1990). Contributions of specific cell information to judgments of interevent contingency. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 509-521.
- West, R. F., & Stanovich, K. E. (2002). Is probability matching smart? Associations between probabilistic choices and cognitive ability. Manuscript submitted for publication.
- Whiten, A., & Byrne, R. W. (Eds.) (1997). Machiavellian intelligence II: Extensions and evaluations. Cambridge: Cambridge University Press.
- Williams, G. C. (1985). A defense of reductionism in evolutionary biology. Oxford Surveys in Evolutionary Biology, 2, 1-27.
- Williams, G. C. (1988). Huxley's Evolution and Ethics in sociobiological perspective. Zygon, 23, 383-407.
- Williams, G. C. (1992). Natural selection: Domains, levels and challenges. Oxford: Oxford University Press.
- Wright, R. (1994). The moral animal: Evolutionary psychology and everyday life. New York: Vintage Books.

## Footnotes

1. In probability learning, or choice situations, both animals and humans generally approximate probability matching when reinforcement is delivered on variable-interval schedules. However, things appear to be more complex when reinforcement is delivered on variable-ratio schedules. Although humans still tend to approximate probability matching (Estes, 1964, 1976, 1984), animals often maximize (Herrnstein & Loveland, 1975; MacDonall, 1988; but see Gallistel, 1990; Graf, Bullock, & Bitterman, 1964; Sutherland & Mackintosh, 1971).
2. We strongly caution that the term "bias" is used throughout this chapter to denote "a preponderating disposition or propensity" (The Compact Edition of the Oxford Short English Dictionary, p. 211) and not a processing error. That a processing bias does not necessarily imply a cognitive error is a point repeatedly emphasized by the critics of the heuristics and biases literature (Funder, 1987; Gigerenzer, 1996a; Hastie & Rasinski, 1988; Kruglanski & Ajzen, 1983), but in fact it was always the position of the original heuristics and biases researchers themselves (Kahneman, 2000; Kahneman & Tversky, 1973, 1996; Tversky & Kahneman, 1974). Thus, the use of the term bias here is meant to connote "default value" rather than "error". Under the assumption that computational biases result from evolutionary adaptations of the brain (Cosmides & Tooby, 1994b), it is likely that they are efficacious in most situations.
3. It should be noted that the distinction between evolutionary and instrumental rationality is different from the distinction between rationality<sub>1</sub> and rationality<sub>2</sub> utilized by Evans and Over (1996). They define rationality<sub>1</sub> as reasoning and acting "in a way that is generally reliable and efficient for achieving one's goals" (p. 8). Rationality<sub>2</sub> concerns reasoning and acting "when one has a reason for what one does sanctioned by a normative theory" (p. 8). Because normative theories concern goals at the personal level, not the genetic level, both of the rationalities defined by Evans and Over (1996) fall within what has been termed here instrumental rationality. Both concern goals at the personal level. Evans and Over (1996) wish to distinguish the explicit (i.e., conscious) following of a normative rule (rationality<sub>2</sub>) from the largely unconscious processes "that do much to help them achieve their ordinary goals" (p. 9). Their distinction is between two sets of algorithmic mechanisms that can both serve instrumental rationality. The distinction we draw is in terms of levels of optimization (at the level of the replicator itself--the gene--or the level of the vehicle); whereas theirs is in terms of the mechanism used to pursue personal goals (mechanisms of conscious, reason-based rule following versus tacit heuristics).
4. We will continue the practice here of using the metaphorical language about genes having "goals" or "interests" in confidence that the reader understands that this is a shorthand only. As Blackmore (1999) notes, "the shorthand 'genes want X' can always be spelled out as 'genes that do X are more likely to be passed on.'" (p. 5) but that, in making complicated arguments, the latter language becomes cumbersome. Thus, we will follow Dawkins (1976) in "allowing ourselves the licence of talking about genes as if they had conscious aims, always reassuring ourselves that we could translate our sloppy language back into respectable terms if we wanted to" (p. 88). Dawkins points out that this is "harmless unless it happens to fall into the hands of those ill-equipped to understand it" (278) and then proceeds to quote a philosopher smugly and pedantically admonishing biologists that genes can't be selfish any more than atoms can be jealous. We trust, Dawkins' philosopher to the contrary, that no reader needs this pointed out.
5. Frequency representations of probabilistic information are one example. However, although it is claimed that frequentist representations can eliminate cognitive illusions (Gigerenzer, 1991), this claim remains controversial. Over (this volume), for example, presents evidence suggesting that the favorable evidence with respect to frequentist representations may have resulted from the use

of problems with transparent logical forms. Furthermore, even if frequency representations sometimes attenuate cognitive illusions, they do not remove them entirely (Evans, Simon, Perham, Over, & Thompson, 2000; Girotto & Gonzalez, 2001; Harries & Harvey, 2000; Macchi, 1998; Mellers et al., 2001; Over, this volume; Sloman & Stibel, 2001; Tversky & Kahneman, 1983).

6. Deary (2000) has written a book summarizing the evidence on the relationship between reaction time and other speeded tasks and intelligence. It turns out to be difficult to explain why elementary information processing tasks correlate with intelligence at all. Stanovich (2001) conjectured that it is not because they measure some inherent "mental speed" (Deary reviews evidence indicating that the RT-IQ relationship is virtually unchanged when differences in nerve conduction speed are partialled out). But the speed component of these IP tasks may not be the critical thing. Rather, they all may serve as indirect indicators of the computational power available in the brain's connectionist network--computational power that is available to sustain the simulation of a serial processor. Of course, there are other more direct indicators of the computational power available to sustain serial simulation, such as working memory, and not surprisingly these indicators show larger correlations with intelligence.
7. Although the outcome of disputes about whether general intelligence is a byproduct or adaptation does not alter our argument, it should be noted that theorists such as LaCerra and Bingham (1998) and Foley (1996) argue that the changing online requirements of the ancestral hominid environment would, contra the massive modularity thesis, have required a flexible general intelligence (see also, Nozick, 1993, p. 120, for a philosophically-oriented version of a similar argument).
8. This is opposed to what they call evoked culture, which to Tooby and Cosmides (1992) is merely domain-specific mechanisms being triggered by local circumstances (culture on a short-leash as Lumsden & Wilson, 1981, have argued; see also Sperber, 1996).
9. Indeed, genetic engineering for purposes of human health and longevity is perhaps the ultimate triumph of Dawkins' (1976) so-called "survival machines" (the human vehicles) over their creators--the replicators. With the technology of genetic engineering, we, who were built by the replicators to serve as their survival machines, use them for our own goals--goals that are not the genes' goals (e.g., survival past our reproductive years). Williams (1988) uses such an example to counter Stent's (1978) argument against Dawkins (1976) that rebelling against one's own genes is a contradiction. Williams (1988) notes that Stent "apparently missed the relevance of major technologies (hair dyeing, tonsillectomy, etc.) based on such rebellion" (p. 403).
10. Those not committed a priori to a relativistic denial of the notion of cultural advance might well argue that the history of civilization reflects just this trend (the emancipation of women and the control of our reproductive lives come immediately to mind).
11. To be precise, we are doubting whether there are people who say they value their genome and have an accurate view of what they are valuing when they say this. For example, in such a case, the person would have to be absolutely clear that valuing your own genome is not some proxy for valuing your children; be clear that having children does not replicate one's genome; and be clear about the fact that the genome is a subpersonal entity.

Table 1

Mean SAT Total Scores as a Function of Response Given on a Selection Task Using the Destination Rule (Number of Subjects in Parentheses)

P,NQ (correct)	1190	(24)
P	1150	(38)
All	1101	(21)
P,Q	1095	(144)
P,Q,NQ	1084	(14)
Other	1070	(53)

Table 2

Mean SAT Total Scores as a Function of Strategy Choice on a Probabilistic Contingency Problem (Number of Subjects in Parentheses)

Strategy A	1151	(15)
Strategy B	1163	(64)
Strategy C*	1160	(168)
Strategy D**	1215	(150)
Strategy E	1148	(48)

\* = the probability matching response

\*\* = the normatively correct utility maximizing response

Table 3  
The Terms for the Two Systems Used by a Variety of Theorists and the Properties of Dual-Process Theories of Reason

	System 1 (TASS)	System 2 (Analytic System)
Dual-process theories:		
Sloman (1996)	Associative system	Rule-based system
Evans & Over (1996)	Tacit thought processes	Explicit thought processes
Evans (1984, 1989)	Heuristic processing	Analytic processing
Evans & Wason (1976)	Type 1 processes	Type 2 processes
Reber (1993)	Implicit cognition	Explicit learning
Levinson (1995)	Interactional intelligence	Analytic intelligence
Epstein (1994)	Experiential system	Rational system
Pollock (1991)	Quick & inflexible modules	Intellection
Klein (1998)	Recognition-primed decisions	Rational choice strategy
Johnson-Laird (1983)	Implicit inferences	Explicit inferences
Fodor (1983)	modular processes	central processes
Chaiken, Liberman, & Eagly (1989)	Heuristic processing	Systematic processing
Gibbard (1990)	Animal control system	Normative control system
Norman & Shallice (1986)	contention scheduling	supervisory attentional
Shiffrin & Schneider (1977)	Automatic processing	Controlled processing
Posner & Snyder (1975)	Automatic activation	Conscious processing
Properties:	Associative Holistic Automatic Relatively undemanding of cognitive capacity Relatively fast Acquisition by biology, exposure, and personal experience Highly contextualized	Rule-based Analytic Controlled Demanding of cognitive capacity Relatively slow Acquisition by cultural and formal tuition Decontextualized
Goal structure	Short-leash genetic goals that are relatively stable	Long-leash goals that are utility maximizing for the organism and constantly updated because of changes in environment

## Figure Captions

Figure 1. Dennett's (1996) Tower of Intellect.

Figure 2. Nature of processing control when a System 1 response is overridden by System 2.

Figure 3. Goal structure of a Darwinian creature. The areas indicate overlap and nonoverlap of vehicle and genetic "interests".

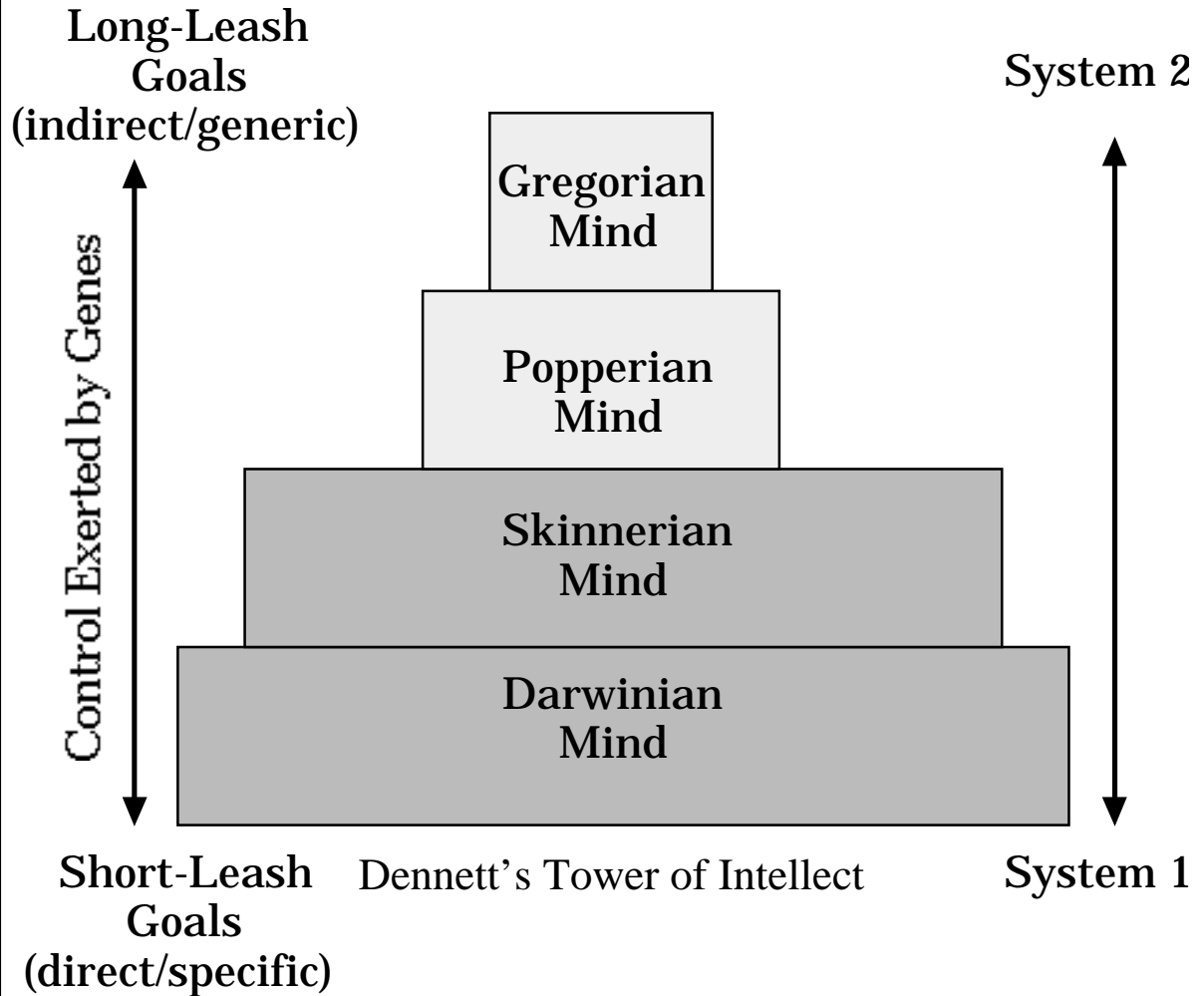
Figure 4. The logic of the goal structure in a human.

Figure 5. Genetic and vehicle goal overlap in the two systems.

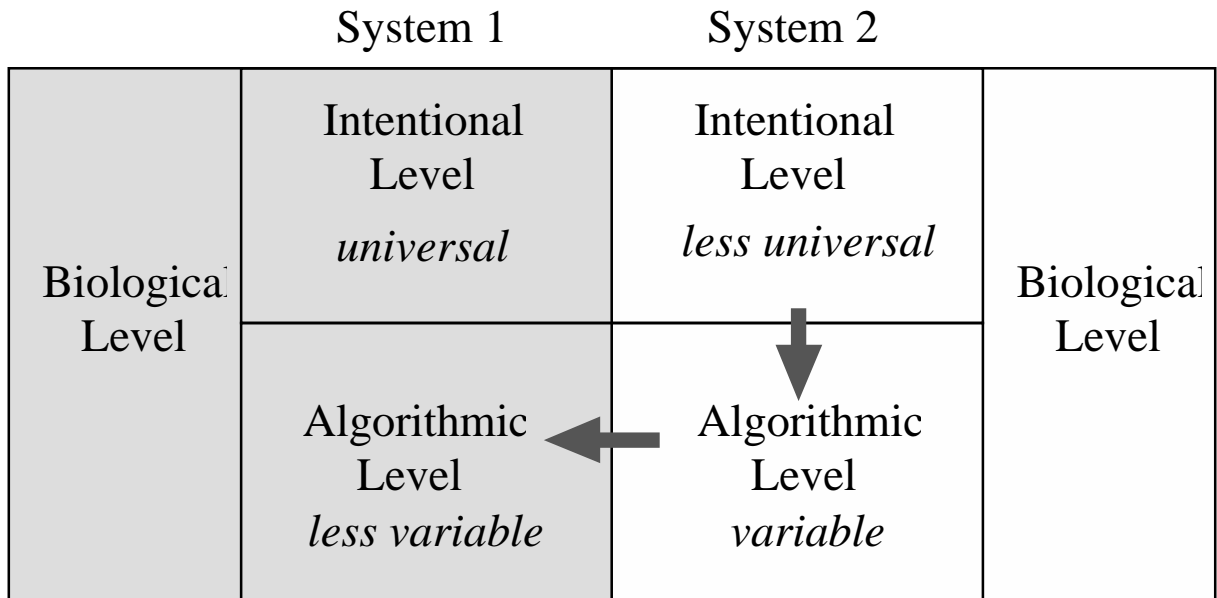
Figure 6. Hypotheses about the way that gene-installed goals and meme-installed goals are distributed across System 1 and System 2.

# Four Kinds of Minds

(Simultaneously Operative)

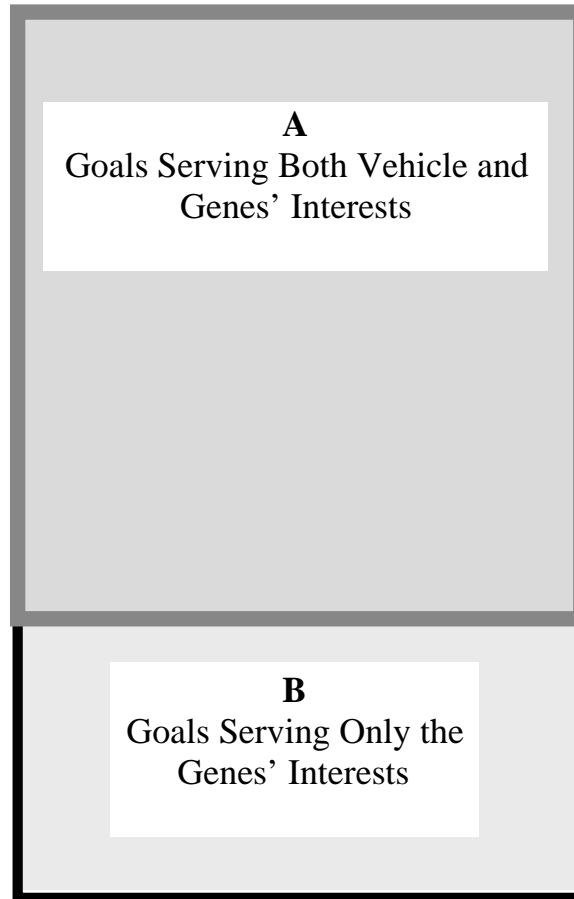


# Control Structures

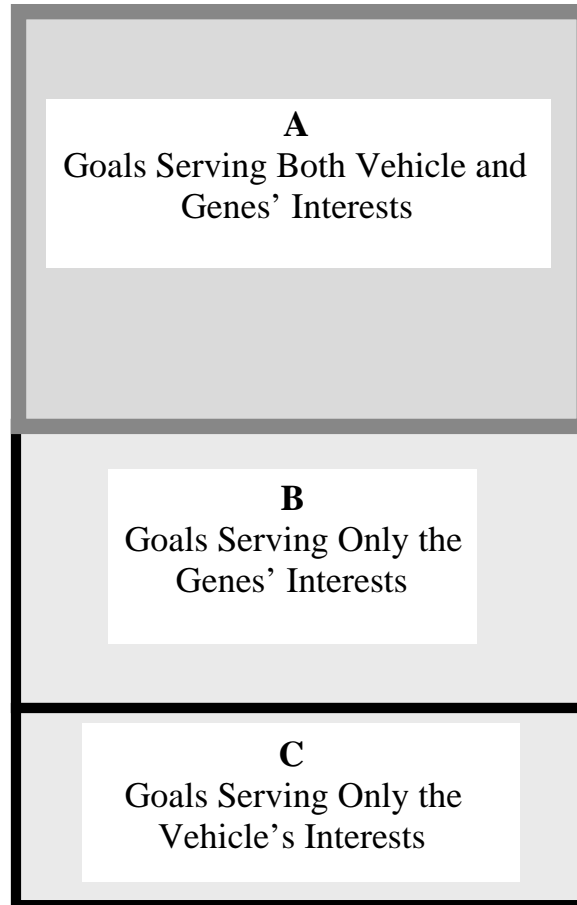




## Goal Structure Darwinian Creature



## Goal Structure Humans



### Goals Structure

