To appear in: *Evolutionary Psychology: Foundational Papers*, by John Tooby & Leda Cosmides, MIT Press: Cambridge.

Tooby, J. & Cosmides, L. (1992). Ecological rationality and the multimodular mind. Center for Evolutionary Psychology Technical Report #92-1. Santa Barbara, CA: University ofo California, Santa Barbara.

Ecological Rationality and the Multimodular Mind:

Grounding normative theories in adaptive problems

John Tooby and Leda Cosmides

John Tooby
Dept. of Anthropology
University of California
Santa Barbara, CA 93106 USA

Leda Cosmides
Dept. of Psychology
University of California
Santa Barbara, CA 93106 USA

The paradox of human competence

Modern research into human reasoning and decision-making has produced a formidable paradox. The consensus among many modern reasoning and decision researchers, particularly within the heuristics and biases school, is that the faculty of human reasoning is riddled with crippling defects. It is thought that -- in the absence of explicit training and, usually, even after training in formal methods -- human reasoning and decision-making is governed by crude and error-prone heuristics. According to this view, the design of human reasoning mechanisms bears little or no resemblance to the logician's or decision theorist's ideal. Indeed, these prescriptive ideals or normative theories are used to provide the standard or metric against which human reasoning has been measured, and found wanting. In this research tradition, reasoning errors are defined as having occurred whenever performance departs in any respect from what experimenters decide these theories dictate. These normative theories include the propositional calculus, hypothesis testing by falsification, Bayes' Rule, sampling theory, and other standards drawn from logic and mathematics. As Kahneman and Tversky put it, "The presence of an error in judgment is demonstrated by comparing people's responses either with an established fact (e.g., that the two lines are equal in length) or with an accepted rule of arithmetic, logic, or statistics" (1982, p. 123). When such standards have been used to evaluate human performance on problems that require deductive reasoning, substantial evidence supports the conclusion that the human mind does not embody the rules of inference of the propositional calculus (Wason & Johnson-Laird, 1972; Johnson-Laird, 1982). Similarly, on the basis of a large set of experimental findings, many believe that inductive reasoning does not operate according to Bayes' Rule or any other widely accepted calculus of probability (Tversky & Kahneman, 1974). In fact, reviews of human reasoning research are customarily presented as lengthy catalogs of "errors," "fallacies," and "biases" -- and the standard textbook treatments organize their discussions around the seemingly endless ways in which human reasoning and decision-making depart from experimenters' normative ideals of rationality.

Yet despite this apparent human ineptitude on laboratory-administered artificial problems, natural reasoning systems -- human and nonhuman minds alike -- negotiate the complex natural tasks of their world with a level of operational success far surpassing that of the most sophisticated existing artificial intelligence systems. On virtually every natural inferential problem that has been carefully investigated -- from grammar induction, semantic induction, and speech perception to vision, object recognition, and color constancy -- organisms perform better than the systems that cognitive scientists have been able to construct, even though these scientists have had full access to modern logics, statistical decision theories, and other formal methods of inference. Equally striking, evolutionary biologists have found that animals with truly miniscule nervous systems, such as bees, make judgments under uncertainty during foraging that manifest exactly the kind of well-calibrated statistical induction that the human brain is widely thought of as "too limited" to perform (e.g., Real, 1991; Staddon, 1988).

Thus, we have two widely contrasting sets of findings that present an apparent contradiction. If our present handful of normative theories about how reasoning should be conducted are in fact the appropriate standards of ideal design, then why aren't artificial systems that are equipped with them truly intelligent?[1] These normative theories have consistently proven insufficient to construct systems that perform powerfully on the natural tasks that humans and other species do well. And if humans are not, in fact, equipped with these "ideal" systems of knowledge production, then how can we routinely solve problems that bedevil the artificial systems that *are* equipped with them? If the consensus view in psychology is correct -- if humans are equipped with reasoning mechanisms that operate according to fallacies and defective principles -- then we should be consistently outperformed by systems equipped with superior principles. We are not. It follows that either (1) human reasoning is, for some reason, better than current experimental findings indicate, (2) some other set of reasoning principles are superior to those "accepted rule[s] of arithmetic, logic, or statistics" which are currently used to define human experimental responses as errors, or (3) both.

The paradoxical state of the literature raises two linked questions:

(1) Are the normative theories that researchers have been using as their standard of ideal performance the appropriate normative theories to use to evaluate and understand the structure of human reasoning?

(2) Is human reasoning performance truly as weak and error-prone as the large existing experimental literature indicates?

In this chapter we argue that, despite widespread belief to the contrary, human reasoning may be generated by a very well-engineered collection of mechanisms, and that reasoning only seems to be governed by crude and simple processes of primitive design because it has been evaluated using the wrong normative theories and experimental paradigms (Cosmides, 1989; Cosmides & Tooby, 1989, 1992, under review; Gigerenzer, 1991; Gigerenzer, Hell & Blank 1988; Gigerenzer, Hoffrage & Kleinbolting, 1991; Gigerenzer & Hug, 1992; Gigerenzer & Murray, 1987). According to this view, a highly sophisticated collection of human problem-solving mechanisms has been nearly invisible to standard experimental approaches for a series of reasons, of which we will mention three.

First, if the mind has a constellation of diverse evolved reasoning specializations (for social exchange inference, syntax induction, object recognition, threat intepretation, hazard avoidance, rigid object mechanics, belief/desire analysis of other minds, event frequency tracking, and so on), then each task-appropriate specialization should have a subsystem that activates it only when -- or primarily when -- it detects cues indicating that a problem of the relevant type is present. Mechanisms will remain dormant -- and, therefore, undetectable by experimental means -- unless subjects are exposed to the cues that activate them. Experimental protocols must pair domain-defined contents (such as threats) with matching tasks (such as bluff detection), if specialized inference engines are to be discovered. Moreover, experimenters must be sensitive to the fact that mechanisms may misfire if subjects are presented with problems of one type -- say, a problem requiring the detection of mistakes -- when these problems are accompanied by cues indicating a different problem-type -- say, a problem requiring the detection of cheaters (Cosmides & Tooby, 1992). Such mismatches between cues and problem types are far more likely under artificial experimental conditions than under natural conditions. Reasoning researchers rarely consider what the issue of ecological validity may mean in experimental contexts because most lack a theory of what ecological validity would mean with respect to reasoning.

Second, we join Gigerenzer in suggesting that human reasoning mechanisms respond far more sensitively to important but subtle features of experimental situations and of the world than most researchers expect or look for (Gigerenzer, 1991). When subjects are appropriately sensitive to dimensions of the experimental situation that the researcher is ignoring, their performance can appear erratic, unmotivated, and irrational. For example, subjects' nonconscious mechanisms may need exposure to ecologically valid cues that a sample is in fact random before they will take a base rate derived from that sample into account in Bayesian decision problems (for evidence supporting this view, see Gigerenzer, Hell & Blank, 1988). If this is true, many "failures" to use base rates would result from a sophisticated design feature of a nonconscious statistical inference competence, rather than from a design defect or from the operation of a crude heuristic.

Third, evolved cognitive mechanisms have been shaped to conform to normative standards that differ sharply from those familiar to most psychologists. Most experiments have not been designed to detect conformity to these very different standards. For example, on reasoning problems involving social exchange, the answer that is adaptively correct often differs from the answer that is logically correct (Cosmides, 1989; Cosmides & Tooby, 1992; Gigerenzer & Hug, 1992). In this case, an effective problem-solving design would be evaluated as "irrational" and defective when it is, in fact, a superior computational design for solving adaptive problems involving social exchange. Just as in classical times, Greeks judged non-Greeks to be "barbarians" because they spoke Greek poorly, so now we may be judging human reason harshly because our criteria for recognizing sophisticated performance have been parochial. If the mind speaks other languages of rationality than those we are familiar with, we need to learn what these languages are and discover their functional significance. To track down these alternative forms of rationality, we need to analyze what types of rationality the evolutionary process should have built into the human mind.

Ecological rationality:
Devices designed to navigate through a stably structured world

We suggest that our cognitive mechanisms are primarily designed to embody what might be called *ecological rationality* rather than the more traditional methods of rational decision-making that have been formally developed by mathematicians, philosophers, and decision theorists to analyze problems of widely ranging contents and contexts. From an evolutionary functionalist perspective, human task environments have tended to manifest recurrent configurations of conditions, enduring relationships, and stable statistical distributions of properties -- what may be termed *ecological structure*. Ecologically rational mechanisms contain design features that anticipate and exploit this recurring structure to greatly augment their problem-solving power -- features such as content-sensitive algorithms, topic-specialized representational formats, and conceptual primitives. Many statistical and structural relationships that endured across human evolution were "detected" by the process of natural selection, because designs that were coordinated with these enduring relationships successfully solved adaptive problems that designs lacking them could not. Mutations that injected such advantageous reasoning specializations into ancestral cognitive architectures caused them to make better decisions within those ecological circumstances, and hence they outpropagated alternative designs. In this way, natural selection built computational machinery whose problem-solving strategies corresponded to the demands and opportunities created by these regularities, as well as to other persistent features of adaptive problems. Devices that embody such problem-solving strategies can go far beyond the fragmentary information provided by monitoring immediate external circumstances -- and far beyond what can be discovered by operating on such data with general principles alone.

Ecologically rational computational devices can generate effective decisions and reconstruct reliable knowledge that would have been unobtainable by any methods that did not

an example here

incorporate specializations designed to exploit this encompassing ecological structure (Cosmides & Tooby, 1987; Tooby & Cosmides, 1992). The key point is that the world contains relationships that are logically arbitrary, but that happen to be both true and stable over evolutionary time (objects are solid, the world is three dimensional, human language reflects universal grammatical patterns, etc.). Under natural circumstances, "problems" not only have a logical structure, but also come clothed or embedded in a rich set of additional relationships -- what Gigerenzer (1991) has called "surplus structure" -- that permit many more computational pathways to solutions. To limit inference to the logically derivable is to radically debilitate the cognitive architecture. Through the incorporation of content-specific reasoning principles, inference engines can be designed that make at least as much use of these logically arbitrary relationships as of logically nonarbitrary ones. Indeed, perceptual systems are known to depend on -- in fact, to consist of -- such specialized inferential components; removing such content-specific inference mechanisms from our cognitive architecture would leave us blind, not only figuratively but literally.

Traditional content-independent and goal-neutral rational methods ignore this recurring structure because (among other reasons) they have been developed by modern humans to be applicable to as wide a range of culturally novel problems as possible. Ecologically rational methods -- however powerful -- are designed to perform well only on problems whose forms reflect enduring evolutionary relationships. Thus, the traditional drive among logicians and statisticians for increasingly general-purpose methods of inference may have made good sense given their particular aims. After all, a central ambition was to develop formal analytic tools that could be used to tackle the inherently unpredictable array of new scientific problems, including hypothetical or counterfactual problems. Sensible or not, this push towards universality and generality has been so strong that at least since Leibniz some have been tempted to claim that certain rational methods should hold in any "possible universe."

In starkly provincial contrast, however, ancestral humans not only evolved solely in this particular universe, but on this planet, embedded in a narrowly delimited and densely structured social, physical, and biotic environment. Moreover, they were not bombarded with a random sample out of the universe of all possible problems. Instead, ancestral humans encountered clustered distributions of adaptively consequential problem types -- such as foraging, vision, language acquisition, cooperation, threat assessment, hazard avoidance, rigid object manipulation, face recognition, incest avoidance, wayfinding, and coalition formation -- within strongly structured task-environments. Given the problems and conditions characteristic of ancestral hunter-gatherer life, selection would not have assembled cognitive devices primarily or exclusively designed to solve computational problems of the widest possible generality, defined in the most abstract possible terms (Cosmides & Tooby, 1987, in press; Tooby & Cosmides, 1992; Tooby & DeVore, 1987).

cannot be expected to have

Instead, hominid cognitive mechanisms and their constituent design features would have been selected on the basis of how well they solved the inferential problems that they were actually exposed to, in proportion to the frequency they were encountered and their adaptive importance. For these reasons, we expect that the hominid-encountered ecological structure of adaptive tasks should not be irrelevant to the evolved reasoning mechanisms of the human mind. Rather, evolution should have created a mesh between the computational principles of the human mind and the regularities of the world (Shepard, 1984, 1987a). The organizational specifics of our evolved inferential devices should have evolved to complement the problem-relevant pieces of recurrent ecological structure. For example, the organizational specifics of Chomsky's proposed "language acquisition device" must, in effect, assume that the local language to be acquired exhibits certain unobservable grammatical relationships universal to all human language communities. Without such assumption-embodying inferential procedures, language acquisition would be an unsolvable computational problem for the child (Chomsky, 1959; Pinker, 1979, 1984, 1989). According to this ecological perspective on rationality, then, one should not necessarily expect to discover the logic of the propositional calculus in the human mind, but rather procedures that embody a logic of rigid object mechanics, a logic of social exchange, a logic of syntax induction, a logic of contagion, a logic of foraging, a logic of threat, a logic of trajectories in three dimensions, and so on.

The subsets of this enduring ecological structure that are relevant to particular problem-families constitute orderly domains that can be efficiently analyzed by devices configured to mesh with them. Thus, face-recognition devices should have design features that mesh with the enduring relationships in the visual arrays presented by human faces; bluff detection mechanisms should reflect the recurring relationships and goals present in situations of threat; dead reckoning mechanisms should reflect the spatial, dynamic, and temporal relationships inherent in movement across surfaces; food selection mechanisms should reflect facts such as the relative value of various nutrients, chemical cues indicating their presence, and the physiological lag time between ingestion and the detection of toxins in the bloodstream. In short, ecologically rational devices can be designed -- and, we argue, would inevitably have been designed by natural selection -- to exploit this cross-generationally recurrent ecological structure in order to solve the adaptively consequential inferential problems our ancestors faced. As we have argued elsewhere (Cosmides & Tooby 1987; Tooby & Cosmides 1992), such methods could solve these problems far more reliably and efficiently than more general methods could alone, even on those rare occasions when general methods could solve them at all. Indeed, general techniques will rarely prove to be the best problem-solving strategy when problems are predictable in structure.

Thus, the claim that the human mind exhibits ecological rationality is the claim that our species-typical cognitive architecture has specialized inferential devices that are designed to mesh with the enduring regularities of task-environments and task demands, in addition to whatever general competences it may have (Cosmides & Tooby, 1987; Gigerenzer, 1991; Shepard, 1987a; Tooby & Cosmides, 1992). Views emphasizing the importance of the structure of the environment were pioneered earlier in this century by psychologists such as Brunswik (1950, 1955) and Gibson (1966, 1979). More recently, within the reasoning community, Gigerenzer and his colleagues have spearheaded an effort to reorganize the study of inductive reasoning and statistical inference along these lines. By considering the structure of environments and by applying a more sophisticated knowledge of the broad variety of statistical theories than is customary among cognitive psychologists, Gigerenzer has been able to both reinterpret prior subject performance as potentially more rational than is usually believed and to successfully predict well-calibrated performance on new experiments (Gigerenzer, 1991; Gigerenzer, Hell, & Blank, 1988; Gigerenzer, Hoffrage & Kleinbolting, 1991; Gigerenzer & Murray, 1987; see also Cosmides & Tooby, under review). In the area of deductive reasoning, we and others have argued that a wide range of content and context effects can be explained by hypothesizing the existence of numerous domain-specific inference engines that were designed to solve evolutionarily recurring problems such as social exchange and threat (Cosmides, 1985, 1989; Cosmides & Tooby, 1989, 1992; Gigerenzer & Hug, 1992; Tooby & Cosmides, in prep.).

Although this view is relatively rare within the community of reasoning and decision-theory researchers, it is has become increasingly prominent in many other fields of psychology.

Shepard, for example, has long defended an ecological view with respect to perception and psychophysics (Shepard, 1981, 1984, 1987a; see also Gibson, 1966, 1979; Neisser, 1982), and supported it with a wide-ranging array of results about, for example, the perception and representation of object motion (Shepard, 1981, 1984, 1987a; Freyd, 1987; see also Shiffrar & Freyd, 1990, for evidence of procedures specialized for the perception of biomechanical motions such as arm movement). Indeed, the inferential mechanisms subserving perception reliably construct interpretations that are far more accurate than is logically warranted solely on the basis of sensory information itself (Marr, 1982). They can do this because they incorporate specialized procedures that assume certain relationships are characteristic of the environment being perceived, and impose them on the interpretation. Researchers have found that to develop computational models of vision that are even minimally effective, they must build in such ecologically rational inferential procedures (Marr, 1982; Poggio, Torre, & Koch, 1985).

Of course, the central claim of Chomskyan psycholinguistics is that our species-typical cognitive architecture includes a battery of reasoning procedures specialized for syntax induction. These procedures are organized to exploit the fact that some grammatical relationships reliably occur in the minds of adult language users, whereas others do not (Chomsky, 1957, 1959, 1975, 1980; Pinker, 1979, 1982, 1984, 1989; Wexler & Culicover, 1980). Inferential devices that lack procedures specialized to exploit this ecological structure have repeatedly been shown to be inadequate; such devices cannot induce the correct grammar from the kinds of information children are routinely exposed to, because an infinite number of grammars are consistent with any finite set of utterances (Pinker 1989, 1991; Pinker & Prince, 1988; Wexler & Culicover, 1980).

More generally, the entire field of cognitive development is being revolutionized by this same organizing logic. A flood of recent findings supports the view that the reasoning principles infants and children spontaneously bring to the tasks of learning are organized to reflect the recurrent structure of specific problem domains, such as object construal and motion, the differences between artifacts and living kinds, physical causality, animacy, the nature of others' minds, social categorization, and so on (Astington, Harris, & Olson, 1988; Avis & Harris, 1991; Baron-Cohen, Leslie, & Frith, 1985; Brown, 1990; Flavell, Zhang, Zou, Dong & Qui, 1983; Gardner, Harris, Ohmoto & Hamazaki, 1988; Gelman, 1990; Gelman & Markman 1986, 1987; Keil, 1989; Hirschfeld, 1989; Leslie, 1987, 1988; Leslie & Keeble, 1987; Leslie & Thaiss, 1992; Markman, 1989; Perner, 1991; Premack, 1990; Spelke, 1988, 1990, 1991; Springer, 1992; Springer & Keil, 1991; Wellman, 1990; Wimmer & Perner, 1983; for an overview, see chapters in Carey & Gelman, 1991, Hirschfeld & Gelman, in press; and Volume 14 of *Cognitive Science*). Obviously, a "Piagetian" or other content-free rational architecture that had to laboriously discover the evolutionarily long-enduring structure of the world would be a poor and inept design compared to one that spontaneously organized its knowledge according to principles that, though non-deducible, are stably true (e.g., Shepard, 1987a). And the theory that inferential specializations govern much, if not all, nonhuman cognition and learning is by now firmly established (Garcia, 1990; Rozin & Schull, 1988; Staddon 1988; Gallistel, 1990).

The advantages accruing to devices designed to exploit ecological structure are enormous and varied: Any evolutionarily stable feature of the world -- including statistical relationships or orderly contingencies -- need not be deduced or detected within an individual's lifespan by an evolved problem-solving system. Evolution can build into the architecture procedures that simply assume and rely on the presence of such problem-relevant relationships, or at least "bet" on them. Such problem-solving strategies appropriately organize problem-spaces, immensely increase the internal store of information immediately available to derive inferences from, supply innumerable uncertainty-reducing constraints, allow detectable cues to signal the state of otherwise nondetectable states of affairs, and so on (Cosmides & Tooby, 1987, 1992, in press; Tooby & Cosmides, 1992). As long as such devices operate in the environments for which they were designed, they can radically outperform content-independent general-purpose rational problem-solvers (Tooby & Cosmides, 1992).

From the "accepted rules" to well-performing rules

Of course, the end result of an evolutionary process that favors efficient solutions to content-defined sets of adaptive problems is a collection of computational devices whose information-processing strategies often will not correspond to the "accepted rule[s] of arithmetic, logic, or statistics." Such devices may often operate using non-traditional inferential principles that diverge sharply from the "accepted rules." These principles may be content-specific and context-sensitive, reflecting the structure of the targeted problem-domain, the task-environment, and the task demands. Such devices may interpret problems in terms of contentful primitives (e.g., benefit, cheat, noun, verb phrase, edge, surface, false belief, desire, animate, solid object, person, snake, mother) that are linked to content-specific procedures. They may impose on problem-spaces a rich series of assumptions and categories. And, generally, they may look strange to those schooled purely in formal decision theory, inferential statistics, mathematics, and logic. (For a detailed example of what such content-sensitive principles of problem-solving might look like, see the proposals about the design of social contract algorithms in Cosmides, 1985; Cosmides and Tooby, 1989, 1992; and Gigerenzer & Hug, 1992.) Domain-specific methods may often have extremely elaborate, densely engineered designs, far different from the clean, simple, and content-free methods familiar to most cognitive psychologists.

On this view, traditional formal methods of inference are not the only methods of productive inference, nor even very good methods of inference for most recurrent human purposes. On the contrary, they are only a small subset out of a potentially unlimited class of methods that might be productively used if matched with appropriate problem-types. Traditionally recognized methods (e.g., the propositional calculus, Bayes' Rule) are special only in that they represent a few "degenerate" or limiting cases out of this larger class of methods. They are limiting cases in the following sense: The less that is known in advance about the problems to be solved, the more special principles and useful assumptions must be stripped away from a problem-solving design. In other words, the more *ignorant* the computational system is about the world and the problems it will be tackling, the less it can presume, and the more it must fall back on increasingly content-independent and context-general methods. Correspondingly, the more *knowledgeable* the system is about the world and the problems it will be encountering, the less often these content-independent methods will prove to be the most effective problem-solving designs. Arguably, content-independent methods are normatively correct (that is, most effective -- see below) only when the problem-solving system is exceptionally ignorant about the world it is facing and the problems it is confronting. They are not ideal methods that should be universally applied; instead, they should be applied only on those rare occasions when the organism is ~~very~~ almost completely ignorant of the problem and the context in which it is embedded (and, only when the organism "knows" that it is ignorant). It seems likely that such circumstances would have been relatively rare for our hunter-gatherer ancestors during human evolution.[2]

Nevertheless, despite their many drawbacks and limitations, the identification and refinement of universally applicable rational methods has been a central aspiration of Western philosophers and scientists for millenia, exerting a more powerful pull even than the search for universal scientific laws. To develop general inferential methods, philosophers and mathematicians fractionated the reasoning process into two categories -- "methods" and "content" -- which, if not entirely arbitrary inventions, are at least more flexible in boundary than is usually realized. In order to arrive at methods that could be applied generally across situations, all aspects of the reasoning process that vary from instance to instance were distilled out of the "methods" category and injected into the "content" category. Within this tradition, all situation-specific structure becomes "content" -- and content exists solely as arguments that are plugged into content-free formal operations.

Hence, the ambition to discover rational methods that can be applied as universally as possible has historically led to a fixation on domain-general and content-independent rational methods and to the relative eclipse of situation-specific rational methods. Of course, there exists a whole universe of potential alternative computational designs where at least some content is inherent in the structure of the procedures themselves, rather than merely plugged in as arguments. Examples can be found in such fields as artificial intelligence, psycholinguistics, and evolutionary psychology, and include grammar induction systems (Pinker, 1989), various computational models of vision (Marr, 1982; Poggio, et al., 1985), and the social contract algorithms we have proposed to explain human reasoning about social exchange (Cosmides & Tooby, 1989, 1992). But the historical impetus to develop broadly or universally applicable inferential methods has funnelled the development of inferential methods onto one narrow path, so that the great preponderance of effort has gone into the construction and elaboration of the family of content-free, domain-general formal methods.[3] The relative neglect of domain-specific inferential methods among logicians, statisticians, economists, and decision-theorists has led many psychologists into the mistaken belief that certain content-independent methods of problem-solving are the only rational methods that exist, that they are normatively prescriptive, and that their exclusive use even defines what it is to be rational (see, e.g., Tversky & Kahneman, 1974). For participants in this program, the question *Are humans rational?* is the easy to answer question, *Do humans use certain content-independent methods to reason?*, rather than the more complex question, *Which, if any, of the many effective procedures capable of solving recurrent adaptive problems are embodied in the human mind?*

The mirage of normatively self-justifying rational methods

The existence of innumerable alternative computational methods and the performance differences among them throw normative questions into a very different light. If there were only one possible set of rational computational methods, then their exclusive use could arguably be considered normatively prescriptive and the valid definition of rationality. Indeed, within the heuristics and biases school, rationality is considered to be conformity to certain privileged methods of arriving at judgments, rather than the success or accuracy of the judgments themselves by an independent standard. In experiments conducted within this research program, the "correct" answer is defined as the answer that results when the method believed to be normative is applied to the problem, regardless of whether the judgment leads to a successful consequence in the world. Computational principles that do not conform to the "accepted rules" are by this standard automatically and definitionally irrational and defective, no matter how well they may perform on average at problem-solving on various real world tasks. It is as if winners in archery contests were selected not by who hit the target most often but by who displayed the judges' idea of the best form while aiming. Those whose targets may be bristling with arrows may nevertheless lose the contest, because the judges never even look at the targets. These method-based definitions of the normative are what have allowed the present paradoxical situation to arise, in which the mechanisms of the human mind often outperform normatively ideal formal problem-solving systems on many natural real-world tasks, yet are nevertheless judged to be normatively defective because they use non-standard methods to achieve their often superior outcomes. Given that there are a wealth of alternative methods, we suggest that there is a better and more coherent framework for addressing normative questions. Instead of defining a method as normative a priori, the normative standing of a method should be assessed by gauging its operational success in attaining an independently defined goal. Only in this way can alternative methods even be compared.

Before proceeding further, it will be useful to distinguish two distinct but often conflated normative questions. One normative question that could be asked about humans (or, indeed, any problem-solving system) is, What problem-solving method is normative? That is, what method should the person or system be using? A very different question that also needs to be asked is, What justifies the normative theory that is being used to define some outcomes as successes or solutions and other outcomes as failures or errors? The first question asks what is the best means of attaining goals; that is, Given that we know that the good is, what should we do to accomplish it? The second question asks what justifies selecting one potential set of outcomes over others as the goal that the system should seek; that is, What is the good? The sensible ordering of such issues is to define the goal first and then, second, to determine what method works best in reaching it.

Within the heuristics and biases community, however, the means-normative question has entirely swallowed up the goals-normative question through the act of defining the goal of reasoning as adhering to the set of inferential methods prejudged to be normative. (And the method usually treated as unquestionably normative is one particular interpretation of Bayesian statistical inference; Gigerenzer, 1991). Within this paradigm, alternative methods simply cannot be evaluated on the basis of how well they solve problems; there are no independent criteria for successful problem-solving other than adherence to the preselected methods. Thus, the "good" is whatever decisions are reached by applying these methods, regardless of what consequences they may lead to in the real world.

This odd approach looks reasonable when practiced because natural problems and real world consequences are rarely if ever considered in such reasoning experiments. Artificial problems are constructed and evaluated solely on the basis of their formal or "logical" structure, with the surplus structure they would ordinarily reflect in the real world stripped away or ignored. The fact that subjects respond so differently to problems of the same "logical" form but with different contents is taken to indicate irrationality, when it may just as easily indicate the activation of different domain-specific mechanisms, each of which was designed to exploit the different ecological structures that characterize different content domains (Cosmides, 1989; Cosmides & Tooby, 1989, 1992).

In any case, the entire notion that there exists a goal-independent normative ideal depends for its coherence on there being one and only one method that can be sensibly applied to each problem (see discussion in Gigerenzer & Murray, 1987; Gigerenzer, 1991). If the "correct" answer can only be defined as the answer that results when the sole applicable method is used, then a pluralism of methods cannot be tolerated. But if there were a multiplicity of accepted normative methods, then a problem would have many inconsistent but equally "correct" answers, a situation that advocates of a unitary rationality reject. Lacking any principled system for justifying the use of one method over another, the heuristics and biases paradigm depends on there being an authoritative consensus among methodologists privileging only one method for each problem. Indeed, researchers in the heuristics and biases program do often speak as if there existed a single monolithic system for statistical induction and for other forms of reasoning. They routinely and unproblematically use phrases such as "the normative theory of prediction" (Kahneman & Tversky, 1973, p. 243), the "normative principles of statistical prediction" (Ajzen, 1977, p. 304), "the calculus of chance or the statistical theory of prediction" (Kahneman & Tversky, 1973, p. 237), or "an accepted rule of arithmetic, logic, or statistics" (Kahneman & Tversky, 1982, p. 123). As Nisbett and Ross put it, "We follow conventional practice by using the term `normative' to describe the use of a rule when there is a consensus among formal scientists that the rule is appropriate for a particular problem" (Nisbett & Ross, 1980, p. 13).

But, as discussed above, there is an indefinitely large class of non-traditional computational methods that could be applied to problems, and these would inevitably produce many different answers to the same problem. Without independent standards of successful problem-solving, there is no way to decide which of these many answers and methods is genuinely "normative" -- as opposed to normative by social convention. Moreover, even if one arbitrarily excludes from consideration everything except those traditional methods already developed within the community of professional probabilists and statisticians -- principally content-independent, general-purpose methods -- one still cannot sensibly speak of any consensus about which formal methods are uniformly the best because the community is divided by a rich diversity of opinion on scores of issues (Gigerenzer & Murray, 1987; Hacking, 1965). Different professional statisticians looking at the same problem will often give completely different answers to it -- that is, they will make contradictory claims about which answer is normatively correct. an example of contradictory claims is needed This extends all the way up from minor methodological disagreements to disagreements encompassing the most major and fundamental issues in the field -- such as whether probability refers to a subjective degree of belief (the Bayesian interpretation favored by many within the heuristics and biases program) or a relative frequency of events (the frequentist interpretation favored by most professional statisticians). It is important to recognize that the mathematization of inductive reasoning over the last three centuries has still not solved Hume's puzzle; there is still no universally accepted or universally applicable solution to the problem of induction (Gigerenzer, Swijtink, Porter, Daston, Beatty, & Kruger, 1989). The mirage of the single true or privileged, all-encompassing, formally demonstrable system -- of which Euclidian geometry was for so long the prototype -- has finally disappeared in mathematics and logic over the last 150 years but, surprisingly, still survives in the minds of many psychologists.

Although the methods courses taught to psychologists and other scientists customarily present probability and statistics as if they constituted a single, integrated deductive system, exposure to the primary literature immediately reveals that what is taught to nonstatisticians represents a reduced and simplified stew of often contradictory approaches, methods, theories, and assumptions. As Gigerenzer puts it, the standards that many cognitive psychologists have been using as their normative yardstick to assess human reasoning are "a caricature of the present state of probability theory and statistics" (Gigerenzer, 1991, p. 103), ~~As he comments~~, "When claiming `errors' and `fallacies', cognitive and social psychologists have largely ignored conceptual and technical distinctions fundamental to probability and statistics" (Gigerenzer, 1991, p. 86). Although subjects' answers to problems may be errors according to some statistical theories, they are often perfectly correct according to others (Gigerenzer, 1991; Gigerenzer & Murray, 1987). Even so, substituting a more technically sophisticated version of modern statistical theory will not allow one to isolate a single, incontrovertible set of normatively ideal methods. The field of modern probability theory and statistics simply cannot be made to supply a unitary, self-consistent definition of normatively prescribed because the wealth of alternative and even contradictory methods and views coexist within its compass.

The inability to find a unitary definition of normative methods among logicians, statisticians, and probabilists does not represent an intellectual failure that will be eventually overcome by some long-awaited conceptual breakthrough. By the nature of things, there cannot be any methodological equivalent of the physicists' anticipated "theory of everything" that demonstrates that one methodological system is the ideal to be applied to all problems. The pluralism of formal inferential methods is, among other things, a direct consequence of the pluralism of computational goals, the pluralism of problem-types, and the pluralism of problem-solving conditions. The best tool depends, quite simply, on what you want to accomplish and the environment you are to operate in.

Grounding methods in goals and conditions: Performance-justified rationality

Many questions relating to rationality can be analyzed far more sensibly once one has distinguished the issue of which means are normative from the issue of which goals are normative. Once a goal is specified, any number of alternative methods can be compared by examining how well each performs in reaching the goal. The better a method is at reaching the goal, the more normatively preferable it is. On this view, the normativeness of a method -- or the rationality of a problem-solving system, such as a human -- would consist of its relative operational success compared to known alternative computational designs. This makes rationality or normativeness a matter of degree, and its definition relative to specific goals. For this reason, we will use terms such as *well-designed* or *well-engineered* as substitutes for *rational* or *normative*. Thus, we would express the question of rationality as, *Is the computational system being investigated well-designed or poorly designed for reaching the specified goal?* Relative performance in goal-attainment provides defensible and objectively testable criteria for identifying some methods as more normative than others -- criteria that are independent of appeals to authority, with all of their attendant drawbacks, inconsistencies, and susceptibilities to custom and prejudice. It opens up an entire range of new computational possibilities for consideration by reasoning researchers, such as domain-specific and ecologically rational methods, as well as newly invented procedures and approaches. Finally, it eliminates the intellectually indigestible possibility that the methods identified as normative will be worse at problem-solving than the methods identified as defective.

Several obvious points follow from this framework:

(1) The best method will depend on which goal is selected. Different methods will perform best according to different definitions of success. For example, a chess-playing program would require very different inferential rules if it was designed to lose as opposed to win. "Goals" in this sense include all of the different issues of costs and benefits relevant to alternative computational systems and decision consequences. For example, which kinds of errors are costly and which kinds are cheap (what, e.g., is the cost of being afraid of a

nonvenomous snake versus the cost of being unafraid of a venomous one)? What is the cost (in time, metabolic energy, processing load, and so on) of one system of computation as opposed to another?

(2) The best method will depend on the distribution of background conditions within which problem-solving is to take place (Manktelow & Over, 1990a, p. 153-154). Different methods will perform best in different problem-solving environments. Natural problem-solving tends to take place in complex environments with certain stable or statistically recurring features. To understand why a particular computational method will prove more effective in one environment than another, one needs to answer such questions as: What is always true in the task-environment, what is statistically true, and what is never true? What do detectable cues predict about the undetectable features of the environment? What information is routinely available? How stable are the variable dimensions of the task environment? And so on.

(3) The best method will depend on the ecological distribution of different problem-types that the problem-solving system encounters. Because computational strategies ordinarily involve trade-offs, different methods will perform best against different composite problem-populations. An appendage that is usually used to swim and rarely to walk (as on seals and turtles) will be different from an appendage that is usually used to walk and rarely to swim (as on dogs or humans).

Thus, the answer to the question *Which method is normative?* is not and cannot be invariant and universal. The normativeness of a method is always relative to the goal to be reached (or the total array of values and trade-off functions), to the background conditions that it which operates in, to the total problem-population to it will be applied, and to other factors as well. This is no more than common sense. For example, the answer to the question, *Which is a more rational method of transportation, a horse or a car?*, is indeterminate without supplying additional information. What is the starting point and the destination? Are there paved roads between the starting point and the destination and, if not, how rugged is the terrain to be traversed? What is the relative availability and expense of forage versus gasoline? What is your budget? How quickly does the trip have to be completed? The most rational method of transportation depends on factors like these, and cannot be identified in isolation from them. Similar considerations apply to the best method of statistical inference or the best method of deductive inference. Whether one is talking about a vehicle or a computational system, which system is judged best depends on the uses to which it will be put and the conditions under which it must perform.

In short, we propose that the seductive mirage of a formally derivable, ideal inferential system be replaced by an experimental, engineering sensibility. Indeed, even when goals and task environments are identified, formal derivation is not always possible. The relative problem-solving merits of alternative computational systems can depend on a great many complex and scarcely investigated real world factors. Consequently, formal analyses that attempt to encompass every potentially relevant factor are often impractical or impossible ways of evaluating (or developing) computational systems or models of cognitive functions. In such cases, formal analyses may still have great heuristic value for exploring well-defined parts of a question. Nevertheless, formal analyses must often be treated as potentially flawed artificial simplifications and only cautiously integrated with an experimental and engineering approach to understanding how the design-features of various computational systems interact with the complex structure of problem-environments. This way of thinking is, of course, completely familiar to applied mathematicians, artificial intelligence researchers, engineers, and evolutionary biologists, and we believe it will prove equally useful for reasoning researchers as well.

Accepting these points changes the interpretive framework one applies to the experimental investigation of human rationality. To begin with, one cannot use a specific result -- say, the finding that individuals fail to use a base rate on a particular problem -- to conclude that humans are irrational. Individual judgments may be successful or unsuccessful at attaining a specified goal, but not rational or irrational. This is because rationality, on this view, is a function of the *global* performance of a set of methods used to arrive at judgments, and not of any specific judgment taken in isolation. Individual examples of problem-solving failure cannot, by themselves, establish the existence of poor design or irrational methods because situations can always be found in which a given set of computational procedures will fail. (Indeed, traditional inferential methods are usually incapable of solving routine adaptive problems, such as echolocation or syntax induction.) The real question is not, *Does the system fail when confronting some situations?*, but, *Is there any known alternative computational system whose overall performance would be superior over the entire set of problems the computational system must contend with?* If not, then the system certainly qualifies as rational or well-engineered for that problem-population and set of task environments, and its problem-solving failures are simply cost-effective by-products of its design features. To call a system irrational is to claim that you have a specified model of a system that performs better on the total problem-population.

In fact, this engineering perspective inverts the implicit standard that humans have been held to within the heuristics and biases program. Within this program, discrepancies from ideal performance (defined by conformity to favored methods) have been considered evidence of irrationality and defective design, and it is the patterns of failure that tend to be emphasized, named, and categorized as the phenomena to be investigated (the conjunction fallacy, overconfidence, base rate neglect, the fundamental attribution error, and so on). Implicitly, the production of problem-solutions has been treated almost as if it were the state of nature, and only departures from this state of nature were phenomena that needed to be explained.

In contrast, from a computational (and evolutionary) point of view, the initial state of nature is a lack of organized responses to encountered problems -- that is, problem-solving failure. Indeed, before systems are functionally organized, they will not even recognize problems as problems, and their output will be unlikely to show any special tendency to solve meaningfully defined inferential problems. Rocks, oranges and waterfalls do not reason particularly well. Actual inferential systems were constructed and organized by evolution, by artificial intelligence researchers, or by logicians in order to solve targeted problem sets; commonly, over time, the incorporation of additional features progressively improves the performance of the system. If the initial state of nature is random performance with respect to problem solution, then increasing rationality is caused by the discovery and inclusion of increasingly better designed sets of procedures.

Moreoever, to know how well a system performs, one must know how large the space of possible answers is, how small the target of successful outcomes is, the structural difficulties intrinsic to isolating the solution set, and how much more often than chance the computational system actually succeeds in hitting the target. Depending on the complexity of the problem, it may take a great deal of intricate machinery to improve performance so that it is notably better than random. Indeed, difficulties inherent in the problem may set upper bounds on performance that cannot be exceeded (e.g., guessing coin flips), so that high rates of errors are inevitable even for optimally designed computational systems. Therefore, frequent operational failure to reach the goal is not necessarily caused by a failure to use a well-designed or even an optimally designed method.

For researchers operating within the heuristics and biases framework, on the other hand, perfectly error-free operation seems trivially easy to design into a system, because to qualify as error-free the system only needs to blindly operate according to "the" normative method. Under this interpretation, perfect rationality seems easy, unguided human reasoning seems unnecessarily inept, and the power of traditional rational methods to solve problems and render good judgments is a central article of faith. We suggest that the widespread acceptance of such beliefs has led many researchers to underestimate the difficulty of successful problem-solving on natural tasks, where success is defined by consequences and not by methods. Within this intellectual climate, the customary restriction of the term "rationality" to cases of error-free performance has blurred the vital distinction between well-designed and poorly-designed computational systems, both of which may make large numbers of errors. In a world full of difficult computational problems, an organic mind or a Voyager probe can be full of superbly well-engineered machinery yet still make many errors. If making errors qualifies a system as irrational, then both superbly performing and poorly performing systems -- very much worth distinguishing -- are lumped together into the same category as "irrational." We suggest that a more descriptive and less misleading word to use in this context is "infallible," allowing "rational" to be applied to any system that is well-designed to solve the relevant problem-population. To most cognitive and social scientists, the present practice of calling humans "irrational" carries with it the strong (and, as of yet, empirically unjustified) connotation that the human cognitive architecture is poorly-designed and generally performs far worse than easily imagined alternative designs. Based on the presently available evidence, however, all that can be said is that humans aren't infallible. It still remains to be determined whether humans are rational in the sense of embodying inferential mechanisms that are well-designed for the task environments and problem-populations humans usually confront or once confronted. No causal system can be omniscient or omnipotent. Unerring goal-attainment is simply not possible for any buildable computational system operating on nontrivial problem-sets. Therefore it ought not to be used as the definition of rationality.

Thus, an engineering perspective reverses the question. Instead of seeing any deviation from "the" normative theory as an error, bias, or fallacy -- i.e., as evidence of defective design -- one flips the interpretive framework. From an engineering perspective, deviations from random performance in the direction of successful problem-solving are evidence for the existence of functional organization -- what an evolutionary biologist would call an adaptation. One needs to explain not so much how a system makes errors as how it achieves its successes -- to identify what features of the system's design cause it to perform better than chance with respect to relevantly defined goals. Like an arrow hitting a target from a distance, correct performance is not the state of nature, from which deviations must be explained. Instead, systematic deviations towards the target are the phenomena to be explained.

Thus, each aspect of successful problem-solving must be explained (1) *proximately*, by mapping the computational design features responsible for success, and (2) *ultimately*, by describing the factors that introduced these function-enhancing features into the computational system. What is it about the design of the system that leads it to perform better than randomly (often phenomenally better) in solving problems? And why did the system come to have one design rather than another? For artificial intelligence systems, of course, the ultimate source of functional design is the programmer and her interpretation of the task-environment and goals of the computation. But for the human cognitive architecture, the ultimate cause of functional design is natural selection in structured ancestral environments. Both types of computational system have artificers that built in functional organization. Understanding a computational system consists of (1) mapping its architecture and (2) relating the design of its functional components to the nature of the problem-population and task environments they were designed to operate over. For such a functionalist program to be coherent, however, one needs a valid way of identifying what goals our reasoning mechanisms were designed to accomplish.

Adaptive problems justify normative theories of reasoning goals

We have been proposing that a distinction be made between normative methods and normative goals, arguing that a reasoning method should be considered normative when it is found to be the most effective method for achieving specified goals within specified task environments. But what, then, justifies privileging some outcomes as goals rather than others? What constitutes success? Where should normative theories of *goals* come from -- theories that would allow us to develop normative theories of reasoning methods? At the first level, proximate goals can be justified based on how well their achievement contributes to the realization of designated higher level values or goals. For example, the subgoal of inferring the local grammar allows one to acquire language and hence to communicate with others -- a higher level goal. But this leaves the question to be asked all over again of the higher level goals. How can goals themselves be normatively grounded?

Most philosophers and scientists who have considered the question have concluded that there is no way of ultimately justifying some goals or values over others -- that is, for deriving "ought from is" -- a view that we will not dissent from here. It is only after a goal is chosen -- "derive all truth-preserving deductions possible from a set of axioms"; "interpret magnetic flux measurements to reconstruct a three dimensional image of the brain" -- that it becomes legitimate or even possible to develop a set of normatively preferable methods appropriate to the goal. Even worse, humans have historically pursued an apparent multiplicity of proximate goals -- correct medical diagnosis, the elimination of poverty, service to the goddess Kali, profit-maximization, bringing Italy under the control of the House of Savoy, the elimination of the Albigensian heresy, and so on. Once humans acquire culturally specific or individually idiosyncratic goals, they often deliberately construct particular analytic methods to accomplish them -- mortgage amortization formulae, Bayesian probability revision, monsoon-timed planting rules. The investigation of such culturally acquired goal-specific competences could form a wide-ranging research program. Given any known goal, one could test subjects to see whether they happened to be, say, "mortgage amortization normative," determine whether they showed systematic errors and biases by such standards, and decide whether they were, for example, "mortgage amortization irrational." But is this a complete framework for the study of human rationality?

No one doubts that the individuating events of ontogeny lead different individuals (within and across cultures) to possess somewhat distinct sets of information and reasoning skills. But beneath the level of individual and cultural variability there exist sets of universal cognitive procedures and developmental programs that (1) regulate and drive the acquisition of individual-specific skills when exposed to specific ontogenetic environments, and (2) constitute the intuitive competences that all normal humans share in common. What many researchers, including ourselves, are after is a precise characterization of these species-typical cognitive and developmental mechanisms, with the individuating effects of ontogeny factored out. A variety of cognitive subcommunities are engaged in the parallel (and overlapping) enterprises of mapping the universal computational procedures underlying human vision, semantic induction, syntax induction, social cognition, object construal, emotion recognition, motor control, mate selection, and so on.

Although not always clearly articulated, the goal of mapping our universal computational architecture is shared by a number of reasoning researchers, including many of the cognitive psychologists who study judgment under uncertainty. For example, Tversky and Kahneman liken cognitive heuristics to perceptual heuristics and the study of cognitive illusions to the study of perceptual illusions (Tversky & Kahneman, 1974; Kahneman & Tversky, 1982): "we use illusions to understand principles of normal perception" (Kahneman & Tversky, 1982, p. 123). The message from their claim that cognitive biases "seem reliable, systematic, and difficult to eliminate" (Kahneman & Tversky, 1972, p. 431) is that these are reliably developing features of human cognitive processes, and not accidents of personal history, such as whether one happened to be schooled in statistics or logic. Indeed, if cognitive scientists thought that the documentation of normative fallacies established nothing more important about the human mind than how assorted subjects happened to be schooled, then the heuristics and biases program would be no more significant than a research program documenting that native speakers of Jivaro with no schooling in English commit "errors and fallacies" when they try to conjugate English verbs. Obviously, the common goal is to explore the underlying universal machinery. For this reason, when we use the terms *architecture*, *design*, *mechanism*, *module*, *intuitive competence*, or *mental organ*, we are referring to this level of species-typical developmental programs and to the reliably developing, effectively universal computational systems that they construct. We are not referring to any idiosyncratically acquired information, process, or skill.

Given a research enterprise that is directed toward the discovery and mapping of our human-universal cognitive architecture, we need not engage in the problematic attempt to analyze all the causes that regulate why individuals select some goals over others. Nor do we need to flounder in a sea of culturally particular goals, wondering which are germane. For the purpose of pursuing this research program, the question *What justifies normative theories of goals?* can be translated into the question, *What goals were our species-typical reasoning mechanisms designed to accomplish? What was the human mind designed to do?* This question, unlike some of the earlier ones, can be answered systematically and with some precision.

If the research program is the exploration of our species-typical cognitive architecture, then we can develop normative theories of goals that are relevant to understanding the computational designs of our reasoning mechanisms. We can define the tasks our reasoning mechanisms were organized to solve, and we can define what counted as solutions to these tasks. Simply put, these tasks were the cross-generationally recurring adaptive problems our hunter-gatherer ancestors encountered in their daily lives. The correct "solutions" to these problems were decisions that, on average, led to survival and enhanced reproduction within these ancestral contexts. Pleistocene hunter-gatherer conditions provide richly informative definitions of the problem-solving environments our mechanisms were constructed to operate in successfully.

Why are hunter-gatherer tasks in particular relevant to the design of our cognitive mechanisms? And why should inferences and decisions that promoted survival and reproduction -- as opposed to, for example, the derivation of universal truths -- be considered "correct" solutions? The answer is straightforward: These tasks, and this definition of problem-solving success, directly caused the assembly of whatever functionally organized cognitive mechanisms we humans now have. Ancestral conditions and problems are uniquely related to our species-typical cognitive architecture as cause and effect. So, the answer to the question, *What justifies normative theories of reasoning goals?* is: natural selection and the adaptive problems that recurred over human evolutionary history. (This answer is, of course, valid only for the purpose of discovering, understanding, and explaining the designs of the reliably developing reasoning competences of biological systems; it is not, for example, a general answer to how artificially intelligent systems should be constructed or how scientific inference ought to be conducted.)

In simple terms, a design feature has caused a biologically successful judgment -- that is, it has correctly solved an adaptive problem -- if and only if, on average, the judgments it generated caused the design feature to be reproduced into subsequent generations at a higher rate than the existing alternatives. For example, a device that correctly infers the presence of a hidden predator in the tall grass based on the behavior of prey animals in overhanging branches may lead the hominid it is situated in to survive long enough to reproduce the device into her children. Within this framework, the only thing that makes an inference "correct" is its adaptive consequences; not its truth, nor its validity, nor its warranted derivation, nor its consistency (for discussions of natural selection and rationality from contrasting viewpoints, see Dennett 1978, Stich 1985, and Manktelow & Over, 1990a). Design features that are successful at problem-solving by this definition (and only by this definition) will cause their own spread, until they ultimately become a species-typical component of that species' cognitive architecture. The presence of such design features in our cognitive architecture is thus explained by their systematic consequences on their own propagation under ancestral conditions. Hence, natural selection imposed non-subjective and non-arbitrary definitions of success and failure on the reasoning judgments of our ancestors. These definitions of problem-solving success and failure are the criteria that regulated which features were incorporated over evolutionary time into our architecture and which features were eliminated. They are the foundations for all of the valid normative theories that our reasoning mechanisms can be expected to embody in their designs.

This is not to say that mathematicians, moralists, economists or cooks don't each have alternative, and quite valid, definitions of successful problem-solving. Nevertheless, evolutionary definitions of successful problem-solving (i.e., of which goals are normative) are the only applicable definitions for psychologists interested in discovering the species-typical functional organization of the human mind. Any other definitions (truth-preserving, utility-maximizing, loss-minimizing, hypothesis-testing, and so on) acquire their interest for research into our cognitive architecture solely to the extent that they happen to parallel the evolutionary definitions.

In short, operational success on adaptive problems (rather than conformity to some method) was the criterion by which computational designs were selected or eliminated during the evolutionary process. Ancestral environments and lifeways were the engineering proving ground within which alternatively designed cognitive devices were tested against one another. For our evolved reasoning competences (as opposed to our culturally acquired skills), *normative theories of the goals of reasoning are grounded in and solely derivable from the dynamics of natural selection, the realities of ancestral hunter-gatherer life, and the structure of ancestral environments.* Only by integrating these factors can cognitive scientists define the function of a cognitive mechanism. If cognitive psychologists want to apply a functionalist approach, it needs to be an evolutionary approach as well.

With the discovery of an objective and analyzable source of definitions of successful reasoning, the components of a program for investigating species-typical human reasoning mechanisms can finally be fit together. First, models of the task environments and problem-populations our ancestors encountered can be consulted or developed. Second, normative theories of goals, defining what counts as successful reasoning can (and must) be built out of models of the selection pressures and adaptive problems humans encountered during their evolution. Third, once the goals of reasoning are defined within a domain, then rational methods -- methods capable of reaching those goals -- can be formulated. Fourth, alternative methods can be normatively evaluated based on their relative performance. Fifth, experiments can be devised that test for the presence of normative performance, so defined. Sixth, experiments can be devised to determine which set of algorithms and representations accounts for the reasoning performance found.

This is, in essence, the ecological and functionalist approach that David Marr used in developing computational models of vision (Marr, 1982), and the approach that we have argued elsewhere is equally applicable to the study of reasoning (Cosmides, 1989; Cosmides & Tooby, 1987, 1989, 1992). Briefly described, Marr's approach involves: (1) defining the goal of the computation -- that is, what counts as successful processing; (2) discovering the stable constraints and properties of the task and the task environment; (3) developing candidate procedures that could, in interaction with the recurrent structure of the task environment, accomplish the goal; and (4) experimenting to see which proposed procedures, if any, the mind appears to embody. This functionalist approach has been productive in every area of psychology where it has been applied. As Marr put it, "an algorithm is likely to be understood more readily by understanding the nature of the problem being solved than by examining the mechanism (and the hardware) in which it is embodied" (Marr, 1982, p. 27).

necessary? repeats the chapter 1.1+

At least two of the key elements in this approach -- discovering the structure of the task and task environment and defining success differently depending on the problem-type and situation -- tend to be ignored in most reasoning research, because of the general-purpose, domain-general premises such research proceeds from. If reasoning is a general-purpose competence applicable to any possible task from any possible task environment, then what is present in one task environment will be absent in others. Consequently, tasks and task environments will have no significant recurrent structure that general-purpose reasoning mechanisms could exploit. If the mechanisms are truly general-purpose, then studying the ecological structure of environments contributes nothing to reasoning research. Similarly, the practical goals of the computation will also differ from task to task; hence, they cannot be incorporated into a general theory of rationality. A Marrian approach makes little sense if reasoning is conceptualized as a unitary faculty applied uniformly to every problem.

The functional organization studied by cognitive scientists was created by natural selection.

Cognitive issues always inhabit an encompassing evolutionary landscape, and certain features of this landscape merit attention by cognitive psychologists.

(1) The starting point is the recognition that our cognitive architecture and its constituent mechanisms is the assembled product of the evolutionary process acting over hominid and vertebrate evolutionary history. Inescapably, evolutionary principles and evolutionary history provide the explanation for why the species-typical human cognitive architecture (as well as every other species-typical biological system) has the specific organization that it does. A great deal is now known about how evolution systematically acts, and much can now be confidently

concluded about the adaptive problems our ancestors encountered and what constituted biologically successful solutions to them (see, e.g., Barkow, Cosmides & Tooby, 1992; Daly & Wilson, 1984; Krebs & Davies, 1987).

(2) Evolution operates according to only two classes of processes, chance and selection, and they collectively explain all aspects of species-typical architectures. Moreover, chance and selection each cause different aspects of species-typical architectures, with virtually all complexly organized functionality injected by natural selection. Of course, there are many nonselectionist processes in evolution by which species are modified away from ancestral designs (e.g., drift, macromutation, hitchhiking, developmental by-product). But selection, by retaining variants that are more functional, is the only process that has a systematic tendency to propel the evolving system in the direction of otherwise improbable but increasingly functionally organized arrangements, instead of into the immeasurably vaster space of uncoordinated, disorganized, and nonfunctional arrangements that the system could move to at each of the innumerable choice points in the evolution of designs (Dawkins, 1986; Pinker & Bloom, 1990; Tooby & Cosmides, 1990a).

In contrast, a feature's degree of functionality plays no role in determining whether nonselectionist processes will cause it to be retained or eliminated. As a result, chance processes push evolving designs through design-space in a random walk (Kimura, 1983), and these stochastic processes have produced nothing in our cognitive architecture more organized or functional than can be accounted for by chance (Tooby & Cosmides 1990a, 1990b). Therefore, the complexly functional aspects of architectures were created virtually exclusively by natural selection, and hence reflect the design principles of natural selection (for good summaries of these design principles, see Daly & Wilson 1984; Dawkins, 1976, 1982, 1986; Krebs & Davies, 1987; Williams, 1966).

It is important to stress that natural selection is the only known explanation for the existence of the functional complexity (adaptations) present in the designs of organisms (Williams, 1966; Dawkins, 1986; Pinker & Bloom, 1990; Tooby & Cosmides, 1990a). Therefore, everything that is complexly functional in our cognitive architecture is an adaptation built over evolutionary time by natural selection. For this reason, cognitive psychologists, like physiologists, are usually studying adaptations and their effects, and they can use adaptationist analytic tools to inform their research (e.g., Cosmides, 1989; Cosmides & Tooby, 1989, 1992; Freyd, 1987; Gallistel, 1990; Gigerenzer & Hug, 1992; Jackendoff, 1992; Leslie, 1987, 1988; Marr, 1982; Pinker & Bloom, 1990; Ramachadran, 1990; Rozin, 1976; Sherry & Schacter, 1987; Shepard, 1981, 1984, 1987a, 1987b; Shiffrar & Freyd, 1990; Staddon, 1988; Tooby & Cosmides, 1992).

(3) Although adaptations may not be perfectly optimal, they tend to be very well-engineered to solve adaptive problems. By dwelling at length on the uncontested point that evolutionarily processes do not necessarily find the one unique optimum in design-space, some scholars have left the impression among nonbiologists that biological structures are therefore at best only crudely functional (see, e.g., Gould & Lewontin, 1979; for discussion, see Dawkins, 1982, 1986). Whenever biologists have known enough about the biological problem faced by an adaptation to investigate it in depth, however, they have typically found adaptations to be extremely well-engineered and impressively functional. To take a relevant cognitive example, no one knows what the optimal design for a visual system would look like because (to take only one reason) there are just too many degrees of freedom in possible designs to construct a description of the space of alternatives within which one could identify optimality. Nevertheless,

What don't sometimes feel is — [illegible] intellectual point (But everything will)

not only is human vision extremely good, but no one has been able to build an artificial system that remotely approaches it. The same is true for speech perception, language acquisition, and so on. It seems reasonable to insist that before the claim that a set of adaptations are poorly designed is given weight, we ought to be able to specify alternative designs that perform better. When this criterion is used, adaptive problem-solving structures like the visual system or the language acquisition device begin to look very well-engineered indeed, and species appear to be collections of well-engineered problem-solving mechanisms (and their by-products). From this perspective, the crude and error-prone appearance of human reasoning seems anomalous as against the known co-existence in the same system of very well-engineered information-processing adaptations such as vision. Before we can conclude that human reasoning is poorly designed, we need to answer several questions, including *What adaptive problems did human reasoning evolve to solve?* and *What would constitute successful reasoning performance from a biological or adaptive standpoint?*

(4) Natural selection coordinates the structure of a recurrent adaptive problem (including the features of the environment in which it occurs) with the structure of an adaptive problem-solver such that the interaction of the two produces the solution to the problem. If selection has created a well-engineered adaptation, then elements that are necessary to solve the problem but lacking from the world are supplied by the structure of the problem-solving device. Equally, that which is reliably supplied by the environment will tend to be left out of the device. So, strictly speaking, one should not look for the complete solution to the adaptive problem in the mechanism itself -- the solution emerges from the complementary interaction of the mechanism and the world. Endogenous circadian rhythms are entrained by the 24 hour periodicty of the sun, for example (Shepard, 1987a); and although we have complex mechanisms for semantic induction, the sound patterns of particular words are stored in a child's social environment, not in its genome (Pinker & Bloom, 1990). To understand the operation and organization of adaptations, it is necessary to understand what regularities reliably permeated the structure of natural problem environments -- the environment of evolutionary adaptedness, or EEA (Tooby & Cosmides, 1990a, 1992). Obviously then, the malfunctioning of devices frequently comes about when a situation lacks cues and relationships that tended to be stably true in the past, and which the device relies on for its successful operation. Locating one's car in a parking lot illuminated by sodium vapor lamps is difficult; our color constancy mechanisms fail under these circumstances because they were designed to mesh with the spectral properties of natural sunlight (Shepard, 1992). This is why one must talk about *ecological* rationality; no causal mechanism can operate properly outside of the context for which it was designed.

(5) Only evolutionarily long-enduring problems last long enough for selection to build complex cognitive problem-solvers. For this reason, complex cognitive machinery will be designed to address problems and conditions that recurred during the several million year period our ancestors spent living as hunter-gatherers. We should not expect to find complex adaptations for solving rare, transient, or novel modern problems that have emerged only in the last few thousand years in agricultural and industrial societies (for argument, see Tooby & Cosmides, 1990b).

In sum, the following points about evolution are relevant to the study of reasoning. The universal and reliably developing organization of our cognitive architecture is the product of evolution. To the extent that there is complex functional organization in this architecture, it was built in by the process of natural selection, and hence should reflect selectionist criteria of functionality. Whatever functionally organized mechanisms exist in the mind are adaptations that evolved to the solve long-enduring problems our ancestors faced as hunter-gatherers. Because these long-enduring adaptive problems are the cause of the functional organization in our cognitive architecture, the design of our cognitive problem-solvers should reflect the persistent structure of these adaptive problems and the environments which accompanied them.

Two levels of ecological structure:
Ontogenetic stability and phylogenetic stability

Where relationships have been cross-generationally stable, domain-specific mechanisms are expected to have evolved to exploit them. But a large part of the hunter-gatherer's world was comprised of rapidly-decaying associations and relationships: the changing spatial and temporal distributions of game, plant foods, predators, social groups, daily weather, and so on. Such short-lived ecological structure disappears too rapidly for evolution to detect it and build corresponding domain-specific adaptations to be inherited by subsequent generations. Nonetheless, many such relationships will be stable enough within a lifetime to be well worth detecting and modeling, and the local sampling of the frequencies of such associations and relationships can provide a reliable basis for prediction and decision-making.

Thus, there is both long-enduring and short-lived ecological structure, and our cognitive architecture is consequently expected to have evolved two broad classes of inferential mechanisms to cope with these two varieties of ecological structure.[4] As noted, it is the stability of encountered relationships over evolutionary time that allows domain-specific inferential devices to evolve and to outperform more general mechanisms. The functional superiority of these devices derives in part from the fact that the tasks and task environments they face often continue to reflect, in certain respects, the structure of past tasks and task environments; consequently, the system can know things about the task that are not logically derivable or sensorily apparent. This problem-solving strategy is foreclosed for tasks requiring the analysis of relationships that decay within a single or a small number of generations. For this reason, cognitive adaptations that are designed to deal with short-lived ecological structure must necessarily operate according to ~~along~~ more domain-general, content-independent principles. Precisely because they are relatively more content-independent, we expect that mathematically derived normative theories might be more appropriate for understanding the functional design of those cognitive components designed to detect and model the short-lived ecological structure humans encounter. In short, one expects architectures that embody more content-free normative theories to process just those kinds of information that our domain-specific inductive architectures do not (Cosmides & Tooby, under review).

For this reason, we expect the human mind to also be equipped with mechanisms that detect, over the course of ontogeny, the short-lived ecological structure of the local world they inhabit, continually updating its changing features. Brunswik (1950, 1955) argued vigorously for such a view, and Gigerenzer and his colleagues have more recently advanced specific proposals of how humans construct probabilistic mental models of their environment (Gigerenzer, Hoffrage & Kleinbolting, 1991). Indeed, humans do appear to be equipped with the automatic frequency encoding mechanisms such a view would predict (Hasher & Zacks, 1979).

Consequently, reasoning researchers have two kinds of ecological validity to contend with in constructing and interpreting their experiments. Not only will subjects automatically embed reasoning problems in a web of domain-specific inferences derived from evolutionarily long-enduring ecological structure, but they will (quite reasonably and, no doubt, unconsciously) embed reasoning tasks in assumptions derived from their ontogenetically extracted models of local, short-lived ecological structure. If the experimental protocol contains features that clash with either the ontogenetic ecological structure to which subjects are calibrated *or* the phylogenetic structure their cognitive architecture is designed to expect, then what would have been good reasoning in the subjects' normal environment will be judged as poor reasoning by the experimenter (e.g., Cosmides & Tooby, 1992; Gigerenzer, Hell & Blank, 1988; Gigerenzer, Hoffrage & Kleinbolting, 1991).

Therefore, experimenters need to remain keenly aware that humans evolved to make consequential decisions in real environments and were under no selection to perform hypothetical reasoning operations abstracted from the actual environment that they acted in. Our cognitive architecture should be saturated with information about the ecological structure in which it operates, including both its short-lived and its evolutionarily enduring relationships. Because it rarely (or never) would have been called on to make consequential decisions outside of its normal environment, an architecture that is well-engineered by evolutionary standards would be designed to automatically apply its ecological knowledge to every reasoning task it faces. (Indeed, a feature of this kind could account for people's notorious inability to transfer a solution from one experimental protocol to a formal isomorph having radically different content.) There would have been no selection for an amnesic mind -- a mind that suspends its knowledge of the world when reasoning.

For these reasons, automatic and nonconscious cognitive mechanisms are expected to "contaminate" reasoning on artificial, ecologically invalid problems with additional tacit assumptions drawn from the ecological structure of the subject's environment. Because the best (most rational) decision will differ depending on the exact array of relevant information available to the decision-making system, experimenters must take into account subjects' "knowledge" (including their nonconscious procedural knowledge) before they can render a meaningful judgment about whether human reasoning performance is good or poor. The information presented to the subject as the "problem" is not a complete inventory of this knowledge. (ref to imported assumptions

The investigation of ecological rationality

In this chapter, we have been defending the view that the role of ecological structure is an important but missing element in the analysis of human reasoning and that our evolved cognitive architecture is designed to rely on, detect, and exploit this structure in solving the inferential problems it confronts. Moreover, this explanation has the ability to resolve the paradox of how humans can solve so many problems that bewilder formal systems, and yet appear to perform so poorly on laboratory problems when the standard of good performance is conformity with traditional general-purpose normative theories. We are not arguing that the traditional normative theories of rationality are wrong -- we are simply pointing out that they are only a few out of many alternatives, and happen to perform more poorly than others when operating within certain ecological contexts. Traditional rational methods simply lack the additional design features that would make them specially efficient in dealing with the ecologically structured problems humans have tended to confront.

Accordingly, reasoning researchers need to experimentally decide between two opposing meta-hypotheses about human rationality. Does the human mind operate according to a small number of rather primitive heuristics? Or does our cognitive architecture contain a large community of well-designed inferential specializations (both domain-specific and domain-general), each of which exploits the adaptive problem-solving opportunities afforded by the existence of ecological structure (both evolutionarily enduring and short-lived)?

We have already mentioned a broad range of studies drawn from a number of different cognitive subcommunities that adds support to the hypothesis that the human mind is ecologically rational. The fields of perception, psychophysics, syntax induction, semantic induction, the categorization of the biological world, cognitive development (including physical causality, biological causality, the analysis of artifacts), and others all show evidence of a rich pluralism of human-universal, domain-specific inferential procedures. Within the field of reasoning research specifically, we would like to touch on two examples worked on by Gigerenzer and his colleagues and by ourselves -- one exploring a reasoning device designed to operate on short-lived ecological structure (inductive reasoning), and one designed to deal with a long-enduring domain (reasoning about social exchange). The basic question we want to address is, Do research programs that fail to consider how cognitive mechanisms were designed to function in structured ancestral environments overlook the presence of well-designed reasoning mechanisms?

at beginning?

Let us take as our first example the current assessment of how humans perform on problems of inductive reasoning. The most widespread view of human inductive reasoning is that, aside from whatever may have been inculcated through explicit training, humans do not have cognitive procedures that are designed for and proficient at statistical inference. As Kahneman and Tversky put it, "[i]n making predictions and judgments under uncertainty, people do not appear to follow the calculus of chance or the statistical theory of prediction. Instead, they rely on a limited number of heuristics which sometimes yield reasonable judgments and sometimes lead to severe and systematic errors" (1973, p. 237). Equally widely accepted is the conclusion that "In his evaluation of evidence, man is apparently not a conservative Bayesian: he is not a Bayesian at all" (Kahneman and Tversky, 1972, p. 450). In the years since these statements were made, a very large experimental literature has accumulated that supports and documents these conclusions.

Yet during their evolutionary history, humans and other species were continually called upon to make decisions under uncertainty. From an evolutionary-functionalist viewpoint, it seems implausible that would lack well-engineered mechanisms designed to cope with the probabilistic nature of the world under ecologically valid conditions. Indeed, behavioral ecologists who study foraging have found evidence of very sophisticated statistical reasoning in organisms with nervous systems that are considerably simpler than our own, such as certain birds and insects (e.g., Real, 1991; Real & Caraco, 1986). Moreover, John Staddon (1988) has argued that in animals from sea snails to humans, the learning mechanisms responsible for habituation, sensitization, classical conditioning and operant conditioning can be formally described as Bayesian inference machines. Humans in the Pleistocene would have needed mechanisms to detect and model short-lived, ontogenetically changing ecological structure, and yet the evidence appears to indicate that they lack them. Could such mechanisms exist, and yet have been overlooked because of the lack of an evolutionary, ecological perspective in reasoning research?

Yes. The overwhelming majority of studies that have been used to support the view that humans lack a calculus of probability are based on experimental designs that ask subjects to calculate the probability of a single event (Gigerenzer, 1991). In overconfidence research, for example, a subject is typically asked to judge the probability that she answered a particular question correctly -- i.e., the probability of a single event -- instead of being asked to judge how many of the last 100 questions she answered correctly -- i.e., a relative frequency. As Gigerenzer emphasizes, the distinction between single event probabilities and frequencies is fundamental to frequentist approaches to probability, and human intuitive statistical competences may be organized to reflect frequentist approaches (Gigerenzer, 1991). Although in the modern world we are continually exposed to numerically expressed statistical information, our hominid ancestors were not (Cosmides & Tooby, under review). Reliable numerical statements about the probability of a single event were rare or nonexistent in the Pleistocene -- a conclusion reinforced by the relative poverty of number terms in modern band level societies. And, significantly, the probability of a single event is intrinsically unobservable. No sense organ can discern that if we go up the west slope, there is a 0.2 probability that today's hunt will be successful. One can only observe, after the fact, that it was or it wasn't. No organism can evolve cognitive mechanisms designed to receive as input, and then reason about, information of a type or format that did not regularly exist. For human hunter-gatherers, the primary or sole available source for ontogenetically shifting statistical information would have been direct individual encounters with the local frequencies of events -- for example, that we were successful 4 out of the last 20 times we hunted on the west slope. Our hominid ancestors were immersed in a rich flow of observable frequencies that could be used to improve decision-making under uncertainty, given procedures that could take advantage of them. So if we have well-engineered mechanisms for inductive reasoning, they should be designed to accept and process information in the only format it was regularly available to them -- as event frequencies, rather than as single event probabilities (Cosmides & Tooby, under review).

If this were the case, experiments about single event probabilities would fail to activate such mechanisms, and researchers would falsely conclude that humans lacked well-designed mechanisms for statistical induction (Gigerenzer, 1991; Cosmides & Tooby, under review). Using a problem famous in the literature for eliciting poor statistical reasoning performance even from technically educated subjects (Casscells, Schoenberger, & Grayboys' (1978) Medical Diagnosis Problem), we found that correct Bayesian reasoning could be elicited from 76% of subjects simply by expressing the problem in terms of frequencies, rather than single event probabilities, and from 92% of subjects through the use of visual representations of the frequencies involved. In short, by improving only this one dimension of ecological validity, extremely high rates of good Bayesian performance were obtained. These results contradict the consensus views that the human cognitive architecture is not capable of well-calibrated Bayesian statistical inference (Cosmides & Tooby, under review; see also McCauley & Stitt, 1978).

This study is not an isolated case. It is one of a growing body of reports where performance on tasks that ask for the probability of a single event has been compared to performance on similar tasks that ask for the answer as a frequency. "Errors" that are reliably elicited by the single-event task disappear on the more ecologically valid frequency task. For example, Fiedler (1988) showed that the "conjunction fallacy" virtually disappears when subjects are asked for frequencies rather than single-event probabilities. The same manipulation can also cause the "overconfidence bias" to disappear. "Overconfidence" is usually defined as a discrepancy between one's degree of belief (confidence) in a single event and the relative frequency with which events of that class occur. But when one instead asks subjects to estimate frequencies, and then compares their estimates with actual frequencies, as Gigerenzer, Hoffrage & Kleinbolting did, "overconfidence" disappears; subjects' judgments turn out to be quite accurate. According to their *probabilistic mental model* theory, people accurately encode and store frequency information from their environment. Their data supports this assumption -- indeed, it is difficult to see how subjects' performance could be so well-calibrated unless they do (Gigerenzer, Hoffrage & Kleinbolting, 1991). These results also fit with the literature on automatic frequency encoding. One would expect an organism that relies on frequency information in making judgments under uncertainty to be constantly encoding accurate frequency information from the environment in a way that does not interfere with the organism's ongoing activities, and we appear to have mechanisms designed to do just that (e.g., Hasher & Zacks, 1979).

This growing body of research illustrates how ignoring the hypothesis that the mind is ecologically rational can lead one astray. The failure to consider what forms certain types of problems would have taken in the Pleistocene led researchers to overlook the existence of a family of cognitive mechanisms that do just what most researching researchers think the human mind cannot do: reliable statistical inference.

The second example we want to touch on is the case of domain-specific reasoning about social exchange. Nearly a decade ago we specifically selected the narrow but evolutionarily ancient domain of social exchange as a test case, to see how productive an ecological, evolutionary and functional approach to reasoning research could be. Through a series of experimental investigations using the Wason selection task, we and others have found that human reasoning about social exchange does display the special properties that we predicted it would have -- properties that reflect the evolutionarily long-enduring ecological structure of social exchange relationships, and the peculiar and specialized computational goals required by this particular adaptive task (e.g., Cosmides, 1985, 1989; Cosmides & Tooby, 1989, 1992; Gigerenzer & Hug, 1992). Because we have already written at length elsewhere about both the relevant theory and experiments we won't review this literature here (for alternative interpretations see, e.g., Cheng & Holyoak 1985, 1990; Manktelow & Over, 1990a, 1990b, 1991; for responses to these interpretations, see Cosmides & Tooby, 1992).

For the purposes of this chapter, we want to focus on the following three points:

(1) When subjects are asked to look for violations of conditional rules on the Wason selection task -- a test of logical reasoning -- they routinely fail;

(2) In all cases where subjects have been asked to perform the logically isomorphic task of detecting cheaters on social contracts, their performance was reliably and substantially enhanced; and

(3) In all cases where social contract problems were presented in which the answer predicted by social exchange logic was diametrically opposed to the answer required by the propositional calculus, subjects followed the social exchange logic, not the propositional calculus.

For the central argument of this chapter, this last point is the most important. In these experiments, subjects were successfully solving a complex adaptive problem. Yet if they were being evaluated according to traditional normative criteria -- i.e., by whether they reasoned according to the canons of formal logic -- their reasoning would have been scored as fallacious, illogical, and irrational. In reality, however, they were reasoning correctly according to the ecologically rational normative standard appropriate to this adaptive problem-type. By being equipped with reasoning procedures that embody a social exchange logic, they were able to reason more effectively than they could have if they had only been equipped with the rules of inference of the propositional calculus. This is a clear and specific case of how, by being ecologically rational, the human mind can be a more powerful problem solver on adaptive tasks that it would if it were merely rational by traditional standards.

A substantial body of evidence now supports the view that (1) social exchange constitutes a bounded domain, (2) defined by its content and by evolutionarily important adaptive demands, which (3) has led to the evolution within our cognitive architecture of its own distinct, adaptively specialized, and "nonlogical" reasoning procedures that (4) are designed to exploit the evolutionarily long-standing structure of the problem with particular efficiency (for the most recent examinations of the evidence, see Cosmides & Tooby, 1992; Gigerenzer & Hug, 1992). It represents another case of a reasoning mechanism whose existence would have been overlooked if the evolutionary and ecological structure of the reasoning problem had not been explicitly considered. Similar experiments using the Wason selection task have provided evidence for the existence of two other domain-specific inferential devices, one for analyzing threats (Tooby & Cosmides, in prep.) and one for analyzing precautionary rules (Manktelow & Over, 1990b, 1991). As more adaptive problems are considered, we expect that a growing number of other domain-specific inferential devices will come to light. Many cognitive scientists are uncomfortable with the prospect of a rapidly proliferating inventory of functionally specialized reasoning devices. But consider the number of mechanisms that are functionally specialized for vision -- indeed, the number that are specialized for the one subtask of depth perception. Against this background, it seems reasonable to assume that the number of devices devoted to "everything else in the world put together" might be at least as numerous.

Conclusion

A skeptic may call into question whether, after the rhetoric is stripped away, there remains any substantial dispute between the heuristics and biases approach and the ecological rationality approach. Both sides agree that humans are rarely if ever using traditional normative reasoning methods. Both sides agree that humans are reasoning according to other principles. Both sides agree that these alternative principles or procedures "sometimes yield reasonable judgments and sometimes lead to severe and systematic errors" (Kahneman & Tversky, 1973, p. 237). Where are the differences?

The first difference lies in which set of normative theories human reasoning is measured against -- traditional content-independent theories borrowed from mathematics and logic or normative theories grounded in the adaptive problems the human mind evolved to solve. The second difference lies in whether human reasoning procedures are expected to commonly be domain-specific in their design and their range of application. The third difference lies in the attribitution of major causes of errors. From an evolutionary perspective, many errors will be caused by well-engineered mechanisms confronting either evolutionarily novel tasks or tasks lacking ecological validity. From a heuristics and biases perspective, errors are the easy-to-generate consequence of using a small set of simple heuristics that are prone to biases, fallacies, and limitations.
systematically produce

But the paramount difference is the clashing expectations about the number, nature and, especially, the quality of the mechanisms that govern human reasoning. If the human mind is organized to be ecologically rational, then the human cognitive architecture will look considerably different than most cognitive scientists anticipate. According to this view, the human mind is likely to be full of complex and well-engineered inferential specializations, whose organization relates in sophisticated ways to the nature of the problems to be solved. In contrast, partisans of the heuristics and biases program have argued that heuristics and other non-normative mechanisms that govern human reasoning must be simple in order to fit within the purportedly severe information-processing limitations imposed by our cognitive architecture. Heuristics "reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations" and, hence, "make them tractable for the kind of mind that people happen to have" (Tversky & Kahneman 1974, p. 1131; p. 1124; Kahneman, Slovic & Tversky 1982, p. xii).

These conflicting expectations about the quality of our cognitive machinery motivate very different research programs. One research program involves testing human reasoning against standard normative theories drawn from statistics and logic, and considers the analysis virtually complete when "errors" have been discovered and inventoried. The other program, with its greater expectations for the quality of our cognitive mechanisms, is suspicious of such easy conclusions about erroneous performance. This program involves using evolutionary biology, hunter-gatherer studies, and a more pluralistic vision of the range of possible problem-solving methods to construct new normative theories and new hypotheses about the probable design of our computational devices. These can be used to construct more ecologically valid experiments, ones that are capable of activating well-engineered mechanisms that may otherwise remain dormant in the laboratory. Ecologically normative theories can also function as organs of perception, allowing us to recognize in patterns that were formerly thought of as "errors", the operation of well-designed mechanisms whose existence had been obscured by applying the wrong normative theories.

If the human mind has been judged irrational, the fault may lie not in ourselves, but in our standards.

Footnotes

1. That is, if you're so smart, why ain't you rich?

2. That is, the exclusive use of many of these general methods may be appropriate only within a narrow and odd envelope of conditions: The computational system must know enough about the problem to be able to identify the general method appropriate to its solution (e.g., the propositional calculus, Bayes' Rule, Neyman-Pearson decision theory); it must know enough to ascertain whether the assumptions these methods depend on have been met (e.g., independence, random sampling); and, yet, at the same time it must not know anything that is relevant to the problem that would allow a more effective domain-specific problem-solving procedure to be applied. Moreover, because they are weak, general procedures should only be applied to the residue of cases that cannot be handled by the total array of more powerful, more structured mechanisms (Tooby & Cosmides, 1992). It is an interesting and open question whether the residual cases that meet all of these conditions would have been significant enough to select for the evolution of domain-general mechanisms of this type, together with the complex guidance system they would require to regulate when they should be applied.

3. It is an interesting question why domain-specific methods have been so neglected, given how powerful they often are. There are some obvious disadvantages, of course. In the first place, while a domain-specific method may be a powerful problem-solver, by its nature it will only solve a limited set of problems. In contrast, the allure of a procedure that could solve all problems is compelling. Second, the use of a domain-specific method involves complications. One needs to define the domain across which the method is effective, discover cues and decision rules that predict whether a problem is in the domain, and then examine each problem to see if it qualifies. Still, this need not be more difficult than the assumption-checking required by so many content-independent methods (the random sampling assumption, assumptions about background distributions, etc.). Perhaps most importantly, to construct domain-specific methods, one has to discover a great deal about the domain in question, a formidable and laborious enterprise. Universally applicable methods often seem to hold out the seductive promise of insight without effort -- through logic, one can be an instant expert on everything, without having to go to the work of learning about each field. Finally, because content-domains are often complex, the domain-specific methods that solve them often turn out to be intricate, elaborate, and cluttered systems rather than elegantly simple procedures. The development of such methods may often appear more like engineering, and less like mathematics. In consequence, developing domain-specific methods may often seem like too much work for a pay-off that is bounded in its applicability.

4. More precisely, we expect that the two dimensions of ecological structure have each shaped our cognitive architecture and, accordingly, that researchers will be able to discover specific design features that were designed by natural selection to exploit each component of ecological structure. We are emphatically not saying that cognitive mechanisms can be divided into two mutually exclusive sets, one exploiting long-term structure, the other exploiting short-term structure. Indeed, we suspect that "hybrid" designs that integrate both types of information

are extremely common in our cognitive architecture. For example, the cognitive mechanisms underlying snake phobias were designed by natural selection to reflect the evolutionarily long-enduring ecological relationships between snake-appearances and the threat of death or injury by venomous bite. Nonetheless, such mechanisms also process and integrate ontogenetically derived information with phylogentically supplied information to calibrate the intensity of the fear (see, e.g., Mineka & Cook, 1988). Similarly, word acquisition in children is governed by a series of phylogenetically provided assumptions regulating the child's set of hypotheses (e.g., the whole object constraint, the taxonomic constraint, the basic level object assumption, the mutual exclusivity principle); yet this system drives the acquisition of culturally specific associations between a word and a category -- ontogenetically supplied information (Markman, 1989). As Staddon has pointed out, many kinds of learning can be viewed as the integration of a phylogenetically supplied prior probability with ontogentically supplied observations that guide the revision of the hypothesis (Staddon, 1988).

Acknowledgments

We warmly thank Lorraine Daston, Ken Manktelow, Robert Nozick, David Over, Steve Pinker, Roger Shepard, and Ed Stein for many stimulating discussions of the issues discussed in this chapter. We are especially indebted to Gerd Gigerenzer, who has had an evolutionarily long-enduring impact on our thinking about rationality. The financial support for this chapter was generously provided by NSF grant #BNS9157-449 to John Tooby and the McDonnell Foundation. Our very special thanks go to Ken Manktelow and David Over for their patience and understanding, which extended far beyond what any normative theory of rationality could predict.

Bibliography

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social psychology, 35*, 303-314.

Astington, J. W., Harris, P. L., & Olson, D. R., eds. (1988). *Developing theories of mind.* Cambridge, UK: Cambridge University Press.

Avis, J. & Harris, P. L. (1991). Belief-desire reasoning among Baka children: Evidence for a universal conception of mind. *Child Development, 62*, 460-467.

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

Baron-Cohen, S., Leslie, A. & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*, 37-46.

Brown, A. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science, 14*, 107-133.

Brunswik, E. (1950). *The conceptual framework of psychology*. Chicago: University of Chicago Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193-217.

Carey, S. & Gelman, R. (Eds.). (1991). *The epigenesis of mind.* Hillsdale, NJ: Erlbaum.

Casscells, W., Schoenberger, A., & Graboys, T. B. (1978). Interpretation by physicians of clinical laboratory results. *The New England Journal of Medicine, 299*, 999-1001.

Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton & Co.

Chomsky, N. (1959). Review of Skinner's "Verbal Behavior". *Language, 35*, 26-58.

Chomsky, N. (1975). *Reflections on language*. New York: Random House.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Cognitive Science, 14. Special issue on Structural constraints on cognitive development.

Cosmides, L. (1985). Deduction or Darwinian algorithms?: An explanation of the "elusive" content effect on the Wason selection task. Doctoral dissertation, Harvard University. University Microfilms #86-02206.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*, 187-276.

Cosmides, L. & Tooby, J. (1987). From evolution to behavior: Evolutionary psychology as the missing link. In J. Dupre (Ed.), *The latest on the best: Essays on evolution and optimality*. Cambridge, MA: MIT Press.

Cosmides, L. & Tooby, J. (1989). Evolutionary psychology and the generation of culture, Part II. A computational theory of social exchange. *Ethology and Sociobiology, 10*, 51-97.

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

Cosmides, L. & Tooby, J. (in press). Origins of domain specificity: The evolution of functional organization. In S. Gelman & L. Hirschfeld (Eds.), *Domain specificity in cognition and culture*. New York: Cambridge University Press.

Cosmides, L. & Tooby, J. (under review). Are humans good intuitive statisticians after all? Rethinking some conclusions of the literature on judgment under uncertainty.

Daly, M. & Wilson, M. (1984). *Sex, evolution, and behavior.* Second Edition. Boston: Willard Grant.

Dawkins, R. (1976). *The selfish gene.* New York: Oxford University Press.

Dawkins, R. (1982). *The extended phenotype.* San Francisco: W. H. Freeman.

Dawkins, R. (1986). *The blind watchmaker.* New York: Norton.

Dennett, D. (1978). *Brainstorms.* Cambridge, MA: MIT Press.

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research, 50*, 123-129.

Flavell, J. H., Zhang, X-D, Zou, H., Dong, Q., & Qui, S. (1983). A comparison of the appearance-reality distinction in the People's Republic of China and the United States. *Cognitive Psychology, 15*, 459-466.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review, 94*, 427-438.

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

Garcia, J. (1990). Learning without memory. *Journal of Cognitive Neuroscience, 2*, 287-305.

Gardner, D., Harris, P. L., Ohmoto, M., & Hamazaki, T. (1988). Japanese children's understanding of the distinction between real and apparent emotion. *International Journal of Behavioral Development, 11*, 203-218.

Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science, 14*, 79-106.

Gelman, S. & Markman, E. (1986). Categories and induction in young children. *Cognition, 23*, 183-208.

Gelman, S. & Markman, E. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*, 1532-1540.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

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

Gigerenzer, G. (1991) How to make cognitive illusions disappear: Beyond heuristics and biases. *European Review of Social Psychology, 2*, 83- 115.

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

Gigerenzer, G. & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating and perspective change. *Cognition, 43*, 127-171.

Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikean theory of confidence. *Psychological Review, 98*, 506-528.

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

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Kruger, L. (1989). *The empire of chance: How probability changed science and everyday life.* Cambridge: Cambridge University Press.

Gould, S. J. & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London, 205*, 581-598.

Hacking, I. (1965). *Logic of statistical inference.* Cambridge: Cambridge University Press.

Hasher, L. & Zacks, R. T. (1979) Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356-388.

Hirschfeld, L. (1989). Rethinking the acquisition of kinship terms. *International Journal of Behavioral Development, 12*, 541-568.
Hirschfeld, L. & Gelman, S. (in press). Domain specificity in cognition and culture. Cambridge, UK: Cambridge University Press.
Jackendoff, R. (1992). *Languages of the mind.* Cambridge, MA: MIT Press.
Johnson-Laird, P. N. (1982). Thinking as a skill. *Quarterly Journal of Experimental Psychology, 34A*, 1-29.
Kahneman, D., Slovic, P, & Tversky, A. (Eds). (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge, UK: 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. (1982) On the study of statistical intuitions. *Cognition, 11*, 123-141.
Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA: MIT Press.
Kimura, M. (1983). *The neutral theory of molecular evolution.* Cambridge: Cambridge University Press.
Krebs, J. R. & Davies, N. B. (1987). *An introduction to behavioural ecology.* (Second edition). Oxford: Blackwell Scientific.
Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review, 94*, 412-426.
Leslie, A. M. (1988). The necessity of illusion: Perception and thought in infancy. In L. Weiskrantz (Ed.), *Thought without language.* Oxford: Clarendon Press. pp. 185-210.
Leslie, A. M. & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*, 265-288.
Leslie, A. M. & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition, 43*, 225-251.
Manktelow, K. I. & Over, D. (1990a). *Inference and Understanding: A philosophical and psychological perspective.* London: Routledge.
Manktelow, K. I. & Over, D. (1990b). Deontic thought and the selection task. In K. J. Gilhooly, M. T. G. Keane, R. H. Logie, & G. Erdos (Eds.), *Lines of thinking* (Vol. 1). London: Wiley.
Manktelow, K. I. & Over, D. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition, 39*, 85-105.
Markman, E. M. (1989). *Categorization and naming in children: Problems of induction.* Cambridge, MA: MIT Press.
Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* San Francisco: Freeman.
McCauley, C. & Stitt, C. L. (1978). An individual and quantitative measure of stereotypes. *Journal of Personality and Social Psychology, 36*, 929-940.
Mineka, S. & Cook, M. (1988). Social learning and the acquisition of snake fear in monkeys. In T. R. Zentall & B. G. Galef (Eds.), *Social learning: Psychological and biological perspectives.* Hillsdale, NJ: Erlbaum. pp. 51-73.
Neisser, U. (1982). Memory: What are the important questions? In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts.* San Francisco: W. H. Freeman.

Nisbett, R. E. & Ross, L. (1980) *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: MIT Press.

Pinker, S. (1979). Formal models of language learning. *Cognition, 7*, 217-283.

Pinker, S. (1982). A theory of the acquisition of lexical interpretive grammars. In J. Bresnan (Ed.), *The mental representation of grammatical relations.* Cambridge, MA: MIT Press.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press. Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure.* Cambridge, MA: MIT Press.

Pinker, S. (1991). Rules of language. *Science, 253*, 530-535.

Pinker, S. & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707-784. Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28*, 73-193.

Poggio, T., Torre, V. & Koch, C. (1985). Computational vision and regularization theory. *Nature, 317*, 314-319.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition, 36*, 1-16.

Ramachadran, V. S. (1990). Visual perception in people and machines. In A. Blake & T. Troscianko (Eds.), *AI and the eye.* New York: Wiley.

Real, L. A. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science, 253*, 980-986.

Real, L. & Caraco, T. (1986). Risk and foraging in stochastic environments: theory and evidence. *Annual Review of Ecology and Systematics, 17*, 371-390.

Rozin, P. (1976). The evolution of intelligence and access to the cognitive unconscious. In J. M. Sprague & A. N. Epstein (Eds.), *Progress in psychobiology and physiological psychology.* New York: Academic Press.

Rozin, P. & Schull, J. (1988). The adaptive-evolutionary point of view in experimental psychology. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Steven's handbook of experimental psychology.* New York: Wiley.

Shepard, R. N. (1981). Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization.* Hillsdale, NJ: Erlbaum.

Shepard, R. N. (1984). Ecological constraints on internal representations: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review, 91*, 417-447.

Shepard, R. N. (1987a). Evolution of a mesh between principles of the mind and regularities of the world. In J. Dupre (Ed.), *The latest on the best: Essays on evolution and optimality.* Cambridge, MA: MIT Press.

Shepard, R. N. (1987b). Towards a universal law of generalization for psychological science. *Science, 237*, 1317-1323.

Shepard, R. N. (1992). The perceptual organization of colors: An adaptation to regularities of the terrestrial world? In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture.* New York: Oxford University Press.

Sherry, D. F. & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review, 94*, 439-454.

Shiffrar, M. & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science, 1*, 257-264.

Spelke, E. S. (1988). The origins of physical knowledge. In L. Weiskrantz (Ed.), *Thought without language*. Oxford: Clarendon Press. pp. 168-184.

Spelke, E. (1990). Principles of object perception. *Cognitive Science, 14*, 29-56.

Spelke, E. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind.* Hillsdale, NH: Erlbaum. pp. 133-169.

Springer, K. (1992). Children's awareness of the biological implications of kinship. *Child Development, 63*, 950-959.

Springer, K. & Keil, F. (1991). Early differentiation of causal mechanisms appropriate to biological and nonbiological kinds. *Child Development, 62*, 767-781.

Staddon, J. E. R. (1988). Learning as inference. In R. C. Bolles & M. D. Beecher (Eds.), *Evolution and learning.* Hillsdale, NJ: Erlbaum.

Stich, S. P. (1985). Could man be an irrational animal? *Synthese, 64,* 115-135.

Tooby, J. & Cosmides, L. (1990a). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology, 11*, 375-424.

Tooby, J. & Cosmides, L. (1990b). On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. *Journal of Personality, 58*, 17-67.

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

Tooby, J. & Cosmides, L. (in prep.). The logic of threat: Evidence for another cognitive adaptation?

Tooby, J. & DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. In W. G. Kinzey (Ed.), *The evolution of human behavior: Primate models.* New York: SUNY Press.

Tversky, A. & Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124-1131.

Tversky, A. & Kahneman, D. (1982) Evidential impact of base rates. In Kahneman, D., Slovic, P, & Tversky, A., (Eds), *Judgment under uncertainty: Heuristics and biases.* Cambridge, UK: Cambridge University Press.

Wason, P. C. & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content.* London: Batsford.

Wellman, H. M. (1990). *The child's theory of mind.* Cambridge, MA: MIT Press.

Wexler, K. & Culicover, P. (1980). *Formal principles of language acquisition.* Cambridge, MA: MIT Press.

Williams, G. C. (1966). *Adaptation and natural selection.* Princeton: Princeton University Press.

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103-128.