13 What we reason about and why

How evolution explains reasoning

Gary L. Brase

One of the more distinctive characteristics of evolutionary approaches to understanding human reasoning (or, indeed, to understanding the human mind in general) is its insistence on domain specificity. That is, evolutionary accounts of human reasoning propose that the vast bulk of reasoning that people normally engage in is done by cognitive processes that are specialized to work within a specific topic area (i.e., domain). The best known of these accounts is the work by Cosmides and Tooby (1989, 1992), that focuses on reasoning about social exchanges. A social exchange, briefly, is an exchange between two individuals that has the form of “If you take Benefit X, then you must pay Cost Y” – for example, “If you take this book, then you pay the cashier $20”. A violation of this arrangement (“cheating”) occurs when someone takes a benefit without paying the reciprocal cost (e.g., taking this book without paying for it), and this is a violation regardless of whether or not it is incorrect according to standards such as deductive logic. Other reasoning domains, involving other specialized cognitive processes, have been similarly proposed for reasoning about threats, precautions, social dominance hierarchies, social group memberships, objects, physical causality, artifacts, language, and mental states.

This domain specificity thesis is controversial for a number of reasons. It creates issues regarding how one should parse reasoning up into different domains. Domain specificity on the scale proposed by evolutionary psychologists also raises issues of how the mind can be organized to manage such a constellation of different and distinct mechanisms. For some traditional psychologists studying human reasoning, however, there may be a simpler cause for worry over the domain specificity thesis: Traditional theories of human reasoning are fundamentally based on extremely domain-general assumptions, which consider the reasoning process to be fundamentally the same across all subject matters. In a very real sense, then, domain-specific theories of human reasoning threaten the very foundations of some of the most prominent theories of human reasoning.
WHY NOT DEDUCTION AS THE DOMAIN?

Although the reasoning literature often labels certain theories as being domain specific or domain general, this is actually a false dichotomy. All theories are domain specific to some degree. Traditional reasoning theories such as the mental models approach, for instance, only cover reasoning; they do not address the areas of vision, motor movement, or interpersonal attraction. In fact, it is clear that the mind simply cannot be completely domain general or content independent. Such a general purpose system – with parameters open enough to accommodate vision, reasoning, courtship, and locomotion, to name but a few – would necessarily be a “weak” system (in the artificial intelligence sense; see Newell & Simon, 1972; a system in which the parameters are minimally constrained). Weak systems can solve certain well-defined problems (e.g., finding a pattern of letters in a database of text) and they can be quite powerful within such contexts, but they cannot deal with the ill-defined problems (e.g., finding a suitable spouse) that make up nearly all real-life problems (Simon, 1973). The debate is, therefore, not really if the mind has specific domains or not, but rather about how specific the domains are within the mind.

The conventional view of human reasoning has been that the only domain is that of deduction. The idea of studying deductive reasoning as a circumscribed area of inquiry is based on the notion that there are normative guidelines – typically formal deductive logic – about how reasoning is properly performed and that these guides transcend the particular contents or topics about which one is reasoning. For example, the conditional if $p$, then $q$ allows certain deductive inferences (if $p$ is true, then $q$ is true, and if $q$ is false, then $p$ is false) regardless of what information is used to replace “$p$” and “$q$”. Of course, one of the things everybody can agree on is that people do not always follow these guidelines. Some aspects of deductive logic are very difficult for people to apply in their reasoning, whereas other aspects of people’s reasoning go far beyond what deductive logic would allow. Nevertheless, these theories assert that the computational problem for which reasoning exists is performing deductive logic, and the two systems – reasoning and deductive logic – should therefore be rather similar. The issues, then, became those of describing what aspects of reasoning mirror deductive logic, how those aspects are cognitively represented, and how those representations also lead to the deviations between human reasoning and formal logic. A classic division within deduction-based theories of human reasoning has to do with whether people reason with rules akin to (but not necessarily the same as) deduction (e.g., Braine, 1978; Braine et al., 1995; Rips, 1994), or people reason with more fluid representations, or models, of situations (Johnson-Laird, 1983, 2001; Johnson-Laird & Byrne, 1991, 1993).
Mental rules and mental models

Mental rule theories propose, in general, that the mind uses rule systems to reason. These rules may or may not be the rules of deductive logic, but they approximate deduction and attempt to serve the same purpose – finding the formal truth in situations. Deviations in reasoning from the results expected by deductive logic are considered the anomalies in need of explanations. Thus, the rules postulated in mental rule theories, while formal-logical in nature, also include additional assumptions to explain human reasoning behaviours. For example, the PSYCOP reasoning model proposed by Rips (1994) includes deductive inference rules such as “and” elimination, Double Negative Elimination, and Modus Ponens, but then adds that prior experiences may be used as the basis for some inferences. Other modifications and addenda to formal logic invoke various cognitive processing limitations, accessory processes, or constraints that interfere with (or augment) logical processing. Rule-based reasoning theories are generally quite well specified in terms of computational components, but as a class of theories they are difficult to conclusively invalidate. The reason for this situation is the large number of algorithm sets that could be used to satisfy the particular computational goals established for a rule-based theory. Any one set of rules or procedures is testable and potentially falsifiable, but the rules can merely be modified to avoid conclusive falsification as a class of reasoning theories (just as version 2.0 of a computer software program can be a disaster, but modifications to produce version 3.0 may result in a perfectly successful program; see Johnson-Laird, Byrne, & Schaeken, 1994).

A viewpoint traditionally given as the counter-theory to mental rule theories is the mental model theory, advocated by Johnson-Laird and colleagues (e.g., Johnson-Laird, 1983, 2001; Johnson-Laird & Byrne, 1991). Briefly, the mental models theory of reasoning proposes that people form an internal representation of the reasoning problem (a model), generate a tentative conclusion, and then search for alternative possible models that would falsify the original representation. As there is some uncertainty in these representations – that is, premises can be interpreted in more than one way – there is room for people to generate alternative models for a single syllogism. The theory of mental models is a Popperian hypothesis-falsification strategy (Popper, 1959) with an implicit goal of establishing the truth (though not necessarily via formal logic), assumed to be hampered by a human inability to always perform more complex computations. What initial model is generated can be influenced by semantic knowledge, and how exhaustively the initial model is evaluated is determined by cognitive processing limitations and/or limited cognitive efforts. The mental models theory also adds computational elements (“extra-logical constraints”) that fit the theory to known data (specifically, that conclusion must conserve semantic information, be parsimonious, and assert something new). In mental models the processes are now more vaguely specified than in the

How evolution explains reasoning 311
mental rules account and no longer bound so rigidly to deductive competence, but the essential idea of deduction supplemented by additional processes remains unchanged. Furthermore, although the purpose for which reasoning exists has changed – from finding logical truth to finding any kind of useful and true conclusion – this is more of a shift in emphasis than a theoretical chasm. In contrast to mental rule theories, Johnson-Laird has promoted mental model theory as being simple to refute in principle (e.g., Johnson-Laird et al., 1994), but it is frustratingly difficult to refute in practice due to the vague nature of the mental model’s procedures (Rips, 1986).

Research on models and rules
Both the mental rules and mental models views tend to focus on the syntax of the reasoning situation, as opposed to the content of the reasoning situation, when conducting research. This can easily be seen by looking at some of the reasoning items that studies from either of these perspectives use:

If there is a red or a green marble in the box, then there is a blue marble in the box.
Which is more likely to be in the box: the red marble or the blue marble?

(Johnson-Laird, Legrenzi, Gorotto, Legrenzi, & Caverni, 1999, p. 64)

IF Betty is in Little Rock THEN Ellen in is Hammond.
Phoebe is in Tucson AND Sandra is in Memphis.
IF Betty is in Little Rock THEN (Ellen is in Hammond AND Sandra is in Memphis).

(Rips, 1994, p. 105)

It does not matter what contents are in the above sentences; what matters are the connecting terms (and, or, not, if, then) and their relative positions. In fact, the contents in the above tasks are specifically designed not to relate to anything that people normally encounter in the real world. Thus, although mental models and mental rules could be adequate explanations of how people reach conclusion in tasks like the above examples, reasoning in the real world is almost never about marbles in boxes or four strangers’ locations in four cities. Instead, people reason about the situations, events, and people in their lives, and the contents (details about the situation, event, or person) are as important – if not more important – than the connecting terms. For these traditional views, however, reasoned inferences that are based in some way on the content of the reasoning task rather than on the logical connecting terms are often called “content effects” and are actually viewed as problematic.

Research done on the nature of content effects, from a domain-general viewpoint, have tended to produce fairly general explanations. For instance,
one major explanation for content effects is the belief bias: the tendency to either accept conclusions known to be true or desirable even if they are logically unwarranted or reject conclusions known to be false or undesirable even if they are logically demanded (e.g., “All dangerous drugs are illegal; Cigarettes are dangerous drugs; Therefore, Cigarettes are illegal”; Evans, Barston, & Pollard, 1983; Markovits & Nantel, 1989; Newstead & Evans, 1993; Newstead, Pollard, Evans, & Allen, 1992; Oakhill & Garnham, 1993). Although there certainly is a phenomenon that is well described by the existence of belief biases, the use of this general explanation to resolve any type of inference that does not fit with a domain-general theory of reasoning seems suspect. First of all, this creates a situation in which nearly all research results can be explained post hoc (e.g., either a person reaches the correct conclusion by the ascribed process, or she reaches the incorrect conclusion because of some belief bias). Second, a belief bias is more of a label for what is occurring than it is an explanation of why it occurs, and specifically, it does not provide much guidance on when and why different beliefs will tend to bias reasoning processes in everyday life (Vadeboncoeur & Markovits, 1999). Third, the explanation of belief biases seems to cover a vast range of inference types that may include processes more complex than beliefs simply overriding deductive competence. Fourth, and perhaps most important, reasoning in the real world is always about something; it is always rich with content. People (aside from working logicians) do not just “reason” in an abstract manner. We use reasoning to understand the behaviours of others. We reason to understand why our car won’t start. We use reasoning to decide upon our own behaviour. Although people seem to be able – with some effort – to reason about “If A, then B; given B is true; what follows?” it is not a usual human activity.

In summary, domain-general accounts of human reasoning may provide an explanation of how people solve fairly abstract reasoning tasks in laboratories, but the work of describing real-world, everyday reasoning (sometimes called “practical reasoning”; Audi, 1989) is largely shunted to accessory mechanisms or simply dismissed as content effects of little theoretical significance. The fact that there may be problems with a domain-general account of human reasoning, however, does not automatically lead to the conclusion that reasoning must be domain specific. It is also necessary to show that there are good reasons in favour of domain specificity as the more plausible explanation.

**WHY DOMAIN SPECIFICITY?**

Over the years, different people have argued for increasingly domain-specific abilities in various areas of psychology, with the fields of perception and psycholinguistics having the oldest histories of recognized domain-specific abilities. Within these fields, the tools of brain lesions (accidental and
experimental), comparative animal studies, and computer modelling have shone an intense light on the modular structure of perception and language (e.g., Gardner, 1974; Gopnik & Crago, 1991; Kolb & Whishaw, 1990; Pinker, 1984, 1994; Pinker & Bloom, 1990). From line detectors and motion detectors, to a universal grammar and the sensory and motor homunculi – it has been continually revealed within these fields that the mind must have a functionally domain-specific cognitive architecture simply to account for the things that people manifestly do.

More recently, similar levels of modularization have been proposed within other areas of psychology. Researchers in the field of developmental psychology, for example, have developed theories and experimental support for domain-specific cognitive abilities that infants have at birth (Hirschfeld & Gelman, 1994), as well as domain-specific abilities that have developmental sequences (e.g., Baron-Cohen, 1994, 1995; the Theory of Mind Module [ToMM], Leslie, 1987). A more pervasive view of domain specificity is proposed by a general perspective known as evolutionary psychology. In one sense, evolutionary psychology is simply the study of psychology while taking into account the information known about evolutionary biology and the evolutionary history of the species being studied (i.e., anthropology and archaeology). Evolutionary psychology is furthermore often regarded as entailing some specific conclusions about how the mind is designed and how it works. These basic premises of evolutionary psychology are not shared universally (see, e.g., Holcomb, 2001), but this specific view of evolutionary psychology is important for several reasons: (1) it is clearly the dominant evolutionary view in the research field of human reasoning, (2) it strongly contrasts with the domain-general views of human reasoning, whereas many other evolutionary views can be understood as intermediate positions, and (3) it is clearly enunciated in terms of first principles derived from evolutionary biology and other fields. These first principles that form the foundation of evolutionary psychology are used to generate the conclusions that the mind is domain specific on a very extensive scale (sometimes called multimodular or even “massively modular”), that these domain-specific modules should be species-typical and common to nearly all people, and that the domains should map onto specific aspects of the world.

**Evolutionary selection pressures and modularity**

Look at your thumb for a minute. The human thumb is a remarkable and unique human evolved adaptation. At some point in evolutionary history, our ancestors began using tools. Those who had slightly better tool-handling skills were able to dig out more food, defend themselves a bit better, and attack opponents or prey with weapons a bit more accurately. These slight advantages had some partial basis in genetics, so these good early tool users passed on their predispositions to their children. Generation after generation the slightly better tool users survived and helped their children survive better
than the slightly inferior tool users. The result today is a thumb that is physically different from the other digits, can directly oppose the other digits, and that has muscles specifically for tool use (Marzke & Wullstein, 1996). The forces that created the human thumb over thousands of generations are called evolutionary selection pressures (they select, by differential survival and reproduction, the individuals in a species that are better adapted to that species’ environmental circumstances). The human thumb – the result of evolutionary selection pressures – is an evolved adaptation.

One of the necessary features of an evolutionary selection pressure is that it must be based on some aspect of the environment that posed a challenge to the species in question (e.g., the advantages of tool use included obtaining food to prevent starving, and defence against large predators). If there was no selection pressure “for” a particular adaptation, then the adaptation could not arise. The environment is not just generally a dangerous place, though; it is made up of all sorts of rather specific dangers and problems. There are dangers from large predators (lions, tigers, and bears), small predators (viruses and other pathogens), failing to satisfy your own needs (food, shelter, etc.), and even from conspecifics that mean you harm. In other words, evolutionary selection pressures are typically based on specific problems (i.e., dangers, threats, dilemmas) in the environment. Specific problems, to be effectively and efficiently solved, require specific solutions.

Figure 13.1 The fact that the human body and mind is a collection of domain specific adaptations is illustrated by the absurdity of the opposite alternative.
This basic property of the evolutionary process is a specific example of the principle described earlier, in which it was discovered by computer scientists that problem-solving systems have to be specific to the problem domain in order to be “strong” systems (i.e., able to solve the problem reliably, accurately, and efficiently). There are other examples of this principle as well. Philosophers recognized it as a general problem of induction (Quine, 1960), Chomsky (1975) recognized it in the domain of language acquisition as the “poverty of the stimuli”, it is known in the field of artificial intelligence as the “frame problem”, and in developmental psychology it motivates the idea of developmental “constraints”. In fact, even within the field of logic there is some recognition that specific types of situations call for correspondingly specific forms of reasoning. Evaluating the truth of a conclusion derived from material statements calls for propositional logic; evaluating the violation of social rules calls for deontic logic.

The conclusion from all this is that one of the fundamental tenets of evolutionary psychology is that evolutionary selection pressures led to domain-specific adaptations, and therefore to a mind that is composed of many specific cognitive adaptations (sometimes called modules, leading to the label of multimodular, or massively modular, mind).

A universal human nature

No two people are completely alike; even identical twins can think, act, and to some extent look different from one another. At a certain level, however, all human beings share a common and universal human nature. Part of this universal human nature is the opposable thumb, and the ability to walk bipedally, and the capacity to learn language. From a biological and evolutionary standpoint, humans are all part of a single species and therefore there are certain features that all of us have and that define us as that species. Some people are more aggressive than others, some people are more intelligent than others, some people walk faster than others. But it is universal of all humans that we engage in aggression, thinking, and walking. Furthermore, all these characteristics are part of our human nature because they were developed over evolutionary history in response to various selection pressures. An extension of the fundamental tenet of a multimodular mind is that all normally developing humans have this multimodal cognitive structure, by virtue of being human. Certainly specific individuals can utilize different cognitive abilities (modules, adaptations) to greater or lesser extents than other individuals, but the only persons who should entirely lack one of these abilities are individuals with traumatic brain damage or developmental disorders, or who were raised in extremely unusual environments.

The environment of evolutionary adaptation

What is an “extremely unusual environment”? The answer, from an evolutionary perspective, is not based on the current world in which we all live.
Because our minds are the products of millions of years of evolution (that is, millions of years of evolutionary selection pressures producing adaptive solutions in the form of cognitive mechanisms), the benchmark “normal” environment is something called the environment of evolutionary adaptation (or EEA). The EEA is the world in which our ancestors lived. Many people generate images of the African savannah or a tropical rainforest as the EEA, but that is somewhat misleading. The evolutionary history of humans did not take place in one location, but across an array of settings, across millions of years. In fact, the EEA is a statistical composite of all the different environments in which our human, hominid, primate, and mammalian ancestors lived. Thus, while the EEA almost certainly includes savannah and forest environments, neither is even approximately a complete description. In some ways the generality of the EEA is limiting – one cannot say exactly in what particular environment humans evolved (Lewontin, 1990). The EEA is also, however, quite informative: humans did, at some points in evolutionary history, live in different environments and we should expect adaptations for these different environments to co-exist in human nature. There are also aspects of the human EEA that have reliably existed for many thousands (or millions) of generations: humans have lived in a world with physical objects, gravity, food, predators, prey, and other humans. Studies in cultural anthropology are useful not only for the information they provide about the different physical and social environments in which humans live, but also because they tell us some of the universal features of human environments (Brown, 1991).

WHAT ARE THE DOMAINS?

Having made a case for the insufficiency of domain-general models of reasoning and provided reasons to support the idea that human reasoning should be expected instead to be domain specific, one immediately paramount question is how to split human reasoning into domains. Although general principles of the evolutionary process and of information processing can tell us that domain specificity is the likely structure of the human mind, it does not tell us what the specific domains are within this structure. This is an issue, essentially, of deciding how best to carve nature at its joints.

The most widely recognized evolutionary approach to domain-specific human reasoning is that of Cosmides and Tooby, who clearly indicate that their overall view postulates the existence of many computational reasoning abilities within different domains (see, for example, Cosmides & Tooby, 1992, pp. 179–180). So what specific elements in our evolutionary history can be expected to have constituted selection pressures recurrent enough to have created their own reasoning processes? Put another way, what cognitive tasks have humans faced that could have been important, enduring, and complex enough to have evolved their own reasoning processes? Making a list of this
sort is fraught with conceptual and practical issues and debates, not the least of which is that there can be disagreement about specific items being added to or removed from the list. There are also problems about specifying the level of detail at which one should “carve” different domains from each other (Atkinson & Wheeler, 2001; Sterelny & Griffiths, 1999). Steen (2001) has constructed a much larger list of cognitive adaptations: abilities for which arguments have been made that they are evolved parts of human nature, and in some ways the current list can be viewed as a subset of that (similarly tentative and prospective) list. Specifically, the following list attempts to identify some areas in which there are both theoretical reasons to expect, and experimental evidence to support, the existence of cognitive adaptations that have as part of their functionality some type of reasoning/inference procedure that is invoked within that domain. And in particular, this list focuses on “higher-order” reasoning processes that are partially or fully under conscious control, which is usually the type of reasoning studied under the topic of human reasoning (“lower-order” inferential processes such as recovering visual depth from retinal information can be considered as “reasoning”, but not with the meaning usually considered by human reasoning researchers).

**Physical interaction domains**

From a very young age, people appear to have an intuitive understanding about physical (i.e., non-living) objects and how they behave. Even infants understand that items which move as a unit are whole objects, that objects do not move on their own (i.e., no action at a distance), and that objects generally obey basic laws of causality (e.g., classic Newtonian mechanics; Spelke, 1990; Spelke, Breinlinger, Macomber, & Jacobson, 1992). Older children and adults understand that objects have properties (rather than the opposite relationship), that one object can have several properties along different dimensions, and that different objects can have the same property without being the same in other ways. For example, people think it sensible to say “all rocks are grey” but not “all grey things are rocks”, and “All rocks are grey; my shirt is grey” does not mean “my shirt is a rock”.

In many ways human reasoning based on intuitive physics and the nature of objects appears to be a plausible toe-hold that can be used to learn abstract and non-intuitive reasoning processes such as formal deductive logic. Intuitive physics, which dominates many areas of everyday thinking (because most things in the world are physical objects) may also contribute to the appeal of ideas such as the mental models theory of reasoning, in which there is an implied analogy of “constructing” mental models as if they were physical entities.

Children also develop specific ideas about the nature of biological organisms, as opposed to non-living objects (artifacts; Gelman, 1988; Gelman & Markman, 1986; Keil, 1986; 1994, 1995). Biological entities are
seen as intentional agents and as having an “essence” that cannot be altered. For example, most people have no objection to “If a soup can is filled with birdseed, punched with holes for access, and hung from a tree, it is a birdfeeder” – regardless of what the can was before or what it previously contained. Children over the age of about 7, however, will reject the analogous idea that “If a black cat is painted with a white stripe and smelly stuff is put in a pouch inside it for squirting at other animals, it is a skunk”.

There are several established cognitive adaptations that involve the evaluation of foods in regards to what is (and is not) safe to eat. These adaptations include a learned taste aversion system (Garcia & Koeling, 1966), motion sickness (visual/vestibular discrepancies can indicate the ingestion of a toxin, leading to regurgitation of recently eaten food; Yates, Miller, & Lucot, 1998), and pregnancy sickness (to avoid the ingestion of teratogens in the first trimester; Profett, 1992). While these adaptations certainly involve some forms of inferences (e.g., from sickness to dislike of certain foods), they are relatively subconscious and therefore have only some resemblance to human reasoning as it is usually considered (i.e., higher-order, conscious, and willful inferences). It does appear, though, that evaluations of food can involve conscious reasoning processes as well, particularly in the service of evaluating the social suitability or appropriateness of a food for consumption (Occhipinti & Siegal, 1994; Rozin, Fallon & Augustoni-Ziskind, 1985).

Language and intention domains

Language has long been an area in which domain-specific abilities are thought to exist (Chomsky, 1975). There are many implicit reasoning processes involved in language (e.g., for conjugating verbs and forming tenses) but there are also some higher-level reasoning processes. There is evidence that specific forms of reasoning exist, beginning in childhood, for individuals to infer the meanings of novel words (Pinker, 1984, 1994). For example, children (and adults) tend to infer that a new word provided in reference to a novel item refers to the entire object, but a new word that refers to an item that already has a label is assumed to refer to some part or property of the whole object (Markman, 1987, 1990). Reasoning processes tied to language have been recognized as being relevant to human reasoning generally (e.g., Braine & O’Brien, 1991, 1998; Brooks & Braine, 1996), but within a domain-general approach these effects have largely been seen as phenomena that either impair or extend beyond “correct” reasoning performance. The point here is that these reasoning processes can be better understood and appreciated if viewed as domain-specific inference mechanisms for language development and use.

People reason about intentionality, including at the most basic level of whether or not something has intentions. If something has self-propelled movement then people tend to spontaneously infer that it has intentions (i.e,
it has beliefs and desires that guide its actions; Dasser, Ulbaek, & Premack, 1989; Premack, 1990). People use eye direction as an indication of what specific things an individual has an interest in or wants or desires (Baron-Cohen, 1994, 1995; Lee, Eskritt, Symons, & Muir, 1998). The basic rule that even young children use is that if someone gazes directly at something, they want/are interested in that thing. People also use facial expressions in a similar manner to infer the emotional states of other individuals (e.g., Knutson, 1996; Segerstrale & Molnar, 1997). At a higher level of consideration there is the ability to reason generally about the state of another’s mind (called mindreading, or the theory of minds module; Baron-Cohen, 1994, 1995). This mindreading ability includes reasoning about false beliefs (e.g., Mary thinks there is a King of France), deception (Tom told Mary there is a King of France), and second-order belief states (Tom thinks it is funny that Mary thinks there is a King of France). Furthermore, there are reasoning processes that have been suggested to exist in order to manage and use information about the sources of information (e.g., Information from Tom is unreliable; Tooby & Cosmides, 2000).

Social interaction domains

Many specific categories, or domains, of social interactions also involve specific and adaptively specialized reasoning processes. The most well-studied area of domain-specific social reasoning is that of social exchanges (also called social contracts).

Cosmides and Tooby (Cosmides, 1989; Cosmides & Tooby, 1989, 1991, 1992, 1997) proposed that people possess, as part of the evolved architecture of the human mind, the abilities to initiate, recognize, and evaluate social exchanges: situations in which two people agree (i.e., form a contract) to exchange items for mutual net benefits. These social contracts can be stated in ways that can be analyzed according to deductively correct responses (such as the example at the beginning of this chapter), but the theory predicts that people will eschew formal deduction for the responses that are adaptive within the specific domain of social contracts. Several studies have found that social contract contents, but not similar non-contract contents, lead to the selection of the “cost not paid” and “benefit taken” as those that could violate the conditional rule, regardless of whether these are the deductively prescribed selections or not (Cosmides, 1989; Cosmides & Tooby, 1992).

Similar areas of domain-specific social reasoning have subsequently been proposed, including reasoning about social threats (Rutherford, Tooby, & Cosmides, 1996) and reasoning about precautions and hazard avoidance (Fiddick, Cosmides, & Tooby, 1995, 2000; Pereyra, 2000). In both these areas there is now experimental evidence to support reasoning processes that are not only adaptive within these specific situations, but are different both from social exchange reasoning and from each other.
Finally, there are reasoning processes that have been hypothesized to exist for specific domains having to do with social groups and social networks. Research by Cummins (1996, 1999) has found that the now established findings of domain-specific social contract reasoning are also influenced by the social status of the participants. She has suggested that this reflects collateral (or alternative) reasoning processes that involve the negotiation of social dominance hierarchies. Other research on reasoning about social groups has found specific patterns of inferences that are made in relation to coalitional markers (i.e., the physical indicators such as ribbons, badges, and cards that people use to identify themselves as group members; Brase, 2001). For example, when people are told “If a person is a member of the Fishermen Club, then he wears a silver badge” and “This person wears a silver badge”, they are very likely to conclude that “This person is a member of the Fishermen Club” (which is deductively invalid; compare to the earlier example of a grey shirt/grey rock).

This is almost certainly not a complete list, nor it is a list that is even at this time a foundational list of certain reasoning domains. It is, for better or worse, an initial cataloging of suggestions that have been made by several different researchers regarding evolved domain-specific reasoning abilities. Like a mock-up construction of a car, or a model of a building before it is actually cast in steel, brick, and cement, this tentative list allows us to see the overall effect and the further implications and issues that arise from domain-specific reasoning.

Properties and interactions of reasoning domains

There have been some confusions, debates, and misunderstandings about how these various domains of human reasoning function, both singularly and in connection with one another. One area in which confusion sometimes occurs has to do with the borders of each domain in which reasoning processes are proposed to exist. Initial reactions to Cosmides and Tooby’s work on reasoning about social exchanges illustrate this confusion well. Some argued that any finding of very good reasoning abilities that occurred outside the context of social exchanges would invalidate the theory of domain-specific abilities for reasoning about social exchanges (Cheng & Holyoak, 1989; Manktelow & Over, 1991; Rips, 1994). This argument may be the result of an interpretation of social exchange reasoning as a domain-general theory of human reasoning – i.e., the claim by Cosmides (1989, p. 200) that “no thematic rule that is not a social contract has ever produced a content effect that is both robust and replicable” was read as being a claim that the context of social exchanges was the only domain in which facilitation could ever exist. A closer reading of that very article, as well as both earlier and subsequent writings, shows clearly that this more extensive interpretation was never intended. Instead, what had been proposed is the multimodular mind described earlier.
Another initial reaction to the social exchange theory of human reasoning was to question some of the specific findings of facilitated reasoning that were claimed to be situations of social exchanges. For example, people tend to reason much better about the conditional rule “If you are drinking beer, then you are over 21 years old” than a similar rule (e.g., “you are drinking beer, then you are eating pizza”). How can a person’s age be an item being exchanged? If, as it seems reasonable to assume, one cannot trade in one’s age for a drink, then this is not an example of a social exchange, and there must be something other than social exchange reasoning going on. The problem here is not that of trying to have social exchange theory account for all types of good reasoning (as above), but rather of constraining the scope of application of social exchange reasoning to a very narrow area. This problem raises an important issue created by modularity: How widely (or narrowly) are modules applied?

Sperber has outlined – and provided terminology for – how cognitive modularity can be expected to work in terms of this scope of application issue – that is, how to think of module domains and the application of each module’s mechanisms. First of all there is the proper domain of a cognitive module. The proper domain is the class of situations making up the specific adaptive problem that those cognitive abilities evolved to solve. In other words, the proper domain is “all the information that it is the module’s biological function to process” (Sperber, 1994, p. 52). A social exchange situation clearly within the proper domain for social exchange reasoning would be something like, “If you watch my child for an hour, then I will feed you lunch”: there are clear benefits exchanged in such a way that both parties can end up with a net gain in their own estimation. Many modular abilities are not tightly scope-limited to their original adaptive function, however, and it is therefore also useful to consider what is called the actual domain of a cognitive module. Because humans are able to conceive of aspects of the world – and particularly the social world – in a non-literal fashion, it is possible to invoke the functional machinery of a cognitive module in a situation that is not actually within the original (i.e., proper domain) scope of that module. Specifically, a modular ability (including some reasoning processes) can sometimes be invoked by a situation that manages to meet the module’s input conditions, even if meeting those conditions is achieved by prior processing of the situational information by other systems (e.g., the beer-drinking rule discussed above relies on a conceptual interpretation of being “over 21 years old” as a cost paid in waiting for the benefit of being allowed to drink beer. (see also Brase & Miller, 2001, on reasoning about sexual harassment situations as being perceived as either social exchanges or threats).

Sperber (1994) additionally points out that a sub-part of this actual domain of application for modular abilities is what he calls the cultural domain. The cultural domain is made of those situations that people have purposefully designed so as to meet a specific module’s input conditions,
in order to reliably invoke the inferential machinery of that module. For example, the invention of money in cultures around the world appears to be a method of capitalizing on our ability to reason quite well (that is, adaptively) about social exchanges, while at the same time overcoming the inherent limitations of a barter economy (by creating a general intermediate item to transfer costs and benefits across individuals). The cultural domain of social exchange reasoning thus now extends to cover a bewildering array of situations, rather than just situations in which two people have rather specific items to exchange.

One can think of these proper and actual domains as concentric circles of light, the inner circle (proper domain) being a more intense light than the outer (actual domain). The different modular abilities can then be envisaged as multiple instances of these concentric circle arrangements, sometimes even with parts of these circles overlapping (Brase & Miller, 2001). This is not an unreasonable arrangement to propose as existing within the human mind, and in fact similar arrangements are known to form parts of the architecture of the mind (e.g., visual cues for perception of depth; Holway & Boring, 1941). This arrangement does bring up new issues to be resolved, or rather it gives a new intensity to issues that were implicitly raised before. It is now apparent that the precise input conditions for each form of domain-specific reasoning must be specified, and these specifications should at least try to be clear about the proper and actual scopes that can satisfy each input condition (e.g., see list in Cosmides & Tooby, 1992, p. 177).

Another area in which it is increasingly imperative for theories of domain-specific abilities to be clear is in the ontogeny of the ability. Because domain-specific reasoning abilities are proposed to be universal aspects of human nature (i.e., species-specific and reliably developing), they are often supposed to be heavily biologically determined. The underlying thought is a testament to the power and influence of the cultural relativism viewpoint; any ability that is proposed to be universal among humans must be biologically determined because otherwise cultural factors would quickly change or eradicate it. The problem with this position is that cultural factors simply are not universal trump cards for any and all aspects of human physiology, behaviour, and thought. Within any domain there are aspects that are heavily influenced by genetic and biological factors and other aspects that are much less influenced by genetics or biology. The human ability to produce language, for example, involves universal human features that require very little in terms of environmental input (e.g., a descended larynx, vocal cords, and suitable oral cavity need only proper nutrition, oxygen, and use to develop), and other universal human features that require more substantial environmental inputs – e.g., the universal language acquisition device requires information inputs from the environment about the which phonemes, morphemes, words, and syntax are utilized in the ambient language system(s) (Pinker, 1994). Neither strong cultural relativism (or environmental determinism) nor strong biological determinism will ever carry the day. The only responsible – and in
fact the only feasible – position to take is that the functional development and subsequent operation of domain-specific abilities are canalized by biology and guided by the environment (Cummins & Cummins, 1999).

CHALLENGES TO AN EVOLUTIONARY APPROACH

The evolutionary approach to reasoning is controversial along several fronts. Some of these sources of controversy have been covered already (for example, the challenge to domain-general approaches). There are other challenges, however, some deriving from the very idea of using an evolutionary approach, and others deriving from certain implications within the evolutionary approach.

Objections to the use of evolutionary theory

The evolutionary approach is not the only perspective that has proposed the existence of domain-specific reasoning abilities. The theory of pragmatic reasoning schemas (PRS: Cheng & Holyoak, 1985; Cheng, Holyoak, Nisbett, & Oliver, 1986; Holyoak & Cheng, 1995; Kroger, Cheng, & Holyoak, 1993) proposes that evolutionary considerations are not necessary for the development of domain-specific reasoning abilities. This theory holds that “people reason using knowledge structures that we term pragmatic reasoning schemas, which are generalized sets of rules defined in relation to classes of goals” (Cheng & Holyoak, 1985, p. 395). Pragmatic reasoning schemas produce domain-specific reasoning mechanisms via inductive learning, with no reference to evolutionary selection pressures or adaptations. Instead, specific reasoning schemas are created by individual experiences with recurring types of situations. For instance, the production rules for the permission schema are:

1. If the action is to be taken, then the precondition must be satisfied.
2. If the action is not to be taken, then the precondition need not be satisfied.
3. If the precondition is satisfied, then the action may be taken.
4. If the precondition is not satisfied, then the action must not be taken.

While these schemas are claimed to be “not equivalent to any proposed formal or natural logic of the conditional” (p. 397), the permission schema (along with the also proposed obligation schema) generally describes situations of deontic logic.

Two aspects of the theory of pragmatic reasoning schemas are troublesome in comparison to evolutionary accounts of domain-specific reasoning. First,
there is little theoretical guidance about the scope and limits of these reasoning schemas: Why do situations of permission and obligation and not other classes of goals and situations so definitely lead to reasoning schemas? If other classes of goals and situations do create reasoning schemas, how does one limit the number of schemas created in order to avoid something approaching the representation of thousands or even millions of goals and situations, each with their own unique reasoning schema? The second, and actually related, trouble with the theory of pragmatic reasoning schemas is that the knowledge structures, schemas, or production rules (i.e., computational processes and constraints) are explained as the result of induction: Ordinary life experiences are distilled into general reasoning processes. This explanation is troublesome because, for a large number of general goals (e.g., language, vision, and concept learning), there is a great deal of evidence that it is improbable – if not impossible – for induction alone to provide the basis for human performance (e.g., Gelman, 1990; Pinker, 1979; Wexler & Culicover, 1980; see, however, Holland, Holyoak, Nisbett & Thagard, 1986).

Another objection to the use of evolutionary theory is more general. Gould and Lewontin (1981; see also Lewontin, 1990) have criticized the use of evolutionary theory to create hypotheses about evolved adaptations (including cognitive adaptations) on the basis of a claim that there is not enough understanding of the evolutionary histories of species to draw even tentative conclusions about functional designs. They have summarized this criticism by labelling adaptationist explanations as “just so” stories (i.e., analogous to Rudyard Kipling’s children’s stories about how various animals got their properties). There are some useful cautions to take to heart in Gould and Lewontin’s arguments (e.g., the need for experimental research to support theories), but overall the stronger criticisms of the adaptationist approach in general have been extensively rebuked several times over (Borgia, 1994; Buss, Haselton, Shackelford, Bleske, & Wakefield, 1998; Queller, 1995; Tooby & Cosmides, 1992). Gould and Lewontin’s objection nevertheless reappears occasionally without any reference to the documented problems with this view (e.g., Johnson-Laird et al., 1999; see comment by Brase, 2002).

**Issues within an evolutionary approach**

Evolutionary theories of human reasoning achieve several useful and important ends: they provide a novel explanation for several phenomena, they make significant and new predictions about how people reason, and they move the field of psychology into closer compatibility with other fields such as biology and anthropology. Evolutionary theories about human reasoning, and specifically the notions of domain specificity and a multimodular structure for the mind, also raise new questions and issues that must be dealt with if such an approach is to ultimately prosper.

One issue that arises based on the assumption of a multimodular mind is the need for some higher-level organization and functional integration of
these modules. For example, if there are, even modestly, dozens of domain-specific reasoning abilities (i.e., modules) in the mind, then something needs to exist that cognitively organizes these different modules, determines what information is sent to which modules, and maybe even influences where the results of module functions are then sent. This sounds something like a domain-general, content-independent reasoning ability itself, and some researchers have identified this as a crucial area in which multimodular descriptions of the mind are seriously inadequate (Samuels, 1998). This certainly is a crucial issue, and one that deserves to be given considerable attention. It is not, however, a problem that invalidates the idea of a multimodal mind and it is certainly not a problem that is impossible to solve (e.g., see Barrett, 2002; Brase, 2003).

Another issue that emerges with the consideration of a multimodal mind is the issue of what exactly is a “module”. For instance (and with some liberty with terminology), people have talked about a “cheater detection” module, which is a specific aspect of what some have called the “social exchange reasoning” module, which others have argued is part of a “permissions and obligations reasoning” module. At what level (or levels) do modules exist? Atkinson and Wheeler (2001) have dubbed this the grain problem of evolutionary psychology; some modules are proposed at a very fine-grained level whereas others are proposed at more coarse-grained levels. Do these different proposals now stand in opposition to each other? Are we to allow modules to exist within larger modules? How about modules that partially overlap with several other modules? In some respects these issues may be semantic; we do not seem to have the theoretical or scientific language yet to talk about certain possible solutions to this problem. Other aspects of the grain problem, however, highlight further topics that need to be addressed regarding the functional organization and structure of a multimodal mind.

Lastly, there is an implication of domain-specific, evolutionary approaches to human reasoning that has not been considered very much at all. For many people the study of reasoning is defined in reference to the ideas of deduction, formal logic, and general abilities. In other words, the study of human reasoning is the study of human cognition in relation to these benchmark ideas. The domain-specific, multimodal conception of human reasoning belies this unitary benchmark, replacing it with benchmarks based on adaptive functionality within each domain. In practical terms, this means that research on reasoning will need to interact much more with areas such as social and developmental psychology when developing theories of “correct” reasoning.

ACKNOWLEDGEMENTS

The author thanks Sandra Brase for her unstinting advice and support, and David Farley for permission to reproduce his work in Figure 13.1.
REFERENCES


How evolution explains reasoning 329


