EVOLUTIONARY PSYCHOLOGY'S GRAIN PROBLEM AND THE COGNITIVE NEUROSCIENCE OF REASONING

Anthony P. Atkinson
Psychology Department,
King Alfred's College Winchester,
Hampshire, SO22 4NR, UK
Email: A.Atkinson@wkac.ac.uk

Michael Wheeler
Department of Philosophy,
University of Dundee,
Dundee, DD1 4HN, UK
Email: m.w.wheeler@dundee.ac.uk

Preprint of chapter published in: D. Over (Ed.) (2003), *Evolution and the psychology of thinking: The debate* (pp.61-99). Hove: Psychology Press (Current Issues in Thinking and Reasoning, K. Gilhooly Series Ed.).

Acknowledgements

A.P.A was supported by a Leverhulme Trust Research Fellowship and a grant from the McDonnell Project in Philosophy and the Neurosciences (http://www.sfu.ca/neurophilosophy/) while this chapter was written. Our thanks to Ralph Adolphs and David Over for comments.

INTRODUCTION

In its most general form, evolutionary psychology is simply psychology that is properly grounded in evolutionary biology. But most research that falls under the banner of evolutionary psychology might be given a more specific gloss, as the Darwinian adaptationist programme applied to the mind/brain. How does this idea work?[1]

Viewed through the lens of Darwinian theory, organisms are (for the most part) integrated collections of adaptations, where adaptations are phenotypic traits that are evolved responses to adaptive problems, and where adaptive problems are selection pressures — recurring environmental conditions that influence reproductive success, or fitness, of individual organisms. Adaptations, then, contribute (or once contributed) to the reproductive success of the organisms that have them. Fitness maximisation per se is not a goal of individual organisms, however. Organisms cannot seek directly to maximise their fitness, since what counts as fitness-promoting behaviour in one situation or for one individual is not likely to be so in another situation or for another individual (Cosmides & Tooby, 1987; Symons, 1992). Rather, thanks to their specific adaptations, individual organisms have correspondingly specific goals that are tied to particular aspects of their physical and social environments, and to the lives they lead in those environments, and which affect, directly or indirectly, their reproductive success. In other words, organisms maximise their fitness by solving many specific adaptive problems. Evolutionary psychology 'simply' applies this well-established Darwinian reasoning to the human brain. Thus conceived, the brain is an integrated collection of psychological åmechanisms that evolved because their behavioural effects tended to help maintain or increase the fitness of organisms whose brains contained those mechanisms. The human brain is thus viewed as a system of psychological adaptations, a system shaped by natural selection to solve many specific adaptive problems.

Of course, adaptive problems (psychological or otherwise) are not set in stone: the goal posts often move, because while certain environmental conditions, such as sunlight coming from above (Ramachandran, 1988), will remain constant for aeons, other such conditions will change across evolutionary time. Against this background, a key claim often made by evolutionary psychologists is that the adaptive problem to which some psychological mechanism is an evolved response must be specified by reference to a certain ancestral environment, what is often called the organism's environment of evolutionary adaptedness, or EEA. For humans, this claim is typically unpacked via the following argument: Given the slow pace of evolutionary change, the last time any significant modifications could have been made, by selection, to the functional architecture of the human brain, was during the Pleistocene era. The Pleistocene is thus the

---

[1] It is not our intention here to engage directly with the basic line of reasoning and argument that grounds evolutionary psychology, but what we do wish to do requires that we first briefly summarise its main features (see Atkinson & Wheeler, under review; Bloom, 1999; Cummins & Cummins, 1999; Samuels, 1998, 2000; Wheeler & Atkinson, 2001, for reviews and critical analyses).

most significant part of the human EEA. Put crudely, the thought is that we have ancient hunter-gatherer minds inhabiting 21st century environments.

A major task facing the evolutionary psychologist is therefore to show just how modern human behaviour is produced by psychological adaptations to ancestral environments. For present purposes, however, a rather different issue is to the fore. Several prominent evolutionary psychologists have argued that, on the basis of the kind of evolutionary reasoning sketched out above, we should expect the mind's information-processing architecture to consist in myriad domain-specific devices rather than a relatively small number of domain-general devices (e.g., Baron-Cohen, 1995; Buss, 1999; Cosmides & Tooby, 1987, 1994a, 1994b, 1992; Gigerenzer, 1997; Pinker, 1994, 1997; Sperber, 1996; Symons, 1992; Tooby & Cosmides, 1992, 1995, 2000). As we have explained elsewhere (Atkinson & Wheeler, under review; Wheeler & Atkinson, 2001), these arguments depend on the idea that domains are to be defined in terms of adaptive problems. So a feature of our cognitive architecture will be maximally domain-specific just when it is dedicated to solving one particular adaptive problem, and maximally domain-general just when it can contribute to the solution of any adaptive problem whatsoever. The evolutionary-psychological claim, then, is that since the brain's functional architecture is a system of psychological adaptations that reflect the vast range of adaptive problems that our hunter-gatherer ancestors faced, it must be a multifarious suite of innately specified, domain-specific devices (often called modules). This, in short, is the 'Massive Modularity Hypothesis' (MMH) (Samuels, 1998, 2000; Sperber, 1996).

While the MMH is both widely endorsed within evolutionary psychology and championed vigorously by some leading theorists, not all who are broadly sympathetic to the overall evolutionary-psychological enterprise are ardent advocates of the idea (e.g., Cummins & Cummins, 1999; Shapiro & Epstein, 1998). And the MMH has certainly attracted a good deal of critical fire from outside the discipline. That combination of pre-eminence and controversy has motivated several detailed critical examinations of the MMH’s core concepts and arguments (e.g., Atkinson & Wheeler, under review; Samuels, 1998, 2000; Wheeler & Atkinson, 2001), to which can be added the present chapter.

In their recent book on the philosophy of biology, Sterelny and Griffiths (1999) devote two chapters to evolutionary psychology. Within these two chapters, they devote little more than a couple of pages to the presentation of a problem that they see as posing a significant threat to the viability of an explanatory programme predicated on the MMH. Evolutionary psychology, Sterelny and Griffiths argue, faces a "grain problem". In essence, their conception of this problem is that attempts to identify evolved, domain-specific modules require one to fix a single level of description at which the selection pressures in play are specified, but that the process of fixing that level will often be arbitrary. In the next section, we elucidate Sterelny and Griffiths' view of evolutionary psychology's grain problem, showing that their account in fact contains two threats to evolutionary psychology, not just one. Having suggested that one of these problems is the more worrying for evolutionary psychology, we turn, in the third section, to our own vision of the grain problem. We reveal a second dimension to the difficulty that Sterelny and Griffiths do not notice (or at any rate, do not make explicit), and that transforms the grain problem into a far more serious challenge to evolutionary psychology than is present in their narrower interpretation. In the subsequent five sections, we present a detailed examination of the interdisciplinary, multilevel approach to the study of cognition offered by evolutionarily inspired cognitive neuroscience, using research on social cognition and reasoning as a case study. In doing this we show that evolutionary psychologists can and do live with the two-dimensional grain problem.

THE GRAIN PROBLEM ACCORDING TO STERELNY AND GRIFFITHS

Sterelny and Griffiths (1999) identify a problem for evolutionary psychology (in fact, it seems, for Darwinian selective reasoning in general), which comes into view when one combines the following two observations: (i) at a fairly gross level of description, a range of coarse-grained adaptive problems can be distinguished that might include obtaining sustenance, avoiding predators, selecting mates, reproducing, and caring for offspring; (ii) many such coarse-grained adaptive problems may also be appropriately described as hierarchically organised complexes of progressively finer-grained adaptive problems. Consider, for example, a problem for animals that live in complex social environments, namely, that of predicting the behaviour of conspecifics and adjusting one’s own behaviour accordingly, a solution for which, in humans at least, is the ability to ‘mindread’. What might appear to be a unitary problem at a coarse-grained level of description can nevertheless be decomposed into a set of more specific problems. Candidate component problems include: interpreting movements in terms of goals and desires, attributing perceptual states to others, attributing shared attention and knowledge, and attributing mental states (e.g., Baron-Cohen, 1995; Leslie, 1994). And each of these more specific problems might themselves be decomposed into yet finer-grained problems: attributing perceptual states, for example, might involve the separable abilities to detect eyes and to track their gaze. Or consider mate selection. Sterelny and Griffiths ask, rhetorically, "Is the problem of mate choice a single problem or a mosaic of many distinct problems?" (p. 328). Is choosing a mate a single adaptive problem, or is it a set of related problems, such as: choosing someone of the opposite sex, someone who shows good reproductive prospects, and someone who shows signs of being a good parent? Or at a yet finer-grained level of description, is the problem of choosing someone with good reproductive prospects a single problem or a set of related problems, such as choosing someone who is young, who is healthy, of high status, etc.? (See also Sterelny, 1995.) The implied outcome of this rhetorical interrogation is, of course, that there is no final answer to any of the questions posed, and that this indeterminacy is a feature of evolutionary scenarios in general (or, at least, of significantly many of them), rather than just a local nuisance specific to mate choice. If this is correct, then the dilemma facing the evolutionary theorist is not merely that adaptive problems are, typically, hierarchical and nested. Rather, the full dilemma is that, since no particular level in a selective hierarchy — or, as one might say, no individual descriptive grain — takes explanatory precedence over any other, any decision to promote a single descriptive grain as the one at which the selection pressures in operation ought to be specified is simply arbitrary. This is the difficulty that Sterelny and Griffiths dub the grain problem.[2] (For now we shall use the term 'grain problem' to indicate the grain problem as Sterelny and Griffiths see it. In the next section, however, we shall argue that this is in fact a limited version of the real grain problem.)

But how, exactly, does the fact that attempts to specify adaptive problems are encumbered by a plurality of equally viable descriptive grains come to constitute a problem for evolutionary psychology? Here, with a few clarificatory bells and whistles, is what Sterelny and Griffiths say:

> The grain problem in evolutionary psychology challenges the idea that adaptations [in this context, cognitive devices] are explained by the [adaptive] problem to which the adapted trait [some single cognitive device] is a solution. If (but only if) there is a single cognitive device that guides an organism's behavior with respect to issues of mate choice [for example], then mate choice is a single domain, and these [more specific problems] are all different aspects of the same problem. It is not the existence of a single problem confronting the organism that explains the module [the single cognitive device], but the existence of the module that explains why we think of mate choice as a single problem. (Sterelny and Griffiths, 1999, pp. 328-329).

What are we to make of this short explication of the threat posed by the grain problem? As far as we can tell, the target passage is consistent with two different interpretations of that threat, which are not clearly distinguished by Sterelny and Griffiths. Moreover, these distinct interpretations actually constitute alternative, rather than complementary, arguments against evolutionary psychology. The difference between the two possible interpretations here turns on how we unpack the two key points made in the target passage, namely:

(i) that the grain problem challenges the view that cognitive devices (as adapted traits) are explained by the adaptive problem to which those devices are a solution; and

(ii) that evolutionary psychologists will be in a position to say with certainty why some, but not other, related adaptive problems can be grouped together as a single problem, only in those cases where it is possible [illegible] solution. This principle must surely be an essential plank of any genuinely Darwinian enterprise, and so must be a non-negotiable feature of the evolutionary-psychological treatment of cognitive devices. But how, exactly, is it threatened by the grain problem? Sterelny and Griffiths are disappointingly quiet on this point, but the answer may depend on what commitments we inherit when we endorse the claim that adapted traits are explained by their associated adaptive problems. On one understanding of those commitments, the evolutionary psychologist not only requires there to be some robust and principled method for identifying the unique adaptive problem to which some adapted trait constitutes an evolutionary solution, she also needs that method to be, in a certain sense, independent of our identification of the phenotypic feature that constitutes the trait to be explained. What is meant by 'independent' here is captured by the following thought: the way in which we specify an adaptive problem must not appeal essentially to the particular phenotypic trait that, ultimately, we intend to explain as the evolved solution to that problem, because if we have no option but to appeal to the evolutionary solution in the specification of the adaptive problem, then (so the argument goes) we are building into our explanation the very thing that we are trying to explain, and that is no explanation at all.

What we have identified as point (ii) in Sterelny and Griffiths' exposition of the grain problem states that there is only one principled strategy by which the evolutionary psychologist might pin down the grain at which to specify an adaptive problem. That strategy is first to identify a distinct cognitive device (the adapted trait) subserving a distinct type of behaviour, and then to use that knowledge to constrain the choice of grain. Given the independence criterion, however, point (ii) emerges as nothing more than a powerful illustration of the disastrous failure of evolutionary psychology to solve the grain problem. Why? Because the proposed strategy contravenes the independence criterion, and so (under this interpretation) marks the failure of the evolutionary psychologist to respect the adaptationist principle that evolved traits are explained by the adaptive problems to which they are solutions. Thus if the independence criterion (or something like it) does figure in the correct unpacking of point (i) in Sterelny and Griffiths' exposition, then that fact demands a pessimistic reading of point (ii). (On the plus side, it also provides a much-needed explanation for why the second point comes hard on the heels of the first.) In effect, then, on this first interpretation of Sterelny and Griffiths' analysis, the evolutionary psychologist faces a dilemma: If she relinquishes her commitment to the independence criterion, then she can no longer hold claim to a non-negotiable principle of Darwinian explanation. But if, on the other hand, she holds fast to that prior explanatory principle, then she has to accept that in many cases there may be no non-arbitrary way of pinning down the choice of grain at which to specify adaptive problems, and thus no non-arbitrary way of explaining evolved traits in terms of the adaptive problems for which they are solutions.

A second interpretation of the passage in question comes into view if, in fact, Sterelny and Griffiths do not intend the independence criterion to be required by the view that adapted traits must be explained by their associated adaptive problems. If an adaptive problem can rightly be said to explain an adapted trait, even though that problem can be singled out only by specifying that trait, then, at first sight anyway, point (ii) becomes not an observation that simply illustrates the rout of evolutionary psychology by the grain problem, but rather an initially plausible response to the grain problem, and more specifically to the threat identified in point (i). In other words, on this interpretation, the grain problem is taken to threaten the principle that an adapted trait is explained by its associated adaptive problem simply because it suggests that the choice of grain at which the adaptive problem might be specified is often arbitrary. But without the background presence of the independence criterion, the antecedent identification of a distinct cognitive device subserving a distinct type of behaviour might legitimately be thought to meet this challenge by non-arbitrarily constraining the choice of grain.

This way of understanding Sterelny and Griffiths' argument is wholly in tune with the fact that they follow up the target passage reproduced above by immediately giving several weighty arguments for why (they think) evolutionary psychology has wedded itself prematurely to the idea of a massively modular theory of higher-level cognition, and thus for why (they think) the search for distinct cognitive devices subserving distinct types of behaviour is doomed to fail (see pp. 329-332). In other words, having proposed a possible response to the grain problem, Sterelny and Griffiths immediately mount a secondary offensive to tell us why that response won't work.

We shall summarise one of Sterelny and Griffiths' second-wave arguments against evolved modularity — partly to illustrate their position, but also because, as we shall see in a moment, the conclusion of that argument is telling. It is an undeniable fact that our psychological architecture has evolved in a social environment. Such socially-embedded evolution is inherently interactive, in that one individual's adaptive solution is another's adaptive problem. This leads to a co-evolutionary arms race, in which each of the competing strategies in the population (say, cheating and cheater-detection), always in pursuit of the evolutionary advantage, periodically transforms itself and hence the selective environment of the other, in a series of adaptations and counter-adaptations. But this means, Sterelny and Griffiths claim, that our evolved psychological mechanisms cannot be modules — that is, innate domain-specific mechanisms that are hard-wired to solve long-enduring adaptive problems — simply because, in the co-evolutionary arms races that characterise social environments, there are "no stable problems … to which natural selection can grind out a solution. The 'adaptive problem' is always being transformed in an arms race" (p. 331).

So, on the first interpretation of Sterelny and Griffiths' position, the grain problem is decisive, since even if distinct cognitive devices subserving distinct types of behaviour were there to be found (something which the secondary offensive calls into doubt), the appeal to the antecedent identification of such devices would fall foul of the independence criterion. On the second interpretation, the antecedent identification of distinct cognitive devices would, in principle, enable us to solve the grain problem. In this instance, the case against evolutionary psychology is completed by the secondary offensive, which aims to provide good reasons to think that such devices are simply not there to be found. But which of these options is the correct interpretation?

In our view, the exegetical waters here are far from clear. However, one way to approach the matter is to take a step back, and to ask whether or not adaptationist explanation, in any form, really is committed to the independence criterion. This move highlights a further tension in Sterelny and Griffiths' argument. In constructing the interactive-social-evolution argument against evolved modularity (described just above), Sterelny and Griffiths depend on the claim that evolutionary psychologists adopt a view of adaptation as being essentially a process of accommodation to a pre-existing environment or set of adaptive problems. Such a view of adaptation is at least strongly suggestive of the independence criterion. In contrast, and on the basis of their account of social evolution, Sterelny and Griffiths promote a notion of adaptation as transformation, concluding that "cognitive adaptation often transforms the environment rather than being an accommodation to it. So there will be real troubles in store for a methodology of discovering the mechanisms of the mind that proceeds by first trying to discover the problems that it must solve, and then testing for the presence of the solutions" (pp. 331-332). But is evolutionary psychology inextricably bound to the allegedly problematic view of adaptation as accommodation? The crucial point here is that while there seems to be no room for the independence criterion in the view of adaptation that Sterelny and Griffiths promote, still they do not suggest that explanations in terms of adaptation as transformation are anything other than solidly Darwinian-adaptationist in character. But this means that even if, in practice, evolutionary psychologists do tend to adopt an independence-criterion-friendly view of adaptation as accommodation, they nevertheless could, in principle, avail themselves of the adaptation-as-transformation view, without being duty bound to give back their copy of the Descent of Man. So even if it is true that adaptation must often be conceived as transformation rather than accommodation, evolutionary psychology, as a Darwinian theoretical endeavour, can continue intact. Our conclusion, then, is that it is the second interpretation of Sterelny and Griffiths' argument that has the intellectual legs here, precisely because it does not make the apparently unreasonable assumption that evolutionary psychology is, in principle, wedded to the independence criterion.

THE TWO-DIMENSIONAL GRAIN PROBLEM

Notwithstanding the fact that Sterelny and Griffiths harbour grave doubts concerning the massively modular view of mind, their position (on our favoured interpretation) seems to be that if those doubts were to be allayed, then the appeal to the antecedent identification of distinct cognitive devices would constitute a general solution to the grain problem. In contrast, our view is that it would not.

The difficulty for the devices-first strategy becomes clear once one realises that there will regularly be a second dimension to the grain problem, one which goes unnoticed by Sterelny and Griffiths. The fact is that it is not just adaptive problems that often form the kind of nested hierarchy that makes choosing a level of descriptive grain arbitrary; so do the parts of organisms that are potential solutions to adaptive problems. For example, Lewontin (1978, p. 161) asks: "Is the leg a unit in evolution, so that the adaptive function of the leg can be inferred? If so, what about a part of the leg, say the foot, or a single toe, or one bone of a t sections we shall try to make this claim stick. (Henceforth we shall default to using 'the grain problem' as shorthand for the two-dimensional grain problem. Where we think there is room for confusion, we shall refer to the one-dimensional grain problem and the two-dimensional grain problem.)

LEVELS OF ANALYSIS

Mainstream evolutionary-psychological theorising is predicated on Marr's (1982) well-known conceptual framework for analysing an information-processing system (Cosmides & Tooby, 1987, 1994a, 1994b). Famously, Marr proposed three distinct levels of analysis. (In the interests of keeping things clear, we note that a Marrian level of analysis is not the same thing as a level of descriptive grain.) As Marr's tripartite framework is central to our account, in what follows, of evolutionary-psychological theorising about social reasoning, we here briefly summarise that framework, as it applies within evolutionary psychology. (For a more detailed consideration of Marr’s levels of analysis and how they feature in evolutionary-psychological theorising generally, see Atkinson & Wheeler, under review.)

The backbone of evolutionary-psychological theorising is what Marr called the computational level of analysis — or, as we prefer, the ecological level (Sterelny, 1990). This is the level at which information-processing capacities are broadly characterised in terms of what a system can do, without any commitment as to the processes or mechanisms that enable the system to do those things. The rationale of this level of analysis is that, in order to understand how a system works — that is, in order to construct detailed theories about the mechanisms and processes that underpin and explain its observable behaviour — we must first understand what problems that system can solve and how it can solve them. This involves specifying cognitive capacities relative to the environment or background in which those capacities are displayed. An evolutionary psychologist working at this level draws upon knowledge of the design principles and practices of natural selection to provide a framework that informs and constrains her theories of the underlying psychological architecture.

At the second, algorithmic level, possible algorithms by which the system under investigation might solve a given information-processing problem are specified, along with descriptions of the type of representation required for the input and output. The third level of analysis, the level of hardware implementation, concerns the physical realisation of those algorithms and representations (e.g., in neural structures and processes).

There is an important difference between ecological-level theories on the one hand, and algorithmic- and implementational-level theories on the other. (We are here going beyond Marr's own account of these levels of analysis, although our intention is to elucidate rather than extend Marr's idea.) Ecological-level theories are concerned with capacities and properties of whole systems; paradigmatically, although, as we shall see, not exclusively, they are concerned with the way in which the behavioural responses and psychological capacities of whole organisms are related to the organism's EEA.[3] Algorithmic- and implementational-level theories, in

---

[3] We say that the ecological level is paradigmatically but not exclusively concerned with whole organisms because (a) the parts of whole systems (e.g., the organs of

contrast, are concerned with parts of systems (their information-processing components and relevant inputs and outputs). This distinction between wholes and their parts is significant because whole systems will have capacities that their parts do not have. (See Bechtel & Richardson, 1993, and Bechtel, 1994a, for detailed discussion of this general point.) Moreover, there is an important sense in which system-level characterisations are able to 'ride free' of their subsystem-level explanations. Within the ecological level, there is no necessary commitment as to the functional architecture of the system; ecological-level theories can be neutral with respect to mechanism. Theories at the algorithmic and implementational levels, on the other hand, are proffered as accounts of, and thus presuppose, ecological-level phenomena. (Or at any rate, to be psychologically relevant, algorithmic- and implementational-level theories should be proffered as accounts of ecological-level phenomena. This was Marr’s 1982, point in arguing for the utility, and indeed necessity, of the ecological level for psychology.)

With this Marrian framework in place, we can understand the overall explanatory enterprise of evolutionary psychology as being the search for answers to two general and distinct kinds of question, which are nevertheless interconnected: (1) Working within the ecological level, the evolutionary psychologist wants to know how psychological capacities are related to adaptive problems. Given a psychological capacity, she will ask what selection pressures in the organism's EEA (if any) might have led to the evolution of that capacity. Given an adaptive problem, she will ask what psychological capacities might be solutions to that problem. (2) Moving between the ecological level on the one hand, and the algorithmic and implementational levels on the other, the evolutionary psychologist wants to know how adaptive problems and psychological capacities are related to features of the brain's information-processing architecture. Given a psychological capacity, she will ask whether there is a feature, or set of related features, of the psychological architecture underpinning that capacity. Given a feature, or set of related features, of the psychological architecture, she will ask what psychological capacities might those features underpin, and for what adaptive problem, or set of related adaptive problems, those features might have evolved to solve.

Notice that within each of these two general kinds of question, there are two main sorts of inference employed by evolutionary psychologists, and by adaptationists generally. These are (a) inferring adaptive problems from phenotypic solutions, and (b) inferring phenotypic solutions from adaptive problems. In the next two sections, we draw on recent work in the evolutionary cognitive neuroscience of social cognition and reasoning, in order to show how these two kinds of inference play their part in enabling

---

organisms), when viewed at a lower level of description or organisation, can themselves be considered as whole systems (systems which will often, in turn, be decomposable into yet smaller subsystems), and (b) on our favoured interpretation of Marr's view, each of the three levels of analysis can be applied at different levels of organisation (Atkinson & Wheeler, under review; Bechtel, 1994b; McClamrock, 1991). Thus, on this view, it is possible for parts of whole organisms (or perhaps even groups of whole organisms) to be the subject of ecological-level analyses.

evolutionary psychology to absorb the grain problem.[4] Note, however, that the picture we are about to paint of two kinds of inference constituting two separate explanatory paths is an idealised one. We paint an idealised picture for ease of exposition, but as we shall elaborate later, the reader should bear in mind that these two inferential paths — from problems to solutions, and from solutions to problems — in fact work together to constitute a unified explanatory project, and that, in practice, any one researcher's explanatory endeavours may sometimes involve both kinds of inference.

INFERRING ADAPTIVE PROBLEMS FROM THEIR SOLUTIONS: THE CASE OF REASONING

What psychological mechanisms underlie our ability to reason and make decisions? The burgeoning search for the information-processing and neural substrates of reasoning and decision-making abilities in humans and other primates is beginning to show fruit. By way of illustration, we here concentrate on the neural substrate of social reasoning. Our aim is to highlight the way in which cognitive neuroscientists have appealed to evolutionary theory in order to help them make sense of the cognitive architecture underpinning social reasoning. That is, in this section we wish to illustrate the approach to evolutionary-psychological explanation that begins with proposals concerning how specific neural structures and processes underpin certain psychological capacities (in this case, those related to social cognition) and ends up by positing an adaptive problem or set of related adaptive problems for which those capacities and their architectural substrates are evolved solutions (see Figure 1). Our approach in what follows is briefly to survey the literature on the neural basis of social cognition, with particular reference to social reasoning, while keeping firmly in mind questions of the following sort: Why is the brain structured and organised this way? What are the various mechanisms for? How did they evolve? These are the sort of questions whose answers require appeal to selection pressures in ancestral environments, that is, to adaptive problems. We end this section by detailing one such attempt to provide a general answer to these sorts of questions in relation to social cognition, namely, the Machiavellian intelligence hypothesis.

---

[4] Sterelny and Griffiths (1999) themselves examine these two sorts of inference and the problems with each (see also Griffiths, 1996, 1997, 2001), but do not marry this examination with their discussion of the (one-dimensional) grain problem. The account we develop here is also to be distinguished from Cosmides and Tooby's writings on various strategies in evolutionary-psychological explanation (e.g., Cosmides & Tooby, 2000; Cosmides et al., 1992; Tooby & Cosmides, 1992, 1995, 2000) — see Atkinson and Wheeler (under review) for details.

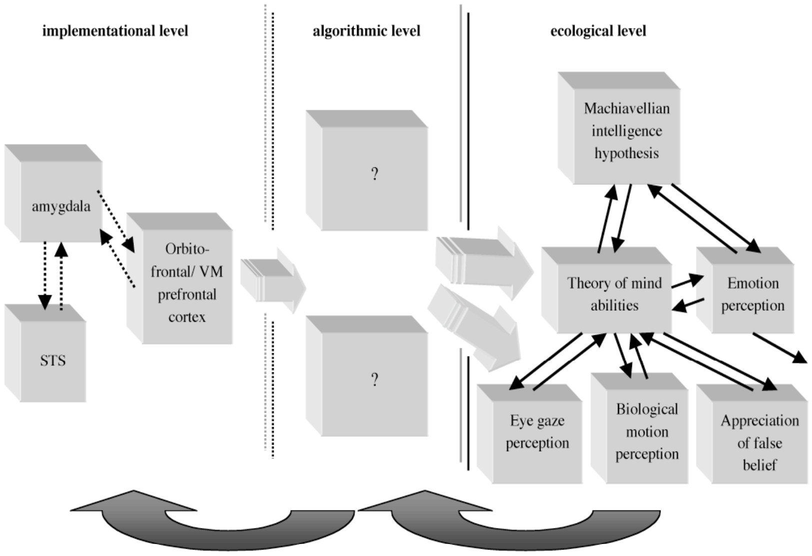

Figure 1 — Levels of analysis from the bottom up: Explanatory paths followed when inferring adaptive problems from phenotypic solutions, illustrated with the case of social reasoning (see text). Arrows represent directions of theorising, except for [illegible] transitions represented by the thick arrows across the vertical dividing lines indicate moves between theorising within Marr's (1982) implementational, algorithmic and ecological levels of analysis. The solid vertical lines between the ecological level and the algorithmic and implementational levels indicates a change from theorising about components of organisms (especially of their brains) to theorising about those whole organisms themselves. The darker right-to-left arrows at the bottom indicate the iterative, continuously evolving nature of theorising in evolutionary cognitive neuroscience (see text).

Recent research has shown that three of the most important areas of the brain for social cognition are the frontal lobes, especially the ventromedial prefrontal cortex; the amygdala; and the superior temporal sulcus. We shall briefly discuss what has so far been revealed about the roles that each of these areas play in social cognition, before going on to summarise the Machiavellian intelligence hypothesis.

The frontal lobes are critically involved in many higher cognitive capacities, such as those that enable us to behave in an organised way, to plan actions, to make decisions, and to decide what to attend to (for reviews see e.g., Damasio, 1994; Duncan, 1986, 1995; Luria, 1966; Miller, 2000; Milner, 1982; Passingham, 1993; Shallice, 1988). Certain memory abilities also involve the frontal lobes: several regions of prefrontal cortex are involved in episodic memory (Henson, Shallice, & Dolan, 1999; Lepage, Ghaffer, Nyberg, & Tulving, 2000), and dorsolateral prefrontal cortex is critically involved in working memory (e.g., Goldman-Rakic, 1990, 1992; Milner, Petrides, & Smith, 1985). Not all higher cognitive capacities are dependent on the frontal lobes, however, for we know that frontal lobe lesions do not produce marked decrements in performance on standard intelligence tests (e.g., Black, 1976; Hebb, 1939; Milner, 1964). But what is most relevant for present purposes is the importance of the frontal lobes in social cognition, and in particular, the subdivision known as the ventromedial prefrontal cortex (VMPFC). (This area includes a significant part of the cortex behind the eyes, that is, part of orbitofrontal cortex or OFC. Hence the two terms are sometimes used interchangeably.)

VMPFC is particularly involved in social reasoning and decision making (for reviews see Adolphs, 1999; Adolphs, Tranel, Bechara, Damasio, & Damasio, 1996; Allison, Puce, & McCarthy, 2000; Barton & Dunbar, 1997; Bechara, Damasio, & Damasio, 2000; Brothers, 1990, 1997; Damasio, 1994; Raleigh et al., 1996; Stone, 2000; Tranel, Bechara, &

Damasio, 2000). For instance, work by Adolphs, Damasio and colleagues (Adolphs, 1999; Adolphs et al., 1996; Bechara et al., 2000; Damasio, 1994) has shown that lesions to VMPFC selectively impair decision-making that depends on associations with emotional experience (e.g., gambles based on hunches) and performance on social contract versions of the Wason selection task. These authors conclude that the VMPFC plays a vital role "in linking perceptual representations of stimuli with representations of their emotional and social significance" (Adolphs, 1999, p. 474; see especially Damasio, 1994).

The ventromedial or orbitofrontal region is also implicated in certain aspects of 'theory of mind', that is, in some of the cluster of abilities involved in explaining and predicting behaviour by attributing mental states. For example, Stone, Baron-Cohen and Knight's (1998) patients, who had bilateral OFC damage, exhibited performance within the normal range on first- and second-order false belief tasks, but were nevertheless significantly impaired in their performance on subtler forms of social reasoning involving mental state attributions, such as recognising a faux pas in a story. Rowe, Bullock, Polkey and Morris' (2001) study suggests that performance on first- and second-order false belief tasks might be compromised by more general but unilateral frontal lesions (their patients' brain damage was limited to either left or right frontal cortex but varied in both extent and exact location). There is also the tentative finding by Stuss, Gallup and Alexander (2001), whose patients had lesions in the region of VMPFC/ OFC, especially in the right hemisphere. These patients were impaired on a task that required them to infer that they were being deceived by the experimental assistant; in contrast, patients with lesions throughout the frontal lobes, but particularly in the right hemisphere, were impaired on a task that required them to infer visual experience in others.

The amygdala has a multifaceted role in social cognition (Adolphs, 1999; Aggleton, 2000; Baron-Cohen et al., 2000; Brothers, 1990; Emery & Amaral, 1999; Kling & Brothers, 1992), as well as in other capacities, such as associative learning (e.g., Bechara et al., 1995; Davis, 1992; La Bar, Le Doux, Spencer, & Phelps, 1995, 1998; Le Doux, 1995; Rolls, 1999; Weiskrantz, 1956). Lesion-based neuropsychological experiments, functional brain-imaging and other neurophysiological measures all converge on the finding that the amygdala plays a key role in the processing of emotional signals, especially those related to fear and in some cases anger, resulting in the triggering of appropriate physiological and behavioural responses (e.g., Adolphs et al., 1999b; Allman & Brothers, 1994; Broks et al., 1998; Calder et al., 1996, 1998; Le Doux, 1995; Morris et al., 1998; Morris et al., 1996; Scott et al., 1997; Tranel, 1997; Young et al., 1995; Young, Hellawell, Van De Wal, & Johnson, 1996). A word of caution here, however: The jury is still out on whether the amygdala's role in humans (and in other primates) is specific to stimuli related to social threat or to threat and danger in general (Adolphs, 1999; Adolphs, Russell, & Tranel, 1999a; Adolphs & Tranel, 1999; Adolphs, Tranel, & Damasio, 1998; Allman & Brothers, 1994; Brothers, 1992; Kling & Brothers, 1992, 1998; Le Doux, 1995). Indeed, some authors propose that the amygdala is not so much involved in dealing with threats, but instead plays a role in dealing with situations of distress (Blair, Morris, Frith, Perrett, & Dolan, 1999) or stimulus ambiguity (Whalen, 1999).

The amygdala, especially the left amygdala, is also implicated as part of a neural circuit underpinning theory of mind abilities (Baron-Cohen, 1995; Baron-Cohen et al., 2000; Fine, Lumsden, & Blair, 2001; Stone, 2000). In particular, research so far has shown that the amygdala is involved in: (a) processing information about biological motion (Bonda, Petrides, Ostry, & Evans, 1996; Brothers, Ring, & Kling, 1990), especially in enabling us to attribute social meaning to moving stimuli (Heberlein et al., 1998); (b) the perception of eye gaze (Kawashima et al., 1999; Young et al., 1995; though see Broks et al., 1998 for a word of caution), allowing us, for example, to judge someone's mental state from eye gaze alone (Baron-Cohen et al., 1999); and (c) the ability to appreciate that people can have false beliefs and the related ability to deploy knowledge about mental states in order to understand jokes and non-literal speech such as sarcasm and metaphor (Fine et al., 2001).

The superior temporal sulcus (STS) has also been implicated as an important structure in theory of mind abilities (Frith & Frith, 1999), again especially via its involvement in the perception of eye gaze and of biological motion (Abell et al., 1999; Emery, 2000; Hietanen & Perrett, 1996; Perrett et al., 1989; Perrett et al., 1991; Perrett et al., 1985), including motion of the head, mouth, hands, and whole body, and even of implied biological motion (see Allison et al., 2000; Carey, Perrett, & Oram, 1997; Jellema & Perrett, 2001, for reviews). Indeed, the STS both sends projections to the amygdala and receives projections from it (Amaral, Price, Pitkanen, & Carmichael, 1992), which goes some way to explaining the shared functions of these structures. Moreover, the amygdala is reciprocally connected to the orbitofrontal/ ventromedial region (Amaral et al., 1992). So the VM prefrontal cortex, amygdala, and STS together form a system for social cognition, including social reasoning.[5]

It is now time to step back from the details of the neurological underpinnings of social cognition and to consider, as some evolutionary psychologists and neuroscientists do, those questions with which we began this section. We can here summarise these questions as follows: How and why did our many specialised cognitive abilities evolve? And in particular, how and why did the specialised circuitry for social reasoning evolve? As indicated in Figure 1, one popular suggestion is the social or Machiavellian intelligence hypothesis (Byrne & Whiten, 1988, 1991, 1992; Dunbar, 1998; Whiten & Byrne, 1997).

The general idea of the Machiavellian intelligence hypothesis is that group living sets up an environment in which each individual is out to get the best or most he can by using strategies of social manipulation, but without causing such disruption that his membership in the group is

---

[5] The overall neuro-computational basis of social cognition is likely to be more complex than this, however, involving these three neural structures in ways not surveyed here, and probably also certain other parts of the brain. See, for example, Damasio's (1994, 1999) ideas about the role of somatosensory cortex in social cognition, and the exciting work on 'mirror neurons' and how they might be involved in certain theory of mind abilities (Gallese, 2000a, 2000b; Gallese & Goldman, 1998; Rizzolatti, Fadiga, Fogassi, & Gallese, 1999).

jeopardised. This hypothesis began with Humphrey's (1976) observation that primates appear to have more intelligence than they need for their everyday tasks of feeding and ranging, and his subsequent suggestion that the social complexity inherent in many primate groups would have been a significant selection pressure acting to increase primate intelligence. This idea has since gained credence from the finding that average social group size and neo-cortex size are positively correlated across species: the bigger the social groupings, the bigger is the neo-cortex relative to the rest of the brain (Aiello & Dunbar, 1993; Dunbar, 1992, 1995; Sawaguchi & Kudo, 1990). When we turn to the question of dedicated psychological mechanisms for social reasoning, the Machiavellian hypothesis predicts that the many specialised cognitive abilities of primates evolved via an arms race in which individuals competed with each other, and formed alliances with some to compete with others, using increasingly sophisticated social strategies. The hypothesis thus predicts psychological mechanisms specialised for pro-social behaviour such as co-operation and altruism, as well as for coercion and deception, and mechanisms for predicting behaviour and for 'reading' the mental states of conspecifics. These psychological capacities have been operationalised as, and their presence in children and nonhuman primates tested by means of tasks probing, for example: self-recognition and self-directed actions, gaze following and joint visual attention, co-operative problem-solving, joint role comprehension, understanding of false beliefs, and several forms of tactical deception and intentional teaching (for reviews and critical discussion see Baron-Cohen, 1995, 1999; Byrne, 1995; Byrne & Whiten, 1988, 1992; Heyes, 1998; Povinelli, 1999; Povinelli & Preuss, 1995; Povinelli & Prince, 1998; Premack & Woodruff, 1978; Tomasello & Call, 1997; Whiten, 1991, 1999; Whiten & Byrne, 1988, 1997).

Our discussion of this work on social reasoning illustrates how theorising in the cognitive and brain sciences involves a sophisticated methodology in which the theorist moves back and forth between the ecological and sub-ecological levels of analysis. In adopting the strategy of inferring problems from solutions, the investigator's explanatory endeavours begin at the ecological level, whether explicitly or not. Working within that level of analysis (refer to the right-hand side of Figure 1), the investigator takes a psychological capacity (e.g., theory of mind abilities) and breaks it down into a set of subcapacities (e.g., the perception of eye gaze and of biological motion, and the appreciation of false belief). She might then continue to decompose each of these subcapacities into yet more specific subcapacities, and so on. At each stage of this decompositional analysis, the postulated subcapacities may be revised, supplemented with additional subcapacities, or replaced by alternatives, in the light of relevant evidence. The next stage in this inferential strategy is to switch to the algorithmic or implementational levels, or both, and ask, for each capacity identified at the ecological level, whether there are in fact any such features of the psychological architecture under which underpin the capacity in question (e.g., the pressures grouped together under the Machiavellian intelligence hypothesis).

INFERRING SOLUTIONS FROM ADAPTIVE PROBLEMS: THE CASE OF REASONING

Evolutionary cognitive psychologists and neuroscientists do not only infer adaptive problems from psychological capacities. They also infer psychological capacities from adaptive problems. For example, in investigating reasoning and decision-making abilities, one might start from a higher-level theory whose essence is that many of the most serious adaptive challenges for primates took place during social interactions with conspecifics. One might start, for instance, with the Machiavellian intelligence hypothesis (refer Figure 2). One might then propose that certain specific pressures in the social environments of primates, including the human EEA, selected for certain psychological mechanisms that enabled our ancestors to make decisions concerning who to form alliances with, who to share precious food and other resources with, who to avoid for fear of being beaten up, who to groom, who to mate with, and so forth.

On the basis of some such account of the likely selection pressures for social decision-making, the evolutionary psychologist would (or at any rate, should) then embark on a two-part search for candidate solutions to those adaptive problems. First, she would remain at the ecological level of analysis, and search for psychological capacities that might have been direct responses to those selection pressures in the organism's ancestors. Having identified a list of specific adaptive problems and candidate solutions to those problems, she would then drop to the algorithmic or implementational levels, or both, and search for underlying architectural features in that organism that subserve or underpin those capacities. This will involve an 'engineering analysis': constructing and testing plausible arguments about how individual phenotypic features function as well-engineered adaptive devices (see e.g., Dawkins, 1986; Lewontin, 1978; Tooby & Cosmides, 1992; Williams, 1966). It should also involve, but often does not, comparative testing of these hypothesised adaptations (Griffiths, 1996, 1997, 2001), that is, analysing the historical development of these phenotypic features across related species.

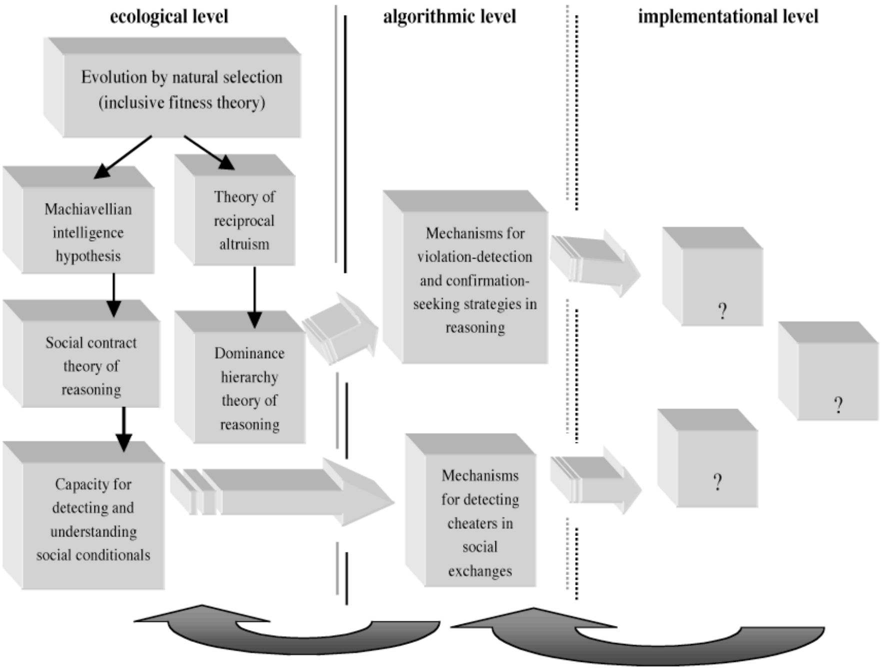

Figure 2 Levels of analysis from the top down: Explanatory paths followed when inferring phenotypic solutions from adaptive problems, illustrated with the case of social reasoning (see text). As with Figure 1, the [illegible] arrows represent directions of theorizing. The [illegible] arrows represent [illegible] the vertical dividing lines indicate transitions between Marr's (1982) three levels of analysis. Thus the sequence depicted in the diagram moves from [illegible] adaptive problems and psychological capacities [illegible] to the psychological mechanisms and processes underpinning those capacities. The darker right-to-left arrows at the bottom indicate the iterative, continuously evolving nature of theorising in evolutionary cognitive neuroscience (see text). (The initial inspiration for Figures 1 and 2 came from Buss, 1999, p. 48.)

There are several good examples of theories of human reasoning and decision-making explicitly motivated by and grounded in considerations of ancestral selection pressures. We shall now briefly introduce two of them — social contract theory and hazard management or precaution theory — in order to illustrate how evolutionary psychologists have engaged in this two-part search for candidate solutions (psychological mechanisms) to adaptive problems (likely selection pressures). (We discuss a third theory in a later section.)

Social contract theory is concerned with deontic reasoning, that is, reasoning about obligations, permissions, and prohibitions. Its particular focus is with reasoning in situations of social exchange — situations in which individuals exchange resources and favours (e.g., Cosmides, 1989; Cosmides & Tooby, 1989, 1992, 1997; Fiddick, 1998; Fiddick, Cosmides, & Tooby, 2000; Gigerenzer & Hug, 1992). The primary proponents of this theory, Cosmides and Tooby, began by drawing upon some ideas and general observations about human social interaction from evolutionary biology, game theory and economics, to derive some conclusions about the selection pressures likely to have shaped social reasoning. One idea they focused upon was that much human social behaviour involves bestowing benefits or inflicting costs on others (helping or hurting them), and that these acts are performed conditionally. For example, a person will generally aid a non-relative only if he has himself received help from that person in the past, or if there is good reason to believe that that person will reciprocate the good deed in the future. This idea is the essence of the theory of reciprocal altruism (Axelrod & Hamilton, 1981; Trivers, 1971), according to which psychological mechanisms supporting altruistic behaviour towards non-relatives can evolve as long as the benefits bestowed on others are reciprocated. So under these lights, human interaction is saturated with social conditionals, that is, with what Cosmides and Tooby describe as "statements or behaviours that express an intention to make one's behaviour contingent upon that of another" (Cosmides & Tooby, 1997, p. 140).

So social contract theory is derived principally from the more general theory of reciprocal altruism. Continuing down the hierarchy of levels of theorising, Cosmides and Tooby derived a specific evolutionary hypothesis from social contract theory, namely, that an important selection pressure in the human EEA was one that selected "for cognitive designs that can detect and understand social conditionals reliably, precisely and economically" (Cosmides & Tooby, 1997, p. 140). But given that this selection pressure is still at a fairly high level of description, it is of more use in informing ecological-level theories about the general kinds of cognitive capacities involved in social reasoning and decision-making than it is in informing theories of the underlying cognitive machinery. Finer-grained selection pressures are required to inform detailed dissection of psychological mechanisms, insofar as having these finer-grained selection pressures to hand allows the evolutionary psychologist to draw up a list of design features that those mechanisms must embody. In our example of the evolutionary psychology of reasoning, it is at this stage in the explanatory enterprise that the concept of social exchange comes to the fore. Cosmides and Tooby reasoned that social exchanges or contracts are an important category of social conditional, and further, that detecting cheaters is an important adaptive problem for those engaging in social exchange. Cheaters in this context are those who break a social contract, that is, those who accept a benefit without paying an appropriate cost, such as someone who receives a favour but does not return it.

So we now have a reasonably specific adaptive problem that was arrived at by decomposing a general or high-level selection pressure into more specific or lower-level pressures. The pressure for forming social contracts and the concomitant pressure for detecting cheaters were derived from the pressure for reliable, precise and economical detection and understanding of social conditionals in general. Famously, Cosmides and Tooby went on to propose that the human cognitive architecture contains mechanisms specialised for detecting cheaters in social exchange, and to test this idea using variations on the Wason selection task (see especially Cosmides, 1989; Cosmides & Tooby, 1989, 1992, 1997; see also Gigerenzer & Hug, 1992). But the important matter for present purposes is that Cosmides and Tooby proposed some specific design features of the psychological architecture underlying social exchange by examining what evolutionary biology and game theory had to say about these selection pressures. They (Cosmides & Tooby, 1992) derived a "partial list" of design features at the algorithmic level of analysis that must be embodied by the psychological mechanisms of a species engaging in social exchange. This list includes algorithms for: assessing costs and benefits, regardless of the goods or services exchanged; recognising other individuals; remembering one's interactions with others; communicating one's values, needs and desires to others; and recognising someone else's values, needs and desires.

One might continue along this explanatory path by decomposing each of these fairly broadly defined algorithms into more specific ones. For example, the ability to remember one's past interactions with others might involve, inter alia, algorithms for remembering whether people were previously co-operators or cheaters, and for remembering and updating records of who owes what to whom. To take another example, person recognition might involve the operation of algorithms for face and voice recognition, and for storing, organising and retrieving semantic and episodic knowledge about people.

Once the evolutionary psychologist has arrived at some such specific hypotheses about the information-processing architecture underlying the psychological capacities in question, these hypotheses must, of course, be tested, by using behavioural experiments, for example (e.g., those employing the Wason selection task). The next major step in the explanatory enterprise, which can also serve to test these algorithmic-level hypotheses, is to search for the neural substrates of these postulated features of the cognitive architecture. Clearly, then, it is in this next step in the explanatory enterprise that a link with the discussion of the previous section emerges. In terms of our idealised account, evolutionary psychologists who infer phenotypic solutions from adaptive problems will eventually be concerned with just those findings with which theorists who adopt the converse inferential strategy begin, namely, data concerning the neural substrate of psychological capacities.

We shall shortly mention some (admittedly initial and tentative) neuropsychological evidence for social contract theory. As that evidence also involves a second and complementary theory of deontic reasoning, we shall first need to introduce that second theory.

Social contract theory explanations of patterns of performance on conditional reasoning tasks are not exclusive, despite implications to that effect in the early literature (e.g., Cosmides, 1989). Cheng and Holyoak (1989) and Manktelow and Over (1990) showed that precaution rules, which lack the cost-benefit structure of a social contract — such as 'If you clear up spilt blood then you must wear rubber gloves' — also facilitate conditional reasoning. Cheng and Holyoak (1989) used such results as ammunition against social contract theory and evolutionary psychology more generally, whereas Manktelow and Over (1990) were only showing that the initial statement of social contract theory was too strong. Subsequently, Cosmides, Tooby, and Fiddick (Cosmides & Tooby, 1997; Fiddick, 1998; Fiddick et al., 2000) ran with Manktelow and Over's view that social contract theory can explain some but not all instances of facilitation on conditional reasoning tasks, and gave it a more detailed evolutionary twist by developing the complementary hazard management or precaution theory.[6]

---

[6] Manktelow and Over (1990), who can be credited with the discovery of precaution rules facilitating reasoning (Cosmides & Tooby, 1992, p. 205), inferred this idea from the simple assumption that people have some basic ability at practical reasoning (rather than a more specific ability at detecting cheaters in social contracts). So hazard management theory did not first arise via the line of theorising we are focusing on in this section, that is, as involving the inference of a cognitive solution from an adaptive problem. Nevertheless, in the hands of Fiddick,

The central idea of hazard management theory is that, since the ability to avoid hazards (predators, poisonous critters, cliffs, etc.) would have been of significant evolutionary advantage to our ancestors, we should expect to find within the human cognitive architecture a system specialised for reasoning about hazards and precautions Moreover, on this view, the hazard management system should be separate from the social contract system, since hazard avoidance and social exchange involve solutions and inferring phenotypic solutions from adaptive problems. Regrettably we cannot. It is early days in the evolutionary cognitive neuroscience of reasoning, and such an integration has not as yet established itself in the literature. We are nevertheless hopeful that it will. In the meantime, what we can do is say what the methodological reflections of last two sections tell us about the prospects for evolutionary psychology in the face of the grain problem.

We have illustrated, with reference to work in the cognitive neuroscience of reasoning, a sophisticated explanatory strategy in which the theorist moves within and between levels of description and levels of analysis. We suggest that once the explanatory endeavours of evolutionary psychologists are seen under these lights, any worry that progress in evolutionary psychology is stymied by the grain problem turns out to be unfounded. Here is why.

The picture that we have painted so far, of two separate explanatory paths is, as we forewarned, an idealised one. At times we gave the impression that different researchers will adopt different approaches: inferring psychological capacities and mechanisms from adaptive problems, or inferring adaptive problems from psychological capacities and mechanisms. But in reality, and as we indicated in our two diagrams, the two approaches are complementary, and many researchers will adopt both approaches, sometimes at the same sitting. Moreover, ideally there is a dynamic and mutually constraining relationship between these two approaches. A theorist who is attempting to infer an adaptive problem from a phenotypic solution will benefit from what her colleagues in psychology, anthropology and evolutionary biology have to say about the likely selection pressures on our ancestors that are relevant to the phenotypic solution in question. And a theorist who is attempting to infer a phenotypic solution from an adaptive problem will benefit from what her colleagues in psychology and neuroscience have to say about the features of our psychological architecture that might be relevant to the adaptive problem in question.

As a result of this mutual guidance and constraint, theories at each level of analysis co-develop, both within a given level of description and especially between levels of description. (Cf. Churchland's (1986) account of the 'co-evolution' of theories at different levels of description.) In other words, theories at a given level of description concerning the adaptive function, the algorithms, inputs and outputs, and the neural hardware relevant to some psychological capacity are informed and constrained by theories about the adaptive function, the algorithms, inputs and outputs, and the neural hardware relevant to that psychological capacity at various other levels of description. Importantly, we suggest, this recurring reciprocal feedback between theories at different levels of analysis and description has the effect of neutralising the grain problem. Recall that the two-dimensional grain problem genuinely threatens the evolutionary-psychological enterprise of explaining adapted traits in terms of adaptive problems. The essence of this threat is that, in many cases, both the choice of grain at which to specify adaptive problems, and the choice of grain at which to specify adapted traits, appear to be arbitrary. But given that, as we have now argued, the very methodology of evolutionary theorising about the mind/brain means that there will often be equally legitimate stories to be told about a given psychological capacity at different levels of description, the pressure to single out a unique level of description at which adaptive problems and phenotypic traits (including cognitive devices) must be described simply dissipates. And once this fact is recognised, the grain problem ceases to be a serious in-principle difficulty.

At this point one might raise the worry that this multilevel explanatory enterprise places no clear restrictions on what counts as a good evolutionary-psychological explanation, such that evolutionary psychologists are free to avail themselves of whatever level of description, or combination of levels of description, will happen to tell a cogent story. But this worry can be met, for implicit in the methodology as we have described it is a means for constraining the choice between sets of levels and thus the boundaries of evolutionary-psychological explanation. In general terms, an overall, multilevel account of a given psychological capacity is constrained by the requirement that the component theories at each of the different levels must ultimately be consistent with, and indeed interlock with, each other. But more specifically, mutual guidance and constraint between theories at different levels, and between the differently directed inferential strategies at work in evolutionary psychology, should (one might think) narrow down the range of possible and acceptable theories, and thus take us closer to nature's real joints. (These points are developed in more detail in Atkinson & Wheeler, under review.)

A DISPUTE DISSOLVED?

In the remainder of this chapter we shall identify some consequences of the foregoing discussion, by investigating one more research programme in which the strategy of inferring architectural solutions from adaptive problems gets played out. We shall then have on the table three theories of deontic reasoning, two of which are complementary and together appear in dispute with the third. However, once we place this supposed disagreement in the context of the kind of multilevel evolutionary-psychological explanation that (we have argued) is resistant to the grain problem, the air of conflict appears to subside.

In contrast to the combination of social contract and hazard management theories, dominance hierarchy theory proposes a single system specialised for all forms of deontic reasoning (Cummins, 1996a, 1996b, 1998, 1999, 2000). Cummins observes that when we reason about deontic rules, we overwhelmingly adopt a violation-detection strategy, whereas when we reason about indicative rules (i.e., descriptions of the world, about what is true or false), we adopt a confirmation-seeking strategy. She argues that this difference in reasoning strategies is fundamental, and that we should expect it to be reflected in the cognitive architecture. We adopt a violation-detection strategy and are thereby proficient deontic reasoners, Cummins argues, thanks to a domain-specific system that evolved as the result of selection pressures related to dominance hierarchies (though see Chater & Oaksford, 1996 for a counter-argument).

Dominance hierarchies are characterised by competition and co-operation amongst conspecifics for limited resources (food, mates, etc). An animal's position within a dominance hierarchy is highly correlated with access to those resources: the higher the rank, the greater the share and control over the available resources. Moreover, rank crucially depends on skilful social interaction and manipulation, especially the ability to form and maintain alliances. Cummins' idea is that a significant component of this set of skills is adeptness in social reasoning, especially of the deontic variety. Reasoning effectively about what oneself and others are obliged to do is central to success in forming and maintaining alliances, and in fulfilling other social contracts, especially when those obligations are reciprocal. Moving up in the world depends on trading favours. And reasoning effectively about what oneself and others are permitted and forbidden to do allows higher-ranking individuals to detect lower-ranking transgressors (so they can be warned or punished) and lower-ranking individuals to avoid transgressions, or to engage in deceptive transgression (thus avoiding threatening encounters and punishment). So moving up in the world also depends on following the law, and on breaking the law while not being seen.

Cummins (1996a, 1996b) cites a variety of evidence in support of her claim that the human cognitive architecture includes an evolved system specialised for deontic reasoning. Numerous studies of reasoning in non-human primates show a 'social content effect', that is, these animals make various sorts of inferences much more readily in social tasks, such as kin and dominance rank discriminations (e.g., Cheney & Seyfarth, 1990; Dasser, 1985), than they do in non-social, object-oriented tasks (e.g., Gillian, 1981; McGonigle & Chalmers, 1977). Developmental studies show that (a) children are considerably more accomplished at deontic reasoning than they are at indicative reasoning, (b) that, like adults, they tend to adopt violation-seeking strategies for deontic rules and confirmation-seeking strategies for indicative rules, and (c) that this differential ability is established early in life, around 3 to 4 years of age (e.g., Cummins, 1996c; Harris & Nunez, 1996; Nunez & Harris, 1998). And there is some cross-cultural evidence for the indicative-deontic distinction, including indirect evidence from members of preliterate societies (see Cummins, 1996b for a review).

Cummins also cites neuropsychological studies in humans and monkeys that show evidence of a selective deficit of deontic reasoning. As we noted earlier, lesions to frontal cortex, especially VM prefrontal cortices, can result in severe impairments in social behaviour and reasoning, such as reduced inhibitions, an inability to organise and plan future actions, impairments in making decisions that depend on associations with emotional experience (e.g., gambles based on hunches), and a lack of concern for oneself and others (see Adolphs, 1999; Adolphs et al., 1996; Bechara et al., 2000; Damasio, 1994 for reviews). Results of this kind lead Cummins to claim that damage to prefrontal cortices "impacts most severely on the capacity to respond effectively to the social rules that underlie the dominance hierarchy" (Cummins, 1996b, p. 176).

In Cummins' work, then, we again see an example of an evolutionary psychologist using the inferential strategy of reasoning from adaptive problems to psychological solutions. At the ecological level of analysis, she began by deriving dominance hierarchy theory from the general claim that the social complexity inherent in many primate groups would have been a significant selection pressure acting to increase primate intelligence (roughly, the Machiavellian intelligence hypothesis). That theory makes certain claims about adaptive problems related to dominance hierarchies, from which Cummins inferred a psychological capacity that plausibly plays a significant role in solving those adaptive problems, namely, adopting a violation-detection strategy in deontic reasoning. She then dropped to the sub-ecological levels of analysis to propose a mind/brain mechanism that underpins that capacity, namely, a deontic reasoning system, and garnered a variety of converging evidence for such a system (at least a significant part of which is supposed be located within the frontal lobes).[7]

So how many distinct psychological mechanisms or modules underpin deontic reasoning? As we have seen, Cummins argues that a single mechanism underpins deontic reasoning, whereas Fiddick, Cosmides and Tooby argue for two separate systems, one for social contracts, the other for hazard management. Cummins' argument is based on the claim that the relevant selection pressures confronting our ancestors selected for the ability to reason about permissions, obligations and prohibitions within dominance hierarchies. Hence, on her view, all deontic reasoning, including that about hazards or precautions, is restricted to the social realm.[8] The arguments of Fiddick, Cosmides and Tooby, in contrast, are based on the claim that there were two separate sets of pressures, both selecting for deontic reasoning abilities: one for the ability to reason about precautions (allowing our ancestors to avoid hazards), the other for the ability to reason about social

---

[7] Here we see that Cummins has moved more-or-less straight from the ecological level of analysis to the implementational level, thus skipping explicit theorising at the algorithmic level. Such skipping of algorithmic-level theorising is quite common in these early days of cognitive neuroscience. This is no more so than in the cognitive neuroscience of reasoning, where there is a marked paucity of formal models and simulations of domain-specific reasoning abilities such as those that we have been discussing. An exception, of course, is the large body of work on formalising and modelling higher-level (and sometimes more domain-general) theories of reasoning, such as the Prisoner's dilemma (see e.g., Axelrod, 1984; Axelrod & Hamilton, 1981; Emshoff, 1970; Hayashi, 1995; Messick & Liebrand, 1993; Nauta & Hoeksta, 1995; Watanabe & Yamagishi, 1999). The ecological rationality programme (Gigerenzer, Todd, & the ABC Research Group, 1999) might also be regarded as a notable exception. This programme proposes a variety of "simple heuristics" ("fast and frugal" procedures that exploit the structure of environments to produce ecologically useful inferences and decisions) as the algorithmic underpinnings of many of our reasoning and decision-making abilities. We should nevertheless note that ecological-level theorising for those pursuing the ecological rationality programme does not focus solely, or even primarily, on adaptive problems in ancesteral environments. On their view, ecological rationality has a wider purview than evolutionary psychology (the latter being grounded in the former), encompassing decision making in modern as well as ancesteral environments (see Over, 2000a, 2000b; Todd, Fiddick & Kraus, 2000, for discussion).

[8] Others have also argued that deontic reasoning is inherently social, though not in terms of dominance hierarchy theory. For example, Manktelow and Over (1991) argue that how one represents a deontic rule, and thus how one reasons about that rule, depends on the social role one adopts with respect to it, viz. whether one represents the utilities associated with the agent who lays down the rule or with the actor whose behaviour is its target.

contracts (allowing our ancestors to trade favours and goods, and to detect cheaters). With respect to deontic reasoning, then, the social realm for Fiddick, Cosmides and Tooby is more restricted than it is for Cummins; it is reasoning about social contracts that is crucial to Fiddick, Cosmides and Tooby's view, not reasoning within dominance hierarchies (which might subsume some or all reasoning about social contracts), or social reasoning in general (which would subsume reasoning about social contracts).

Certainly the neuropsychological data do not yet allow us to decide between dominance hierarchy theory and social contract plus hazard management theories. The neuropsychological study of reasoning is still in its infancy; there is as yet insufficient evidence to make any firm conclusions about how specific types of social behaviour and reasoning map on to specific brain regions. For instance, it is an open question as to whether selective impairment on social contract versions of the Wason selection task after VM prefrontal cortex lesions (Adolphs et al., 1996) result because VMPFC forms a crucial part of a system specialised for reasoning about social contracts or for reasoning about dominance hierarchies (which would include reasoning about social contracts). That being said, there is at least one study that provides more direct support for Cummins' theory over social contract theory, which turns on the difference between the two theories with respect to social rank: dominance hierarchy theory predicts social reasoning to be affected by rank whereas social contract theory does not make this prediction. Work by Raleigh and colleagues (Raleigh et al., 1996) has shown that the density of serotonin receptors in the orbitofrontal cortex of monkeys varies according to rank, and that social behaviour and rank are affected by manipulation of serotonin neurotransmission. (See Cummins, 1996a, 1996b, 1998, 1999, 2000, for further evidence and arguments for dominance hierarchy theory as against social contract theory. And see Fiddick, 1998 for a counter-argument, which centres on the claim that, pace Cummins, all deontic reasoning — specifically, reasoning about precautions — is not intrinsically social reasoning.)

While the empirical adequacy of these theories is undoubtedly crucial, the main point that we wish to highlight here is this: the interdisciplinary, multilevel explanatory enterprise that we have outlined in the wake of the grain problem, brings to the fore the possibility that the positions of Cummins on the one hand, and of Cosmides, Tooby, and Fiddick on the other, may not be in competition after all. For it is plausible that these two sets of theories are pitched at different levels of description, and are thus potentially compatible. Specifically, Cummins' dominance hierarchy theory is pitched at a higher level than social contract and hazard management theories, proposing higher-level selection pressures and corresponding psychological mechanisms. So consider, for example, how one might get from Cummins' position to that of Fiddick, Cosmides and Tooby. One might first seek to decompose the higher-level selection pressure for deontic reasoning within dominance hierarchies into more specific selection pressures at a lower level of description. One might also decompose the ability to adopt a violation-detection strategy in deontic reasoning into more specific or lower-level psychological capacities. The second step in this continuation of Cummins' work would then be to identify the more specific psychological mechanisms at lower levels of description that together compose the larger deontic reasoning system, and whose joint operation underpins the adoption of a violation-detection strategy in deontic reasoning tasks. Interestingly, Cummins herself is open to this move, at times implying the possibility that a single deontic reasoning system might itself be decomposed into subsystems (vide, for example, Cummins, 1996b, p. 177). Once all this has been done, it is at least conceivable that the subsystems of the larger deontic reasoning system that one ends up with are something like the subsystems proposed by Fiddick, Cosmides and Tooby. Of course, arriving at that position would first require one to revise Cummins' initial claim that the selection pressures for deontic reasoning are specific to dominance hierarchies (i.e., the claim that all deontic reasoning is inherently social). Indeed, the main unresolved issue blocking a reconciliation of these two positions is whether reasoning about non-social precautions engages a deontic reasoning mechanism (as Cosmides, Tooby, and Fiddick claim), or not (as Cummins claims). The resolution of this particular disagreement is an empirical matter. Be that as it may, our point is that the multilevel explanatory enterprise allows us to see how a reconciliation might proceed.[9]

CONCLUSIONS

Our take-home message is this. The grain problem is even more serious than Sterelny and Griffiths' account suggests. Nevertheless, that problem is not a silver bullet for evolutionary psychology. So as far as this particular threat is concerned, the explanatory credentials of a truly interdisciplinary, scientific, and evolutionary approach to the human mind remain in good shape.

---

[9] Although we do not pursue the matter in detail here, the foregoing discussion provides support for a more general conclusion that we have argued for in detail elsewhere (Atkinson and Wheeler, under review; Wheeler and Atkinson, 2001). In many cases at least, there will be no stage at which one could provide a definitive answer to the question of whether a psychological capacity is domain-specific or domain-general, because that would require one to fix a unique level of description, from which judgements about domain specificity or domain generality could be made, and the regular failure to fix such a level of description is just what is suggested by the theoretical endeavours that we have been illustrating. If this is right, then regardless of whether the two positions on deontic reasoning highlighted in the main text are compatible, it may well be that at no point in the continuing evolutionary-psychological enterprise could one provide a definitive answer to the question of whether deontic reasoning is domain-specific in the sense suggested by Cummins, or whether it is instead a function of more domain-general abilities. Similarly, at no point could one say definitively, as Cosmides, Tooby, and Fiddick are wont to do, that social contract reasoning and precaution reasoning constitute two distinct domain-specific abilities rather than two facets of a higher-level, more domain-general ability.

REFERENCES

Abell, F., Krams, M., Ashburner, J., Passingham, R., Friston, K., Frackowiak, R., Happe, F., Frith, C., & Frith, U. (1999). The neuroanatomy of autism: a voxel-based whole brain analysis of structural scans. Neuroreport, 10, 1647-1651.

Adolphs, R. (1999). Social cognition and the human brain. Trends in Cognitive Sciences, 3, 469-479.

Adolphs, R., Russell, J. A., & Tranel, D. (1999a). A role for the human amygdala in recognizing emotional arousal from unpleasant stimuli. Psychological Science, 10, 167-171.

Adolphs, R., & Tranel, D. (1999). Preferences for visual stimuli following amygdala damage. Journal of Cognitive Neuroscience, 11, 610-616.

Adolphs, R., Tranel, D., Bechara, A., Damasio, H., & Damasio, A. R. (1996). Neuropsychological approaches to reasoning and decision-making. In A. R. Damasio, H. Damasio, & Y. Christen (Eds.), Neurobiology of decision-making (pp. 157-179). Berlin: Springer-Verlag.

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. Nature, 393, 470-474.

Adolphs, R., Tranel, D., Hamann, S., Young, A. W., Calder, A. J., Phelps, E. A., Anderson, A., Lee, G. P., & Damasio, A. R. (1999b). Recognition of facial emotion in nine individuals with bilateral amygdala damage. Neuropsychologia, 37, 1111-1117.

Aggleton, J. P. (Ed.). (2000). The amygdala: A functional analysis. Oxford: Oxford University Press.

Aiello, L. C., & Dunbar, R. I. M. (1993). Neocortex size, group size and the evolution of language. Behavioral and Brain Sciences, 34, 184-193.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. Trends in Cognitive Sciences, 4, 267-278.

Allman, J., & Brothers, L. (1994). Faces, fear and the amygdala. Nature, 372, 613-614.

Amaral, D. G., Price, J. L., Pitkanen, A., & Carmichael, S. T. (1992). Anatomical organisation of the primate amygdaloid complex. In J. P. Aggleton (Ed.), The amygdala: Neurobiological aspects of emotion, memory and mental dysfunction (pp. 1-66). New York: Wiley-Liss.

Atkinson, A. P., & Wheeler, M. (under review). The grain of domains: The evolutionary-psychological case against domain-general cognition.

Axelrod, R. (1984). The evolution of cooperation. New York: Basic Books.

Axelrod, R., & Hamilton, W. (1981). The evolution of cooperation. Science, 211, 1390-1396.

Baizer, J. S., Ungerleider, L. G., & Desimone, R. (1991). Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. Journal of Neuroscience, 11, 168-190.

Baron-Cohen, S. (1995). Mindblindness: An essay on autism and theory of mind. Cambridge, MA: MIT Press/ Bradford Books.

Baron-Cohen, S. (1999). The evolution of a theory of mind. In M. C. Corballis & S. E. G. Lea (Eds.), The descent of mind: Psychological perspectives on hominid evolution (pp. 261-277). Oxford: Oxford University Press.

Baron-Cohen, S., Ring, H., Wheelwright, S., Bullmore, E., Brammer, M., Simmons, A., & Williams, S. (1999). Social intelligence in the normal and autistic brain: an fMRI study. European Journal of Neuroscience, 11, 1891-1898.

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. R. (2000). The amygdala theory of autism. Neuroscience and Biobehavioral Reviews, 24, 355-364.

Barton, R. A., & Dunbar, R. I. M. (1997). Evolution of the social brain. In A. Whiten & R. W. Byrne (Eds.), Machiavellian intelligence II: Extensions and evaluations (pp. 240-263). Cambridge: Cambridge University Press.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. Cerebral Cortex, 10, 295-307.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Roackland, C., & Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. Science, 269, 1115-1118.

Bechtel, W. (1994a). Biological and social constraints on cognitive processes: The need for dynamical interactions between levels of inquiry. In M. Matthen & R. X. Ware (Eds.), Biology and society: Reflections on methodology. (Canadian Journal of Philosophy Supplementary Vol. 20) (pp. 133-164). Calgary: University of Calgary Press.

Bechtel, W. (1994b). Levels of description and explanation in cognitive science. Minds and Machines, 4, 1-25.

Bechtel, W., & Richardson, R. C. (1993). Discovering complexity: Decomposition and localization as strategies in scientific research. Princeton, NJ: Princeton University Press.

Black, F. W. (1976). Cognitive deficits in patients with unilateral war-related frontal lobe lesions. Journal of Clinical Psychology, 32, 366-372.

Blair, R. J., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. Brain, 122, 883-893.

Bloom, P. (1999). The evolution of certain novel human capacities. In M. C. Corballis & S. E. G. Lea (Eds.), The descent of mind: Psychological perspectives on hominid evolution . Oxford: Oxford University Press.

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. Journal of Neuroscience, 16, 3737-3744.

Broks, P., Young, A. W., Maratos, E. J., Coffey, P. J., Calder, A. J., Isaac, C. L., Mayes, A. R., Hodges, J. R., Montaldi, D., Cezayirli, E., Roberts, N., & Hadley, D. (1998). Face processing impairments after encephalitis: amygdala damage and recognition of fear. Neuropsychologia, 36, 59-70.

Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. Concepts in Neuroscience, 1, 27-51.

Brothers, L. (1992). Perception of social acts in primates: Cognition and neurobiology. Seminars in the Neurosciences, 4, 409-414.

Brothers, L. (1997). Friday's footprint. New York: Oxford University Press.

Brothers, L., Ring, B., & Kling, A. (1990). Response of neurons in the macaque amygdala to complex social stimuli. Behavioural Brain Research, 41, 199-213.

Buss, D. M. (1999). Evolutionary psychology: The new science of the mind. Boston: Allyn & Bacon.

Byrne, R. M. (1999). Human cognitive evolution. In M. C. Corballis & S. E. G. Lea (Eds.), The descent of mind: Psychological perspectives on hominid evolution (pp. 71-87). Oxford: Oxford University Press.

Byrne, R. W. (1995). The thinking ape: Evolutionary origins of intelligence. Oxford: Oxford University Press.

Byrne, R. W., & Whiten, A. (1988). Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans. Oxford: Clarendon Press.

Byrne, R. W., & Whiten, A. (1991). Computation and mindreading in primate tactical deception. In A. Whiten (Ed.), Natural theories of mind: Evolution, development and simulation of everyday mindreading (pp. 127-141). Oxford: Basil Blackwell.

Byrne, R. W., & Whiten, A. (1992). Cognitive evolution in primates: Evidence from tactical deception. Man, 27, 609-627.

Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., & Etcoff, N. L. (1996). Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. Cognitive Neuropsychology, 13, 699-745.

Carey, D. P., Perrett, D. I., & Oram, M. W. (1997). Recognizing, understanding and reproducing actions (in sect. 16: Action and cognition, M. Jeannerod, Ed.). In F. Boller & J. Grafman (Eds.), Handbook of neuropsychology (Vol. 11). Amsterdam: Elsevier.

Chater, N., & Oaksford, M. (1996). Deontic reasoning, modules and innateness: A second look. Mind and Language, 11, 191-202.

Cheney, D. L., & Seyfarth, R. M. (1990). How monkeys see the world. Chicago: Chicago University Press.

Cheng, P. W., & Holyoak, K. J. (1989). On the natural selection of reasoning theories. Cognition, 33, 284-313.

Churchland, P. S. (1986). Neurophilosophy: Toward a unified science of the mind/brain. Cambridge, MA: MIT Press/ Bradford Books.

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 (pp. 227-306). Cambridge, MA: MIT Press.

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

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), The adapted mind: Evolutionary psychology and the generation of culture (pp. 163-228). New York: Oxford University Press.

Cosmides, L., & Tooby, J. (1994a). Beyond intuition and instinct blindness: toward an evolutionarily rigorous cognitive science. Cognition, 50, 41-77.

Cosmides, L., & Tooby, J. (1994b). Origins of domain specificity: The evolution of functional organization. In L. A. Hirschfeld & S. A. Gelman (Eds.), Mapping the mind: Domain specificity in cognition and culture (pp. 85-116). Cambridge & New York: Cambridge University Press.

Cosmides, L., & Tooby, J. (1997). Dissecting the computational architecture of social inference mechanisms. Ciba Foundation Symposium, 208, 132-156.

Cosmides, L., & Tooby, J. (2000). Evolution: Introduction. In M. S. Gazzaniga (Ed.), The new cognitive neurosciences. 2nd ed. Cambridge, MA: MIT Press/ Bradford Books.

Cosmides, L., Tooby, J., & Barkow, J. H. (1992). Introduction: Evolutionary psychology and conceptual integration. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), The adapted mind: Evolutionary psychology and the generation of culture (pp. 9-13). New York: Oxford University Press.

Cummins, D. D. (1996a). Dominance hierarchies and the evolution of human reasoning. Minds and Machines, 6, 463-480.

Cummins, D. D. (1996b). Evidence for the innateness of deontic reasoning. Mind and Language, 11, 160-190.

Cummins, D. D. (1996c). Evidence of deontic reasoning in 3- and 4-year-old children. Memory and Cognition, 24, 823-829.

Cummins, D. D. (1998). Social norms and other minds: The evolutionary roots of higher cognition. In D. D. Cummins & C. Allen (Eds.), The evolution of mind (pp. 30-50). New York: Oxford University Press.

Cummins, D. D. (1999). Cheater detection is modified by social rank: The impact of dominance on the evolution of cognitive functions. Evolution and Human Behavior, 20, 229-248.

Cummins, D. D. (2000). How the social environment shaped the evolution of mind. Synthese, 122, 3-28.

Cummins, D. D., & Cummins, R. (1999). Biological preparedness and evolutionary explanation. Cognition, 73, B37-B53.

Damasio, A. R. (1994). Descartes' error: Emotion, reason and the human brain. New York and London: Putnam/ Picador.

Damasio, A. (1999). The feeling of what happens: Body, emotion and the making of consciousness. London: William Heinemann.

Dasser, V. (1985). Cognitive complexity in primate social relationships. In R. A. Hinde, A. N. Perret-Clemont, & J. Stevenson-Hinde (Eds.), Social relationships and cognitive development. Oxford: Clarendon Press.

Davis, M. (1992). The role of the amygdala in conditioned fear. In J. P. Aggleton (Ed.), The amygdala: Neurological aspects of emotion, memory and mental dysfunction (pp. 255-306). NeD)

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. Journal of Human Evolution, 20, 469-493.

Dunbar, R. I. M. (1995). Neocortex size and group size in primates: A test of the hypothesis. Journal of Human Evolution, 28, 287-296.

Dunbar, R. I. M. (1998). The social brain hypothesis. Evolutionary Anthropology, 6, 178-190.

Duncan, J. (1986). Disorganisation of behaviour after frontal lobe damage. Cognitive Neuropsychology, 3, 271-290.

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. S. Gazzaniga (Ed.), The cognitive neurosciences (pp. 721-733). Cambridge, MA: MIT Press/ Bradford Books.

Ellis, A. W., & Young, A. W. (1996). Human cognitive neuropsychology: A textbook with readings (2nd ed.). Hove, Sussex: Psychology Press.

Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. Neuroscience and Biobehavioral Reviews, 24, 581-604.

Emery, N. J., & Amaral, D. G. (1999). The role of the primate amygdala in social cognition. In R. D. Lane & L. Nadel (Eds.), Cognitive neuroscience of emotion (pp. 156-191). New York: Oxford University Press.

Emshoff, J. R. (1970). A computer simulation model of the Prisoner's Dilemma. Behavioral Science, 15, 304-317.

Fiddick, L. (1998). The deal and the danger: An evolutionary analysis of deontic reasoning. Unpublished Ph.D., University of California, Santa Barbara.

Fiddick, L., Cosmides, L., & Tooby, J. (1995). Are there really separate reasoning mechanisms for social contracts and precautions? Paper presented at the Seventh Annual Meeting of the Human Behavior and Evolution Society, University of California, Santa Barbara.

Fiddick, L., Cosmides, L., & Tooby, J. (2000). No interpretation without representation: the role of domain-specific representations and inferences in the Wason selection task. Cognition, 77, 1-79.

Fiddick, L., Cosmides, L., & Tooby, J. (under review). Does the mind distinguish between social contracts and precautions? Dissociating cognitive adaptations through inference priming.

Fine, C., Lumsden, J., & Blair, R. J. R. (2001). Dissociation between 'theory of mind' and executive functions in a patient with early left amygdala damage. Brain, 124, 287-298.

Frith, C. D., & Frith, U. (1999). Interacting minds — A biological basis. Science, 286, 1692-1695.

Gallese, V. (2000a). The acting subject: towards the neural basis of social cognition. In T. Metzinger (Ed.), Neural correlates of consciousness: Empirical and conceptual questions (pp. 325-333). Cambridge, MA: MIT Press.

Gallese, V. (2000b). The inner sense of action: agency and motor representation. Journal of Consciousness Studies, 7, 23-40.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. Trends in Cognitive Sciences, 2, 493-500.

Gigerenzer, G. (1997). The modularity of social cognition. In A. Whiten & R. W. Byrne (Eds.), Machiavellian intelligence II: Extensions and evaluations (pp. 264-288). Cambridge: Cambridge University Press.

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

Gigerenzer, G., Todd, P. M., & Group, t. A. R. (1999). Simple heuristics that make us smart. New York: Oxford University Press.

Gillian, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. Journal of Experimental Psychology: Animal Behavioral Processes, 7, 150-164.

Goldman-Rakic, P. S. (1990). Cellular and circuit basis of working memory in prefrontal cortex of nonhuman primates. Progress in Brain Research, 85, 325-336.

Goldman-Rakic, P. S. (1992). Working memory and the mind. Scientific American, 267, 111-117.

Goodale, M. A. (1997). Visual routes to perception and action in the cerebral cortex. In M. Jeannerod (Ed.), Handbook of neuropsychology (Vol. 11, pp. 91-109). Amsterdam: Elsevier.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. Trends in Neurosciences, 15, 20-25.

Griffiths, P. E. (1996). The historical turn in the study of adaptation. British Journal for the Philosophy of Science, 47, 511-532.

Griffiths, P. E. (1997). What emotions really are: The problem of psychological categories. Chicago: University of Chicago Press.

Griffiths, P. E. (2001). From adaptive heuristic to phylogenetic perspective: Some lessons from the evolutionary psychology of emotion. In H. R. Holcomb (Ed.), Conceptual challenges in evolutionary psychology: Innovative research strategies. Dordrecht: Kluwer.

Harris, P., & Nunez, M. (1996). Understanding of permission rules by pre-school children. Child Development, 67, 1572-1591.

Hayashi, N. (1995). Emergence of cooperation in one-shot prisoner's dilemmas and the role of trust. Japanese Journal of Psychology, 66, 184-190.

Hebb, D. O. (1939). Intelligence in man after large removals of cerebral tissue: report of four left frontal lobe cases. Journal of General Psychology, 21, 73-87.

Heberlein, A. S., Adolphs, R., Tranel, D., Kemmerer, D., Anderson, S., & Damasio, A. R. (1998). Impaired attribution of social meanings to abstract dynamic visual patterns following damage to the amygdala. Society of Neuroscience Abstracts, 24, 1176.

Henson, R. N. A., Shallice, T., & Dolan, R. J. (1999). Right prefrontal cortex and episodic memory retrieval: a functional MRI test of the monitoring process. Brain, 122, 1367-1381.

Heyes, C. M. (1998). Theory of mind in nonhuman primates. Behavioral and Brain Sciences, 21, 101-148.

Hietanen, J. K., & Perrett, D. I. (1996). Motion sensitive cells in the macaque superior temporal polysensory area: response discrimination between self-generated and externally generated pattern motion. Behavioural Brain Research, 76, 155-167.

Humphrey, N. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), Growing points in ethology (pp. 303-317). Cambridge: Cambridge University Press.

Jellema, T., & Perrett, D. I. (2001). Coding of visible and hidden actions. In W. Prinz & B. Hommel (Eds.), Attention and performance XIX. Oxford: Oxford University Press.

Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., Fukuda, H., Kojima, S., & Nakamura, K. (1999). The human amygdala plays an important role in gaze monitoring. A PET study. Brain, 122, 779-783.

Kling, A. S., & Brothers, L. A. (1992). The amygdala and social behavior. In J. P. Aggleton (Ed.), The amygdala: Neurological aspects of emotion, memory and mental dysfunction (pp. 353-377). New York: Wiley-Liss.

La Bar, K. S., Le Doux, J. E., Spencer, D. D., & Phelps, E. A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. Journal of Neuroscience, 15, 6846-6855.

Le Doux, J. E. (1995). Emotion: Clues from the brain. Annual Review of Psychology, 46, 209-235.

Le Doux, J. (1998). The emotional brain: The mysterious underpinnings of emotional life. New York: Simon & Schuster.

Lepage, M., Ghaffer, O., Nyberg, L., & Tulving, E. (2000). Prefrontal cortex and episodic memory retrieval mode. Proceedings of the National Academy of Sciences, 97, 506-511.

Leslie, A. M. (1994). ToMM, ToBY, and agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), Mapping the mind: Domain specificity in cognition and culture (pp.119-148). Cambridge: Cambridge University Press.

Lewontin, R. C. (1978). Adaptation. Scientific American, 239, 156-169.

Luria, A. R. (1966). Higher cortical functions in man. New York: Basic Books.

Manktelow, K. I., & Over, D. E. (1990). Deontic thought and the selection task. In K. J. Gilhooly, M. T. G. Keane, R. H. Logie, & G. Erdos (Eds.), Lines of thinking: Reflections on the psychology of thinking (Vol. 1, pp. 153-164). London: Wiley.

Manktelow, K. I., & Over, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. Cognition, 39, 85-105.

Marr, D. (1982). Vision. New York: W. H. Freeman.

Maunsell, J. H. R., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. Annual Review of Neuroscience, 10, 363-401.

McClamrock, R. (1991). Marr's three levels: A re-evaluation. Minds and Machines, 1, 185-196.

McGonigle, B. O., & Chalmers, M. (1977). Are monkeys logical? Nature, 267, 694-696.

Messick, D. M., & Liebrand, W. B. (1993). Computer simulations of the relation between individual heuristics and global cooperation in Prisoner's Dilemmas. Social Science Computer Review, 11, 301-312.

Miller, E. K. (2000). The prefrontal cortex and cognitive control. Nature Reviews: Neuroscience, 1, 59-65.

Milner, A. D., & Goodale, M. A. (1995). The visual brain in action. Oxford: Oxford University Press.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren & K. Akert (Eds.), The frontal granular cortex and behavior (pp. 313-334). New York: McGraw-Hill.

Milner, B. (1982). Some cognitive effects of frontal lobe lesions in man. In D. E. Broadbent & L. Weiskrantz (Eds.), The neuropsychology of cognitive function (pp. 211-226). London: The Royal Society.

Milner, B., Petrides, M., & Smith, M. L. (1985). Frontal lobes and the temporal organization of memory. Human Neurobiology, 4, 137-142.

Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. Brain, 121, 47-57.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. Nature, 383, 812-815.

Nauta, D., & Hoeksta, J. (1995). Effective choice in the single-shot Prisoner's Dilemma tournament. Theory and Decision, 39, 1-30.

Nunez, M., & Harris, P. L. (1998). Psychological and deontic concepts: Separate domains or intimate connection? Mind and Language, 13, 153-170.

Over, D. E. (2000a). Ecological rationality and its heuristics. Thinking and Reasoning, 6, 182-192.

Over, D. E. (2000b). Ecological issues: A reply to Todd, Fiddick, & Kraus. Thinking and Reasoning, 6, 385-388.

Passingham, R. E. (1993). The frontal lobes and voluntary action. Oxford: Oxford University Press.

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K., & Ortega, J. E. (1989). Frameworks of analysis for the neural representation of animate objects and actions. Journal of Experimental Biology, 146, 87-113.

Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J., & Thomas, S. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. Experimental Brain Research, 86, 159-173.

Perrett, D. I., Smith, P. A., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., Broennimann, R., Milner, A. D., & Jeeves, M. A. (1985). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report. Behavioural Brain Research, 16, 153-170.

Pinker, S. (1994). The language instinct. London: Allen Lane/ Penguin.

Pinker, S. (1997). How the mind works. London: Allen Lane/ Penguin.

Povinelli, D. J. (1999). Social understanding in chimpanzees: New evidence from a longitudinal approach. In P. D. Zelazo, J. W. Astington, et al. (Eds.), Developing theories of intention: Social understanding and self-control (pp. 195-225). Mahwah, NJ: Lawrence Erlbaum.

Povinelli, D. J., & Preuss, T. M. (1995). Theory of mind: Evolutionary history of a cognitive specialization. Trends in Neurosciences, 18, 418-424.

Povinelli, D. J., & Prince, C. G. (1998). When self met other. In M. D. Ferrari & R. J. Sternberg (Eds.), Self-awareness: Its nature and development (pp. 37-107). New York: Guilford Press.

Premack, D., & Woodruff, G. (1978). Does a chimpanzee have a theory of mind? Behavioral and Brain Sciences, 1, 515-526.

Raleigh, M., McGuire, M., Melega, W., Cherry, S., Huang, S.-C., & Phelps, M. (1996). Neural mechanisms supporting successful social decisions in Simians. In A. R. Damasio, H. Damasio, & Y. Christen (Eds.), Neurobiology of decision-making (pp. 63-82). Berlin: Springer-Verlag.

Ramachandran, V. S. (1988). Perception of shape from shading. Nature, 331, 163-166.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1999). Resonance behaviors and mirror neurons. Archives Italiennes de Biologie, 137, 85-100.

Rolls, E. T. (1999). The brain and emotion. Oxford: Oxford University Press.

Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe lesions. Brain, 124, 600-616.

Samuels, R. (1998). Evolutionary psychology and the massive modularity hypothesis. British Journal for the Philosophy of Science, 49, 575-602.

Samuels, R. (2000). Massively modular minds: evolutionary psychology and cognitive architecture. In P. Carruthers & A. Chamberlain (Eds.), Evolution and the human mind: Modularity, language and meta-cognition (pp. 13-46). Cambridge: Cambridge University Press.

Sawaguchi, T., & Kudo, H. (1990). Neocortical development and social structure in primates. Primates, 31, 283-290.

Scott, S., Young, A. W., Calder, A. J., Hellawell, D. J., Aggleton, J. P., & Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. Nature, 385, 254-257.

Shallice, T. (1988). From neuropsychology to mental structure. Cambridge: Cambridge University Press.

Shapiro, L. and Epstein, W. (1998). Evolutionary theory meets cognitive psychology: A more selective perspective. Mind and Language, 13, 171-194.

Sperber, D. (1996). Explaining culture: A naturalistic approach. Oxford: Basil Blackwell.

Sterelny, K. (1990). The representational theory of mind: An introduction. Oxford: Basil Blackwell.

Sterelny, K. (1995). Review of The Adapted Mind. Biology and Philosophy, 10, 255-285.

Sterelny, K., & Griffiths, P. E. (1999). Sex and death: An introduction to philosophy of biology. Chicago: University of Chicago Press.

Stone, V., Cosmides, L., Tooby, J., Knight, R., & Kroll, N. (under review). Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage. .

Stone, V. E. (2000). The role of the frontal lobes and the amygdala in theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), Understanding other minds: Perspectives from developmental cognitive neuroscience (2nd ed., pp. 253-273). Oxford: Oxford University Press.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. Journal of Cognitive Neuroscience, 10, 640-656.

Stone, V. E., Cosmides, L., & Tooby, J. (1996). Selective impairment of cheater detection: Neurological evidence for adaptive specialization. Paper presented at the 8th Annual Meeting of the Human Behavior and Evolution Society, Northwestern University, Evanston, IL.

Stuss, D. T., Gallup, G. G., & Alexander, M. P. (2001). The frontal lobes are necessary for 'theory of mind'. Brain, 124, 279-286.

Symons, D. (1992). On the use and misuse of Darwinism in the study of human behavior. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), The adapted mind: Evolutionary psychology and the generation of culture (pp. 137-159). New York: Oxford University Press.

Todd, P. M., Fiddick, L., & Krauss, S. (2000). Ecological rationality and its contents. Thinking and Reasoning, 6, 375-384.

Tomasello, M., & Call, J. (1997). Primate cognition. New York: Oxford University Press.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), The adapted mind: Evolutionary psychology and the generation of culture (pp. 19-136). New York: Oxford University Press.

Tooby, J., & Cosmides, L. (1995). Mapping the evolved functional organization of mind and brain. In M. S. Gazzaniga (Ed.), The cognitive neurosciences (pp. 1185-1197). Cambridge, MA: MIT Press/ Bradford Books.

Tooby, J., & Cosmides, L. (2000). Toward mapping the evolved functional organization of mind and brain. In M. S. Gazzaniga (Ed.), The new cognitive neurosciences. 2nd ed. Cambridge, MA: MIT Press/ Bradford Books.

Tranel, D. (1997). Emotional processing and the human amygdala. Trends in Cognitive Sciences, 1, 46-47.

Tranel, D., Bechara, A., & Damasio, A. R. (2000). Decision making and the somatic marker hypothesis. In M. S. Gazzaniga (Ed.), The new cognitive neurosciences. 2nd ed. Cambridge, MA: MIT Press/ Bradford Books.

Trivers, R. (1971). The evolution of reciprocal altruism. Quarterly Review of Biology, 4, 35-57.

Ts'o, D. Y., & Roe, A. W. (1995). Functional compartments in visual cortex: Segregation and interaction. In M. S. Gazzaniga (Ed.), The cognitive neurosciences (pp. 325-337). Cambridge, MA: MIT Press/ Bradford Books.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), Analysis of visual behavior (pp. 549-586). Cambridge, MA: MIT Press.

Van Essen, D., & Deyoe, E. (1995). Concurrent processing in the primate visual cortex. In M. S. Gazzaniga (Ed.), The cognitive neurosciences (pp. 383-400). Cambridge, MA: MIT Press/ Bradford Books.

Van Essen, D. C. (1985). Functional organization of primate visual cortex. In A. Peters & E. G. Jones (Eds.), Cerebral cortex, Vol. 3: Visual cortex (pp. 259-329). New York: Plenum Press.

Watanabe, Y., & Yamagishi, T. (1999). Emergence of strategies in a selective play environment with geographic mobility: A computer simulation. In M. Foddy & M. Smithson, et al. (Eds.), Resolving social dilemmas: Dynamic, structural, and intergroup aspects (pp. 55-66). Philadelphia, PA: Psychology Press.

Weiskrantz, L. (1956). Behavioral changes associated with ablations of the amygdaloid complex in monkeys. Journal of Comparative Physiology and Psychology, 49, 381-391.

Whalen, P. (1999). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. Current Directions in Psychological Science, 7, 177-187.

Wheeler, M., & Atkinson, A. P. (2001). Domains, brains, and evolution. In D. M. Walsh & A. O'Hear (Eds.), Evolution, naturalism and mind (pp. 239-266). Cambridge: Cambridge University Press.

Whiten, A. (Ed.). (1991). Natural theories of mind: Evolution, development and simulation of everyday mindreading. Oxford: Basil Blackwell.

Whiten, A. (1999). The evolution of deep social mind in humans. In M. C. Corballis & S. E. G. Lea (Eds.), The descent of mind: Psychological perspectives on hominid evolution (pp. 173-193). Oxford: Oxford University Press.

Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. Behavioral and Brain Sciences, 11, 233-273.

Whiten, A., & Byrne, R. W. (1997). Machiavellian intelligence II: Extensions and evaluations. Cambridge: Cambridge University Press.

Williams, G. C. (1966). Adaptation and natural selection: A critique of some current evolutionary thought. Princeton, NJ: Princeton University Press.

Young, A. W. (1998). Face and mind. Oxford: Oxford University Press.

Young, A. W., Aggleton, J. P., Hellawell, D. J., Johnson, M., Broks, P., & Hanley, J. R. (1995). Face processing impairments after amygdalotomy. Brain, 118, 15-24.

Young, A. W., Hellawell, D. J., Van De Wal, C., & Johnson, M. (1996). Facial expression processing after amygdalotomy. Neuropsychologia, 34, 31-39.