Contents

Impressum

Evolution and Cognition: ISSN: 0938-2623 **Published by:** Konrad Lorenz Institut für Evolutions- und Kognitionsforschung, Adolf-Lorenz-Gasse 2, A-3422 Altenberg/Donau. Tel.: 0043-2242-32390; Fax: 0043-2242-323904; e-mail: sec@kla.univie.ac.at; World Wide Web: http://www.univie.ac.at/evolution/kli/ **Editors:** Rupert Riedl, Manfred Wimmer **Layout:** Alexander Riegler **Aim and Scope:** "Evolution and Cognition" is an interdisciplinary forum devoted to all aspects of research on cognition in animals and humans. The major emphasis of the journal is on evolutionary approaches to cognition, reflecting the fact that the cognitive capacities of organisms result from biological evolution. Empirical and theoretical work from both fields, evolutionary and cognitive science, is accepted, but particular attention is paid to interdisciplinary perspectives on the mutual relationship between evolutionary and cognitive processes. Submissions dealing with the significance of cognitive research for the theories of biological and sociocultural evolution are also welcome. "Evolution and Cognition" publishes both original papers and review articles. **Period of Publication:** Semi-annual **Price:** Annuals subscription rate (2 issues): ATS 500; DEM 70, US$ 50; SFr 60; GBP 25. Annual subscriptions are assumed to be continued automatically unless subscription orders are cancelled by written information. **Single issue price:** ATS 300; DEM 43; US$ 30; SFr 36; GBP 15 **Publishing House:** WUV-Universitätsverlag/Vienna University Press, Berggasse 5, A-1090 Wien, Tel.: 0043/1/3105356-0, Fax: 0043/1/3197050 **Bank:** Erste österreichische Spar-Casse, Acct.No. 073-08191 (Bank Code 20111) **Advertising:** Vienna University Press, Berggasse 5, A-1090 Wien. **Supported by Cultural Office of the City of Vienna, Austrian Federal Ministry of Science/ Transportation and the Section Culture and Science of the Lower Austrian State Government.**

Jaak Panksepp and Jules B. Panksepp

The Seven Sins of Evolutionary Psychology

Abstract

Modern evolutionary psychology is demonstrating, once again, that an uncritical enthusiasm for the gene's-eye point of view can easily lead to conceptual excesses that go far beyond the available evidence. Seven major flaws in the evolutionary psychology agenda are outlined. With its enthusiasm for human inclusive-fitness issues, this variant of sociobiology has expressed little interest in what we already know about the brains and behaviors of non-human animals—facts that should be of foundational importance for thinking about many human abilities. To create a lasting understanding of 'human nature', we must incorporate the lessons from the past half-century of research on subcortical emotional and motivational systems that all mammals share. Seven examples of how a study of these systems can highlight some of the core problems of evolutionary psychology are outlined. From this perspective, the developmental interactions among ancient special-purpose circuits and more recent general-purpose brain mechanisms can generate many of the 'modularized' human abilities that evolutionary psychology has entertained. By simply accepting the remarkable degree of neocortical plasticity within the human brain, especially during development, genetically-dictated, sociobiological 'modules' begin to resemble products of dubious human ambition rather than of sound scientific reasoning.

Key words

Sociobiology, evolutionary psychology, brain, modularity, emotional systems, epigenetic landscapes, inclusive fitness, human nature.

Editor's Note

Since this article covers a lot of very controversial topics related to different disciplines it is considered as a target article, which will (in the forthcoming issues) be followed by several commentaries.

Only recently have psychologists seriously considered how the various abilities of the human mind were created during the long course of neural evolution. This approach, called evolutionary psychology, has captivated many investigators (see Wright 1994; Betzig 1997; Buss 1999; Cosmides/Tooby 2000), and it has encouraged the conceptualization of a variety of special-purpose evolutionary solutions (e.g., genetically ingrained adaptive functions or 'modules') that may exist within the human brain. The aim of this essay is to analyze the extent to which such approaches are providing unsubstantiated explanations of human behavior rather than clarifying realities of human and animal brain/minds. Many investigators, including ourselves, feel that evolutionary psychology has recently gone too far in its epistemological agenda, as it attempts to uncover the brain 'mechanisms' that constitute 'human nature'. In our estimation, such issues cannot be resolved without a full confrontation with the relevant cross-species, neuro–psycho–behavioral evidence.

Although we now realize that the 20th century image of the whole brain as simply a massive general-purpose learning machine was fundamentally incorrect, investigators have yet to demonstrate the existence of any sociobiological mechanisms that evolved in the massive human neocortex within the Pleistocene Environment of Evolutionary Adaptation (EEA). There is yet no well-established empirical reason for viewing any of those association areas of the neocortex as genetically pre-ordained 'modules' that generate specific types of psychological strategies. Although we have gained a new taste for natural mental kinds (e.g., intrinsic emotional categories) within the human brain (Brown 1991; Betzig 1997), we must remember to be especially cautious in ascribing discrete special-purpose functions to brain association areas that appear at birth to be largely general-purpose 'computational' devices. Many of the apparent special-purpose functions in the higher regions of adult brains may only emerge as a result of specific types of life experiences. In contrast, there are many special-purpose, genetically-dedicated circuits for various emotions and motivations in subcortical regions shared by all mammals.

The interactions between those specific brain operating systems and life experiences can, presumably, mediate the formation of an enormous diversity of 'modularized' software functions in higher neocortical regions of the brain. If this view is largely correct, we must proceed in a more epistemologically disciplined way than has become common practice in modern evolutionary psychology. Although we applaud the willingness of evolutionary psychologists to open up the Pandora's box of innate faculties within psychology once more, we fear that the parochial tendencies of many current views may promote needless controversies reminiscent of those that characterized the 'sociobiology wars' of the past quarter century. Although an appreciation of the power of inclusive-fitness can be incredibly productive in addressing many issues in population genetics and behavioral ecology, it cannot serve as a precise tool to dissect the nature of brain/mind mechanisms. How, then, might we generate credible perspectives that diminish the likelihood of arousing incendiary political passions, such as those that characterized the 'sociobiology wars'?

Biologists have long accepted evolutionary perspectives as historical scenarios for the emergence of all bodily organ systems. However, biologists have also come to recognize that evolutionary viewpoints are not especially useful for most of their ongoing experimental investigations. Evolutionary scenarios provide only marginal insights for guiding the experimental analyses of how biological systems actually function. Scientific demonstration of the functional mechanisms within the brain still need to be achieved through traditional experimental approaches. This poses a great dilemma for modern evolutionary psychological perspectives, for it is much easier to postulate adaptive 'modules' in the brain/mind than to demonstrate their neuropsychological nature. Such considerations lead to one overarching conclusion: Real neural functions across a variety of species should provide definitive constraints on speculation about what evolution did or did not create within human and animal brain/minds.

A new breed of evolutionary psychologists appears to disagree with such a marginal utility view of evolutionary scenarios. For the past dozen years they have been asserting, often with a tone of revolutionary fervor, that our ability to peer into the hazy crystal ball of 'recent' human ancestry will help us fathom the intrinsic nature—the evolutionary epistemology—of the human brain/mind. We, as well as many other scholars who have long accepted evolutionary principles as being ontologically correct, are forced to question this new and potentially virulent strain of dubious neo-DARWINIAN thinking. Without a strong linkage to neuroscientific research, evolutionary psychology has no credible way of determining whether its hypotheses reflect biological realities or only heuristics that permit provocative statistical predictions.

These considerations become especially pertinent when we consider that some evolutionary psychologists now explicitly claim their approaches can shed light on how the brain controls mind and behavior (e.g., see TOOBY/COSMIDES 2000). To us, this seems highly unlikely. Accordingly, we offer the following analysis to help direct psycho–evolutionary thinking in a more balanced and productive direction, where the available empirical riches from the Affective, Behavioral and Cognitive Neurosciences can be used effectively to construct a genuine image of how the human brain/mind is actually organized. In the first half of the paper, we take a conceptual approach, using the 'TOOBY & COSMIDES tradition' as the most prominent example of current thinking in the field. In the second half, we proceed to real brain issues that can be dissected empirically. We will not attempt to summarize specific sociobiological findings in this paper, and we shall assume that readers are reasonably familiar with the types of views that have been espoused by evolutionary psychologists during the past decade. At the outset, we regret that space constraints do not allow us to discuss all of the available evolutionary views in the detail needed for a comprehensive analysis.

The Creative Excesses of Evolutionary Psychology

To begin, we will briefly consider the general historical threads that have led to the present revolution in evolutionary thinking and then discuss several distinct ways to conceptualize the adaptive functions of the brain/mind. What is currently hailed as mainstream evolutionary psychology (i.e., symbolized most commonly by the cognitively-based tradition initiated by BARKOW/TOOBY/COSMIDES 1992) is making radical theoretical claims concerning the human mind, some of which are contrary to what is already known about the mammalian brain. We believe the evidential disparity between their adaptive theory of 'human nature' and current neuroscience understanding is largely due to the separate and remarkably non-interactive paths taken by psychological and biological approaches to the brain/mind during the 20^{th} century.

Once upon a time many philosophers and psychologists believed that the mind was a *tabula rasa* upon which raw experiences were transformed into knowledge through the power of associative learning. That era should have dimmed forever once DARWIN (1872) opened the door to a deeper understanding of human and animal minds. Evolution surely has constructed a variety of robust and perhaps fundamentally similar intrinsic potentials in the ancestral, neuro–mental apparatus of all mammals. However, since behavioral and bodily change are the only things we can directly measure in other animals, the search to uncover the nature of the intrinsic neuropsychological processes of the brain was delayed. Before the neuroscience revolution became fully recognized, modern evolutionary theory, especially with the robust concept of inclusive-fitness, provided psychologists with a substantive way to proceed (HAMILTON 1964). This principle has now achieved seminal recognition by many psychologically oriented investigators who are not, by tradition, accustomed to think in biological terms. Unfortunately, the acceptance of this profound evolutionary principle, which is most clearly applicable to sub-human species, did not promote an evident desire for the assimilation of neuroscientific research into human psychology.

Meanwhile, within ethology, behavioral genetics and comparative neuroscience there has existed a long and practical tradition of evolutionary thinking that continues to remain isolated from modern evolutionary psychology. There is no intellectually coherent reason for keeping the important findings of these fields un-integrated. Together, these disciplines can help create a balanced view of how the human mind was constructed. Empirical evidence indicates that the human mind was created through evolutionary shaping of ancient mammalian brains (MACLEAN 1990). Focusing on these ancestral emotional functions of the brain, to the extent that they are still represented in existing species, provides a unique empirical platform for thinking about the adaptive foundations of the human brain/mind (NESSE 1990; PANKSEPP 1982, 2000a–e).

We believe that the essential character of the human mind was laid down to a substantial extent within very ancient (i.e., subcortical) emotional and motivational neurochemical systems that we share with many other animals. Modern cladistic analyses of ancestral descent permitted by DNA, RNA and protein sequencing, along with the specification of neural systems in which such molecules are found, has dramatically increased our ability to uncover homologous brain functions across all mammalian species that have been studied. It is within the subcortical systems of the brain that the anatomical, neurochemical and functional homologies among mammalian brain/minds are most striking (MACLEAN 1990; PANKSEPP 1998a). It is also within these homologous brain areas where the most definitive human brain/mind 'modules' will be found. If we do not consider these shared proclivities, we will be led astray in trying to identify the abilities that have emerged from the unique higher brain functions of our species—namely from our cortico–cultural 'thinking cap'. The organization of the neocortex, although still constrained by many unknown genetic rules (e.g., KEVERNE et al. 1996; VANDERHAEGHEN, et al. 2000), may be much more of a general-purpose computational device than modern evolutionary psychologists have been willing to concede. Of course, there is abundant room for debate on most of these issues, but hopefully not of the kind that is disengaged from substantial segments of the relevant evidence.

From animal brain research, we know that there are a great number of special-purpose emotional operating systems in the mammalian brain (PANKSEPP 1998a). It may be that much of the potential explanatory power imparted by evolutionary scenarios will be found in the details of the robust and widespread influence that those systems have throughout mammalian brains. The possibility is remote, however, that many unique and detailed epistemological engravings of sociobiological strategies (i.e., modules) exist within the human neocortex. Moreover, the likelihood is high that many human behavioral tendencies, consistent with inclusive-fitness dictates, emerge as functions of individual experiences. Once a basic emotion has been aroused, the mind is wonderfully filled with cognitive processes that can be fertile breeding grounds for unseemly sociobiological strategies (i.e., epigenetically derived 'modules'). Yet, we must remember that many of these emotionally charged cognitive manifestations are driven by the sub-cortical emotional systems that exist in all mammalian brains.

The exploration of emotional systems is a major challenge that is recognized by many evolutionary theorists (NESSE 1990; BUSS 1999; COSMIDES/TOOBY 2000), but, rarely is existing brain evidence incorporated into such discussions. For instance, the adaptive 'fear' module postulated by COSMIDES/TOOBY (2000) does not adequately recognize one 'fear system' that has already been characterized in the brains of other mammals (PANKSEPP 1982, 1990a;

GRAEFF 1994; ROSEN/SCHULKIN 1998). The fear module envisioned by COSMIDES and TOOBY appears to be a master module that coordinates the activity of the many smaller modules dedicated to cognition and autonomic regulation. In fact, the type of fear 'module' that has been revealed by animal brain research appears to have come into existence long before any sophisticated cognitive capacities existed (PANKSEPP 1990a, 1998a). From the way Cosmides and Tooby appear to envision matters, the modularization of fear came after the existence of rather sophisticated cognitive capacities. Their fear module's main purpose is to simultaneously recruit and coordinate cognitive activities during fearful situations, so that the likelihood of an adaptive behavioral response is maximized. In fact, the experimentally demonstrated neuro–emotional systems are extensive, widely-ramifying subcortically situated circuits which posses the intrinsic capacity to modulate and synchronize a large variety of relevant brain and bodily resources. Evolutionary psychologists appear to be seeking specific socio–emotional modules among higher brain functions where the predominant functions may only be general-purpose cognitive/thinking mechanisms.

Only after a great deal of development can the cortex regulate emotional states by creating higher meanings through deliberation over fitness concerns. This view is quite similar to E. O. WILSON's suggestion at the introduction of *Sociobiology*, where he states that the genetically provided "emotional control centers flood our consciousness with all the emotions", not only "to promote the happiness and survival of the individual, but to favor the maximum transmission of the controlling genes" (1975, pp3–4). We now know an enormous amount about these ancient affect-generating, fitness-regulating systems—the "ancestral voices of the genes", to share the felicitous phrase of BUCK (1999). Unfortunately, that knowledge has yet to penetrate the sociobiology revolution that continues presently under the banner of evolutionary psychology.

The Plasticity of Language Cortex in Human Brain

We believe that some currently fashionable versions of evolutionary psychology are treading rather close to neurologically implausible views of the human mind. Although the lower reaches of all mammalian brains contain many intrinsic, special-purpose neurodynamic functions (e.g., basic motivational and emotional systems), there is no comparable evidence in support of highly resolved genetically dictated adaptations that produce socio–emotional cognitive strategies within the circuitry of the human neocortex.

Although there is substantial neuro–evolutionary evidence for the emergence of certain special abilities such as language (DEACON 1997; PINKER 1997), we still do not know with any assurance that such uniquely human abilities emerged from *de novo* genetic shaping of cognitive structures or from the remolding of preexisting adaptations (i.e., exaptations). It is certainly possible that language emerged within a spandrel of evolving multimodal brain complexities selected to generate internal imagery necessary for reflective consciousness, rather than being a deeply ingrained adaptation that emerged from evolutionary sculpting of the fine details in cortico–communicative neural circuits (GOULD 1991). The availability of extra general-purpose 'computational space' in the cortex may have been sufficient to allow language to 'emerge' without the guidance of specific evolutionary selective pressure. As CLARK (1997) put it, language may have adapted to the emerging complexities of the brain rather than the other way around. Although once language did emerge, it most certainly provided new social environments for both biological and cultural evolution (DEACON 1997). However, the result of such selective pressures, we would assume, was to dedicate more cortical tissue towards general-purpose symbolic processing. Once there was an urge to communicate and some new inter-subjective pragmatics (e.g., social intent and social gestures), language may have been guided by cultural evolution as much as by natural selection. In any event, the traditional view that language is a discrete, genetically-dictated specialization residing unconditionally in BROCA's and WERNICKE's areas (a simplification that CHOMSKY himself never accepted) has been crumbling for some time now (DEACON 1997). The reason the receptive aspects of speech find their natural home in WERNICKE's Area may largely be because that part of the neocortex is simply the brain's most multimodal area for integrating information coming from the external senses.

Although such radical degrees of revisionism may seem unlikely from certain popular points of view (e.g., PINKER 1994,1997), it is certainly not a perspective that has been falsified, and thus, should not be scientifically neglected. There are many lines of evidence that commend the view that recent human brain evolution provided the context for the emergence of a very general and flexible form of intelligence (e.g., PLOMIN 1999; SPENCE/FRITH 1999). Much of what traditional evolutionary psychology conceives to have been modularized in recent brain evolution may simply reflect our multi-modal capacity to conceptualize world events symbolically and to relate them to primitive affective feelings that reflect specific fitness concerns. Our massive memorial capacity then allows us to project these feelings and associated thoughts forward and backward in time, as well as onto other minds. There have surely been other remarkable evolutionary advances in the way ancient emotional systems interact with more recent cognitive processes (WIMMER 1995), but empirical knowledge remains meager (PANKSEPP 2000b,d,e). Indeed, without assimilating existing brain evidence, evolutionary psychology and psychology in general will remain on weak evidential footing with regard to the basic genetically-dictated psychological tendencies that guide the behavior of our species (PANKSEPP 1990b, 1998a).

Neuro–foundational Issues—the Evolution of the Brain/Mind

We subscribe to the view advocated by Paul MACLEAN (1990): the human brain is a structure consisting of distinct evolutionary layers, with many more homologies existing in the lower strata of the brain than in the higher cortico–cognitive layers. Our ancient reptilian basal ganglia and paleomammalian limbic system harbor many homologies, in comparison to the enormous species divergences at neocortical levels. Primary process consciousness is obviously based on subcortical circuits (DAMASIO 1999; PANKSEPP 1998a, 1998b). Hence, we doubt if it will be possible to reveal the intrinsic nature of higher aspects of the human brain/mind without first having a solid understanding of the lower aspects—the archetypal emotional-motivational processes that all mammals share.

Although rats and humans diverged in evolutionary history some 80 million years ago, there is evidence that we continue to share some remarkably similar emotional and motivational urges—evidence for an ancestral mind—if we simply consider all the available evidence (PANKSEPP 1998a). At present, it remains possible that most of the higher aspects of the human brain/mind arise largely from the interaction between general-purpose neural systems of the multimodal cortical association areas and the very basic life experiences encoded by more ancestral emotional/mind systems that all mammals share. This is not to say that there won't be many mismatches between the ancient operating systems of the human brain and the modern environments in which we currently live (PANI 2000).

In our estimation, 'minds' started to exist when the evolved complexities of the nervous system allowed organisms to know more than is contained within their reflexive responses to the world (GODFREY-SMITH 1996). In other words, the existence of 'mind' can be inferred whenever a substantial amount of the variability in the behavior of organisms needs to be understood with reference to the intrinsic, evolved 'representational' abilities of the brain. For instance, the ability to experience affect may be an essential antecedent to foresight, planning, and thereby willful intentionality. We place 'representational' in quotes to highlight our suspicion that the central dogma of cognitivism may be deeply misleading when it is applied to many fundamental brain functions, which may intrinsically seek meaning by dynamically 'grasping' the world (Freeman 1999; PANKSEPP 2000b).

In short, 'mind' represents our shorthand way of talking about the more creative ways that brains reach out into the world in their attempts to make sense of internal imbalances and environmental circumstances that can help alleviate those imbalances. In our estimation, the human mind, as well as all other mammalian minds, are fundamentally built upon ancient emotional and motivational value systems that generate affective states as indicators of potential fitness trajectories (DAMASIO 1999; PANKSEPP 1998a, 2000b,e).

By considering the knowledge that has been derived from other animals, we are ready to share an alternative vision of evolutionary psychology that is based solidly on our deep ancestral heritage. Such sources of evidence have been neglected in recent discussions of evolutionary epistemology because many investigators are understandably hesitant to deal with the difficult issue of cross-species comparison. Indeed, this has proved to be a most troublesome approach as the tendency of too many investigators has been to simply remain at the behavioral level, instead of considering the underlying causal (i.e., brain) mechanisms. Our ontological bias is as follows: We take the naturalist-pragmatist's view that all aspects of mind supervene upon the material functions of the brain (for the seminal discussion of *supervenience*, see KIM 1993). Although mind may not be simply reduced to neurophysiological functions, due to the genetic and neuroscientific revolutions, we can finally begin to understand the intrinsic nature of the human mind by scientifically exploring its lower, ancestral manifestations.

During the past decade, the functional architecture of the cognitive brain/mind in other mammals—the evolved dynamics of 'the great intermediate net' that intervenes between inputs and outputs—has become a topic of vigorous discussion (BUDIANSKY 1997; PANKSEPP 1998a; FREEMAN 1999; TOMASELLO 1999; HAUSER 2000). This recent cycle of intellectual activity has the strength to become a solid foundation for understanding the types of mechanisms that evolution truly built into the human brain. Now that there is a growing acceptance that we are thoroughly biological in both mind and body, and that the foundations of psychology make no sense except in an evolutionary framework (paraphrasing DOBZHANSKY's famous statement), let us be constrained by the evidence rather than captivated by the sea-swell of possibilities. In short, we believe that all too many 'stories' of evolutionary psychology may be scientifically explained by the interaction between basic emotional systems and the unique general-purpose intellectual abilities that human beings possess. If so, the foundations of 'human nature' will boil down to an 'animal nature' that was solidified in evolution long before the Pleistocene.

Let us now highlight seven specific 'sins'—flaws in epistemic strategies—we must critically consider before a fully synthesized evolutionary psychology can emerge and prevail in the brain/mind sciences. This summation of shortcomings will be followed by a seven exemplars of how knowledge of sub-cortical functions that humans share with other animals can help solve many of the basic fitness issues that evolutionary psychology has sought to clarify in our own species.

The Seven Sins of Evolutionary Psychology

Some of what follows repeats points already made (allowing each complaint to be read independently), which hopefully will only serve to reinforce their importance for the emergence of more substantive and lasting viewpoints in this field. Indeed, we shall see that many of the 'sins' reflect variants of the failure of evolutionary psychology to conceptualize adequately the emergence of various basic emotional and motivational systems in the mammalian brain and general purpose, cortico–cognitive abilities in the higher reaches of the human brain/mind.

1. Are there really Pleistocene sources of current human social adaptations?

Although all evolutionists recognize that existing organisms are living historical 'texts' that reflect past evolutionary passages, empirically we can only work effectively with the here and now brain/mind processes that are mixtures of evolutionary hand-me-downs and experiential blossoming. We can directly observe little more than strands of DNA, the proteins they help create, and the resulting developmental progression that takes place in specific environments. As is recognized by most, all historical/functional issues are largely hidden from any direct analysis. Since we have no time machine that can promote a reasonably credible analysis of specific phases of our ancestral past, the types of psychological adaptations that evolution provided must be extrapolated from a direct analysis of brain/mind processes that presently exist in other creatures. We can attempt to estimate the emergence of various general principles only to the extent that we have established credible brain structure/function relationships in many related species, and we should only believe narratives where convergent lines of evidence point in the same direction.

It seems to us that much of brain evolution during the Pliocene and Pleistocene eras was based upon the rapid expansion of general-purpose cortico–computational space (which permitted the emergence of foresight, hindsight and language) rather than on any fine-grained molding of special-purpose socio–affective mechanisms. Most special-purpose mechanisms in the brain, of which there are many in sub-cortical regions, evolved long before humans emerged as contenders for the top 'predator' position in the feeding hierarchy. Although those ancient special-purpose systems surely constrained subsequent brain/mind evolution in our line of ascent, we have barely started to fathom the resulting evolutionary epistemology—the 'affect logic'—that can come to exist within the higher reaches of the human brain (SCHAND 1920; WIMMER/CIOMPI 1995; SEGAL/WEISFELD/WEISFELD 1997; BOROD 2000; PANKSEPP 2000b,d).

Although many evolutionary psychologists are wisely backing away from creative speculations concerning the role of specific Pleistocene EEAs that are assumed to have been conducive to the selection of various psychological abilities, it might be wise to have a moratorium on such potentially idle speculations until what is already known about the functional organization of mammalian brains is integrated into evolutionary psychological thinking. Also, as noted earlier, animal husbandry practices and many behavior genetics experiments have indicated that it takes no more than a half dozen generations of selective breeding for robust temperamental differences to be induced into animal lines (Scott/ FULLER 1965; SEGAL 1999). Hence, we should not resist considering the possibility that a significant amount of phenotypic variability in human temperaments and intrinsic cognitive styles have been created by reproductive isolation or 'caste' precipitated assortative mating (FREEDMAN 1979; SEGAL et al. 1997), but such work should be done and communicated with a deep sense of responsibility for human sensitivities. The effects that can obviously be demonstrated, have multiple explanations and they will account for very little in our fundamental understanding of 'human nature'. Indeed, most of the genetically selected differences that have been documented are probably a matter of selection for differential emotional and motivational sensitivities and responsivities rather than any qualitative differences in underlying psychological 'kinds'. Some people may simply never really understand the concept of color because they are color-blind; others may simply not understand 'emotions' because their emotional systems are not as strongly developed as those of others. In any event, our remarkable similarities to other mammals, especially at an emotional/affective level, should not be underestimated. However, we should also recognize that different species, and perhaps different lines of the human family, have dispositionally and developmentally distinct patterns of emotional and motivational expressions (to put an incredibly complex topic in shorthand).

2. Excessive species-centrism in evolutionary psychology

Evolutionary psychology remains enthralled by the human species. Although this species-centrism is understandable considering the anthropocentric biases of most modern schools of psychology, it creates enormous shortcomings in our ability to tackle basic issues of human brain/mind evolution. Many of our basic values and ways of ascribing affective meaning to events may be based on primitive brain processes that are homologous in all mammals. If so, we should be able to effectively study these processes in animal models. A highly restricted focus on human issues is bound to be less productive than judicious integration of demonstrable evolutionary antecedents in our conceptual structures (PANKSEPP 1998a). Recognition of the ancient emotional systems may help frame many of the ideas of evolutionary psychology in more archetypal, evolutionary terms. This may be a more fertile way to understand the primordial sources of our essential affective abilities and emotional tendencies.

We do not believe that it makes intellectual sense to follow a fundamental speciesism in evolutionary psychology, where human proclivities are commonly discussed independently of what we share with other creatures. There is a profound continuity in the subcortical neuro–mental processes among all mammalian species, especially with respect to deep emotional-motivational issues. If we do not fully recognize these shared processes, we can easily construct intellectual houses of cards by solely considering the final products of human socio–cultural life. Despite what some culturologists and social constructivists believe, the modern human mind is still tethered to ancestral animal minds. But, as they also claim, most of what we outwardly do is obviously constructed from our massive and highly generalized intellectual abilities interacting with cultural contexts. Of course, there are bound to be perceptual canalizations that are unique to humans (e.g., PERRETT et al. 1998; PENTON-VOAK/PERRETT 2000). In any event, brain/mind evolution during the past few million years probably operated mostly by generating new general-purpose regulatory mechanisms that could flexibly solve the endogenously generated affective dictates aroused by the ancient emotional systems we share with many other animals.

3. The sin of adaptationism

Little needs to be said on this issue, since we have all become sensitized to the flaws of Panglossian thinking (GOULD/LEWONTIN 1979), and the necessary mechanistic and conceptual distinctions have been acknowledged (e.g., BUSS et al. 1998). In any event, if we do share a variety of basic emotional and motivational systems with all other mammals, there surely should be little doubt that those systems reflect profound cross-species adaptations. However, difficulties become more apparent when we come to consider the evolution of higher cognitive abilities. Since the emergence of massive, general-purpose cortical space, exaptations and spandrels have arisen everywhere we look (GOULD 1991). In this conceptually tricky territory, it is easy to imagine intrinsic brain/mind patterns that appear to be evolutionarily engraved in the higher regions of the brain, whereas in reality there are only social and cultural cortico–constructions that are simply tethered to more primitive attentional, motivational and emotional systems. Ancient emotional systems are able to imbue 'cold' perceptions with 'hot' affective charge. Thus, we must reconsider which higher intellectual tools and perceptual proclivities actually reflect inherent tendencies of the higher reaches of the brain as opposed those arising from learning.

Within the neocortical reaches of the brain, we suspect data-constrained scholars may only agree that all mammals have the ability to perceive objects and events, to compute temporal passages of time between events, to navigate through objects in space (HAUSER 2000; SILVERMAN et al. 2000), and perhaps several other species-typical resource finding capacities (e.g., KANWISHER 2000; PENTON-VOAK/PERRETT 2000). Unfortunately, evolutionary psychology, rather than deriving most of its impact from a discussion of such general cognitive capacities, typically entices us with the allure of much juicier emotional and motivational stories that color human life. However, at present, there is very little evidence for those types of discrete cortico–cognitive adaptations, even though we can anticipate that special forms of *affect logic*—'centers of gravity' for emotion-cognition interactions—emerge in higher brain areas developmentally (SCHAND 1920; FRIDJA 1986; WIMMER/CIOMPI 1995; CIOMPI, 1997; PANKSEPP 1989 2000b). Although sociobiological modular possibilities may certainly exist in the higher regions of the human brain/mind, it is essential to try to demonstrate them rather than to simply argue for their existence based upon commonly observed phenotypic expressions. Many of the postulated cognitive adaptations, upon closer examination, may simply turn out to be emergent properties of development and culture (GOULD 1991; SCHAFFNER 1998).

4. The sin of massive modularity

After the reign of general-purpose behaviorism/associationism was declining, FODOR (1983) opened up the Pandora's box of innate faculties by simply accepting what to most was obvious—that there are brain/mind modules for all of our basic sensory/perceptual and motor processes. However, we can now be equally confident that there is also a great amount of general-purpose computational space (heteromodal tissues). The phenotypic expression of complex mental and behavioral tendencies can be generated by so many different mind-brain processes that any ascription of intrinsic modularity to human association cortex must presently be deemed intuition-trading rather than evidence–based argumentation (SAMUELS 1998). Although it may be possible that some unique genetically-channeled resource-holding mechanisms related to social-emotional needs do exist in higher areas of the brain (e.g., systems for monitoring attractiveness, greediness, degree of social reciprocity and commitment, as well as 'mind-reading' tendencies), it should be the responsibility of each investigator who posits such modules to plot out credible strategies for revealing their inherent nature. Without that and with the recognition of general purpose intelligence mechanisms in the brain, it may be wiser to accept as a default assumption that most socio–modular functions are simply due to epigenetic emergence. Of course, this makes the search for ingrained socio–emotional systems of the brain even more interesting (CARTER/LEDERHENDLER/KIRKPATRICK 1999).

The traditional position, certainly not negated by available evidence, is that most higher cortico–cognitive functions are epigenetically created by the experiences of organisms. Although there is abundant evidence for the emergence of neural complexities that permitted language and a general increase in propositional intelligence (SCHEIBEL/SCHOPF 1997), there is presently no clear evidence that new and refined emotional modular functions emerged in the human brain/mind during the past several million years of human brain evolution. Indeed, when humans have strong affective experiences, higher cortical regions tend to shut down (DAMASIO et al. 2000; FISCHER et al. 2000). Although it is certainly possible that modules for social affiliation, empathy, pride and various other resource holding capacities emerged within the massive neocortical mushrooming that occurred in our line several million years ago (NESSE 1990), it is as easy to envision how such sociobiological processes may emerge from the interaction of basic emotional systems with higher general-purpose propositional abilities. Indeed, our facility with general-purpose representational abilities (e.g., internal imagery and language) may have been as important in generating such emergent adaptive processes as any type of special-purpose epistemological engravings within the expanding neocortex.

Obviously, adaptive behavior can be genetically, experientially and culturally guided. The genetic components of adaptive behavior can only be distinguished by especially stringent criteria. As BUSS and colleagues (1998) have enumerated, but have not yet empirically resolved, genetically ingrained adaptations can be identified by specific criteria such as "complexity, economy, efficiency, reliability, precision and functionality" (p536), but learning can also provide these results. In our estimation, such conceptual criteria can only be cashed out by more rigorous developmental-neurobehavioral approaches than have yet to be implemented by evolutionary psychologists. In any event, the simple postulation of genetically-dictated modules, especially in higher areas of the human brain, may end up being a regressive 'phrenological' strategy rather than a progressive paradigm based on real brain circuit analysis.

We should not forget that it took the better part of brain evolution to create the subcortical systems of the mammalian brain, while the expansion of the cortex was remarkably rapid, guided probably by a small number of genetic regulatory changes. Morphometric analysis suggests that the higher brain as a whole enlarged, with no clear indication for specialized modular development, at least at the gross anatomical level (FINLAY/DARLINGTON 1995). Let us recall that the 'chips' of the human cortex—the columnar structures containing approximately 3,000 neurons each—are very similar throughout the brain and also from one mammalian species to another. These features are suggestive of highly generalized (almost 'random access', chip-type) computational devices.

Although functional specialization must have emerged as cortical columns proliferated and interconnected in increasingly complex ways, there is presently little empirical data, aside from certain perceptual and motor/action processes (e.g., VANDERHAEGHEN et al. 2000), that any types of functional psychological strategies emerged in the cortex via natural selection. Perhaps the identification of the learned inputs into the headwaters of the fear circuitry in the amygdala (LEDOUX 1996), and the reward-association learning circuits of the frontal lobes (ROLLS 1999) might be taken as exemplars, but they may also reflect general learning principles operating in higher brain areas to which intrinsic subcortical emotional circuits project (PANKSEPP 1998a).

Put another way, the relatively homogeneous columnar organization of the neocortex is not straightforwardly compatible with any highly resolved, genetically-governed, modular point of view. Indeed, functional studies suggest a vast plasticity in many of the traditionally accepted cortical functions. For instance, visual cortex can be destroyed in fetal mice, and visual ability will emerge in adjacent tissues (see DEACON 1990, 1997). Accordingly, the heteromodal cortex of the human brain may be better conceptualized as a general purpose cognitive-linguistic-cultural 'playground' for regulating the basic affective and motivational tendencies that are organized elsewhere. In this view, cognitive processes are 'tools' or 'handmaidens' for helping regulate the more basic life concerns.

5. On the conflation of emotions and cognitions

Common human subjective experiences highlight how massively emotional feelings and cognitive attitudes are intermeshed. Clearly cognitions can easily become 'charged' with emotional values. This has led some (e.g., PARROT/SCHULKIN 1993; COSMIDES/TOOBY 2000) to ignore the evidence that the basic emotional circuits of the brain emerged much earlier in brain evolution than the higher cognitive capacities. For those not accustomed to neuro–evolutionary thinking, the 'ancientness' of basic emotional systems is supported by their medial and caudal locations in the brain (DAMASIO et al. 2000; PANKSEPP 1998a), as well as the dating of neurogenesis in the underlying brain zone (i.e., the timing of the fetal development of brain systems seems to parallel the historical pattern of their phylogenetic origin).

In short, the classic distinction between emotional and cognitive processes is sustained by abundant data indicating that the two can be dissociated functionally and anatomically (e.g., ZAJONC 2000; PANKSEPP 1990c, 2000b). Even though emotions and cognitions obviously interact massively, there is no scientifically sound reason to conflate the two, especially in subcortical realms where the power of affect seems to be elaborated (OLMSTEAD/FRANKLIN 1997; PANKSEPP 1998a, 1998b, 2000b) However, perhaps more important for the present argument is the issue of whether genetically guided cognitive specializations exist in higher regions of the neocortical apparatus to generate higher order emotions such as jealously, shame, guilt, pride, etc. There may well be such intrinsic higher 'centers of gravity' (e.g., in subareas of the frontal lobes) for the blending of basic emotional impulses (WEISFELD 1997; DAMASIO 1999; ROLLS 1999; PANKSEPP 1982, 1989), but at present, there is no critical neurodevelopmental evidence to adjudicate how such functions actually emerge. Although there may well be abundant genetically governed canalizations for such processes, it is still equally likely that most are created through cognitive learning and other developmental processes guided fundamentally by the affective power of basic emotional circuits concentrated in subcortical regions of the brain.

6. The absence of credible neural perspectives

Although there is increasing talk of neural circuits for cerebral modules, especially since evolutionary psychology became a compelling view in cognitive neuroscience (e.g., PINKER 1997; TOOBY/COSMIDES 2000). Yet, none of the proposed sociobiological modules have coalesced with established neural realities. In this context, it is remarkable that evolutionary psychology continues to neglect evidence concerning the basic socio–emotional systems of the mammalian brain that have been studied for many years in animal models (e.g., note the absence of existing neuroscience work in a recent chapter by COSMIDES/TOOBY 2000 which is juxtaposed to a summary of such work by PANKSEPP 2000a).

Instead of a solid confrontation with the brain, there is abundant talk of computational-representational views which ignore the fact that many neuroscientists are not convinced that such information-processing metaphors provide much that resembles an accurate perspective on how the brain creates meanings (Freeman 1999). Perhaps the higher cortical systems are 'computational' by some stretch of the digital information processing metaphor, but the subcortical reaches that mediate emotions and motivations are not. Subcortical systems generate many neurochemically-specific mass-action effects in the brain where the patterning of action-potentials carry no psychological or behavioral 'codes', but their population frequencies do control the intensity and patterning of specific action tendencies. We should never forget that the capacity to simulate certain brain functions in a digital processor does not mean those computations reflect physiological realities. Also, we should at least consider that many brain functions are created not simply by digital informational codes but by 'volume-conduction' types of analog mechanisms (explaining why 'pressure' and 'energy' metaphors have been so popular in emotion and motivation research).

Although we do not want to distance ourselves completely from the potential utility of computational approaches for understanding the brain (e.g., MAUK 2000), we should remember that brain 'computations', even in the higher cortical regions, are probably vastly different than those that transpire in digital computers. Biological brains contain massive internally generated background activities that help establish spontaneous, self-organizing, non-linear dynamic capacities of which digital 'brains' seem incapable (FREEMAN 1999; LEWIS/GRANIC 2000). In other words, the foundations of mind are fundamentally 'embodied' by organic processes that are impossible to compute except in the most superficial ways. The exploitation of ever-popular computational-informational metaphors may not really be instructive for understanding the essential organic underpinnings of the human mind. A realistic confrontation with the biology of neural systems most assuredly will. Of course, considering the power of computer algorithms to superficially model essentially anything, it is understandable why many continue to be enticed by traditional computational metaphors.

7. Anti-organic bias or the computationalist/ representationalist myth

Let us consider the previous 'sin' from the perspective of recent flirtations by evolutionary psychologists with proximal mechanistic analyses (e.g., TOOBY/COSMIDES 2000). Evolutionary psychology, as most other forms of cognitive psychology, has been inspired by computer science rather than the molecular biology/neuroscience revolutions of the past three decades. Indeed, at times, it seems that practitioners of evolutionary psychology have an active aversion to organic perspectives. They talk about the brain simply as a modular computational device. Although that view has also been pushed forward by many cognitive neuroscientists, an equally credible alternative is rarely discussed: Much of mental life is fundamentally organic. The only reason the brain-mind appears to be computational is because nerve cells fire action potentials, yielding a surface similarity to SHANNON-WIENER type informational principles (CAMPBELL 1982). However, we could view those microscopic elements of neural activity as mere 'stitches' within a more dynamic, amplitude modulated neural fabric of mind that must be understood fundamentally in organic terms. From this view, action-potentials simply provide the necessary mass-action effects to generate a more global neurodynamic fabric of the mind. The molding and shaping of the population-dynamics arising from ensembles of neurons may be more important in generating a substantive understanding of mind, than is the search for discrete digital codes of action potentials in individual neurons.

We should certainly remember that no sophisticated digital 'neuronal code' for psychological processes has yet been found via a study of action potentials. At best, we have some impressive neuronal correlates for some sensory and perceptual processes, but an equally impressive amount of evidence that specific neurochemical patterns in the brain can create basic psychological states (PANKSEPP 1986b, 1993, 1998a). One could argue that processes such as emotions could be instantiated by any of a variety of neurochemistries or computational devices, but at present that is an empty promissory note and a view (Dare we call it 'dustbowl cognitivism'?) that does not encourage the search for practical knowledge that could link up with basic human and clinical issues, although recent developments in evolutionary medicine and psychiatry certainly do have potential (e.g., NESSE/WILLIAMS 1974; STEVENS/PRICE 1996; MCGUIRE/TROISI 1998).

Just to highlight how remarkably ingrained is the bias against the organic brain, we note that in *How The Mind Works*, PINKER (1997) shares a vivid description of the computational bias. He describes a scenario of information transfer in which one person says something into a telephone receiver, and the 'information' then gets successively converted from nerve cell and muscle activity; to air pressure oscillations (i.e., auditory signals); to electrical signals traveling down wires; to electrochemical reactions within a silicon medium and subsequently back in the completely reverse order. The receiving brain converts this 'information' into a vocal response that can be perfectly and reliably conveyed to a companion sitting on the couch. PINKER then makes the point that the information remained unaltered and in perfect form, independent of the medium in which it traveled. This view almost completely ignores the simple fact that the brain has actively created 'meaning' out of sensory events. Few ideas have been put forward about how this is achieved.

By contrast, an esteemed hands-on neuroscientist working on such profound issues has noted that "The only patterns that are integrated into the activities of the brain areas to which the sensory cortices transmit their outputs are *those patterns they have constructed within themselves*. In colloquial terms, the ingredients received by brains are not direct transcriptions of impressions from the environment inside or outside the brain. All that brains can know has been synthesized within themselves" (FREEMAN 1999, p93, our emphasis). Although we do not suggest that auditory and visual information is not strongly restrained by computational algorithms at the neuronal level, we do assert that, at present, computationalism has added no fundamental understanding to how the brain generates emotions and motivations.

Nevertheless, PINKER uses his example to explain why neuroscience deserves less merit than information/computational theories in attempting to study mind. Yet, he fails to emphasize that the only entities that have ever been demonstrated to possess mind are those that possess neurally- (as opposed to silicone-) based brains. To rephrase a previous point, it is a mistake to believe something is biologically real simply because one can computationally simulate the shadow of an end result. Until someone discovers facts to the contrary, we should continue to acknowledge that there is something remarkably special about the organic brain medium. Abundant examples exist that would be very hard to explain from any existing computational view, but can easily be explained by the organic properties of brain tissue. Thus, it is certainly premature and unwise, for any science of the mind to neglect the brain, as is still too common in evolutionary psychology and most of psychology, in general. (ROBINS et al. 1999).

In sum, from a neuroscience perspective, there is certainly no paucity of sins in current evolutionary psychology. Hubris and reification of verbal concepts are not in short supply. Fortunately, only a few have had the temerity to claim that their speculations are providing the conceptual structures that can effectively guide future neuroscience investigations (TOOBY/COSMIDES 2000). So far evolutionary psychology has only been effective in framing stochastic predictions in terms of presumed distal evolutionary adaptations guided largely by Hamilton's concept of inclusive fitness. However, now that evolutionary psychologists are persistently talking about 'mechanisms' [e.g., note the use of that term three times in the introductory paragraphs of a fine paper by BUSS (2000) on the evolutionary nature of happiness], they need to invest vigorously in the underlying causal and supervenience issues rather than simply making computational assumptions concerning the nature of the underlying processes.

Now that sophisticated brain imaging devices and samples of anatomically well-characterized brain-damaged individuals are available, we anticipate that evolutionary psychologists will attempt to cash out their claims using neuroscientific approaches. However, since those approaches are not likely to yield anything but correlative data, it is bound to be a rocky road toward any substantive understanding of the underlying mechanisms. Thus, we strongly urge the discipline to cultivate good relations with various animal brain research traditions that can help them reveal, in some reasonable detail, the underlying causal mechanisms. To facilitate this hope, we now share some of the intriguing possibilities from animal brain research that can help evolutionary psychology ground itself in a more catholic empirical tradition.

Seven Solutions from Neuroevolutionary Psychobiology

In this section, we highlight seven specific examples, from a vast pool of available evidence, of how existing behavioral and affective neuroscience research can clarify some of the basic neural underpinnings of major problem areas in evolutionary psychology. These examples also help highlight how the interactions between ancient brain emotional systems and general purpose cortico–cognitive abilities can contribute insight into how we might proceed to solve some of the fascinating, but often self-evident (i.e., folk-psychological), human tendencies that evolutionary psychology has helped bring to the intellectual foreground (e.g., see Table 1, BUSS et al. 1998). Although there are a large number of options to choose from, most come from the senior author's long-term research program into the fundamental nature of mammalian emotions. Since evolutionary psychology has such a unique and highly delimited epistemological agenda, these ideas may have little impact on those who already have well-established positions in the field. However, we offer this 'sampler' for all scholars who are seeking a comprehensive understanding of the ancestral roots of 'human nature'.

1. A general purpose foraging system

The mammalian brain contains a powerful subcortical system that can generate the seeking of resources essential to survival (ROBINSON/BERRIDGE 1993; PANKSEPP 1981, 1982, 1986a, 1992, 1998a). This so-called expectancy/SEEKING or 'wanting' system has been the focus of work for decades, typically guided by discrete behavioral hypotheses of reward, reinforcement and more recently pleasure. Now, an increasing number of investigators, taking their lead from ethological rather than behaviorist analyses, are recognizing that this a generalized system for foraging—a system that provides a goad with no fixed goal for exploratory/investigatory activity. This system is capable of helping construct goal-directed behavior patterns based on the confluence of bodily need states, environmentally accessible reward objects, and contextual contingencies. In evolutionary terms, this system could be conceptualized as a generalized positive appetitive 'resource holding potential' system that monitors and promotes fitness issues by instigating vigorous exploratory-seeking activities (PANKSEPP/KNUTSON/BURGDORF 2001).

We will not attempt to even allude to the remarkable behavioral neuroscience work that has been done on this system (for a recent review see IKEMOTO/PANKSEPP 1999), but we highlight this circuitry simply as one example of a variety of widespread subcortical emotional systems shared by all mammals (and probably some other animals). This system has broad implications for a large number of appetitive behaviors that can vary considerably depending on the contexts encountered by such animals. It also surely controls a diversity of human/animal aspirations and desires.

This brain process also helps establish confirmation biases in organisms—coaxing them to behave with causal 'convictions' when only correlations exist in perceptual inputs (for a summary of the relevant 'auto-shaping' literature, see PANKSEPP 1981, 1982, 1986b). We suggest that much of evolutionary psychology, indeed much of science, proceeds on the inductive inferences made by such forward-looking, experience-expectant brain processes. Unfortunately, many resulting conclusions constructed in the aroused cortico–cognitive spaces of the mind turn out to be delusional from formal logical perspectives. We urge evolutionary psychologists to ponder the implications of this general-purpose motivational system for the ways they are seeking to understand 'human nature'. We also would suggest that functionally dedicated subcortical systems for RAGE, FEAR, LUST, CARE, PANIC and PLAY be considered as foundational (PANKSEPP 1998a) for the creation of many of the socio–emotional 'modules' that are currently being entertained. The general point is that despite such dedicated systems for emotions, human behavior and the underlying brain systems are much more plastic than evolutionary psychologists commonly emphasize.

2. General-purpose neurochemical systems for regulating all psycho–behavioral tendencies

Consider just one neurochemical system: Ascending brain serotonin circuits arising from two compact midbrain cell groups ramify widely throughout the forebrain. Through a diversity of distinct receptors, these networks modulate all emotional and motivational process in all mammalian species in essentially similar ways (PANKSEPP 1986a, 1998a). In general, when serotonin is high, mammals appear more relaxed, satisfied and confident. They are less likely to initiate aggression, but also less likely to back down during social confrontations (see MCGUIRE/TROISI 1998). Humans respond likewise

(KRAMER 1993; KNUTSON, et al. 1998). Similar behavioral patterns, modulatory functions, cellular characteristics (autoinhibition) and diffuse projection patterns have also been documented in invertebrate serotonin systems (HUBER/DELAGO 1998; HEINRICH et al 1999; KRAVITZ 2000).

Since such neurochemical issues are probably the most credible means of confirming homologies between animals and humans (PANKSEPP 1986b, 1993, 1998a), let us dwell on this issue at some length: In order to clarify whether domain-specific brain modularity is a physical reality, evolutionary psychology as well as the neurosciences must rely on the existing evidence of brain structure and function to support all claims that attempt to elucidate how the brain/mind operates. Historically, the most thoroughly studied neural systems have been amine-modulatory systems that utilize small amino acid derivatives (dopamine, norepinephrine, serotonin) as neurotransmitters. Evidence from these well-studied systems provides a suitable foundation for brain/mind theorizing in evolutionary psychology. The group of neurochemicals that are utilized by these systems (collectively referred to as biogenic amines) are present in the nervous systems of many animal groups, including molluscs, annelids, crustaceans, and all mammals. Most neurons that produce biogenic amines reside in discrete clusters of cell bodies that are situated near the midline of the nervous system. Their projections stretch over large areas of neural tissue and release chemical messages diffusely, rather then through information-specific synaptic transmissions. In essence, functions of amine-modulatory systems have remained highly conserved across a remarkably wide range of species.

Instead of producing neuronal activity per se, this type of global release regulates ongoing nerve cell activity by changing the response properties of large neuronal ensembles. Amine-modulatory systems have been indicated as key elements in the regulation of a broad range of appetitive behaviors in every species studied. For example, calcium-calmodulin kinase II knock-out mice exhibit marked decreases in serotonin release from the raphe nuclei (i. e. the serotonin–producing cells in the mammalian brain) and have been shown to be less fearful in situations where animals normally exhibit heightened fear responses (e.g., foot shocks, fear conditioning, 'open-field' tasks, re-engaging aggressive conspecifics). Heterozygous mutants are also much more aggressive in various behavioral paradigms (CHEN et al. 1994). Even more notable from the present perspective is the role of biogenic amines in a wide range of human psychological abnormalities. Serotonin-modifying drugs can treat depression, anxiety, hyper-aggressive tendencies and eating disorders (BITTAR/BITTAR 2000).

Alterations in this widespread neuro–modulatory system produce effects that span across large behavioral repertoires. The close association between these systems and global arousal, attentional and appetitive states should make us dubious about any suggestion that would ever attempt to ascribe unique human brain/behavioral propensities to these systems. The basic plans of these neural networks were established long before humans existed. Although these chemistries also came to regulate the types of higher brain tissues that emerged in the humanoid line during the Pleistocene, there is no indication that biogenic amines ever evolved to participate in anything more than non-specific modulation of all attentional, cognitive and emotional functions, albeit at times in remarkably subtle ways. For instance, KRAVITZ (1988, 2000) has proposed that aminergic modulation may function to recruit adaptive behaviors over contra-adaptive behaviors. Unlike the computationally resolved cognitive models that evolutionary psychology suggests underlie mind and behavior, the amine-'spritzer' systems tend to support a more general, organic explanation where global, neurochemically induced field-dynamics set the tone for what the brain is likely to produce.

For 'modular' evolutionary arguments to work, we have to be able to specify how selection pressures can mold specific neural circuits. All perceptual and cognitive specificity appears to be driven by a few excitatory and inhibitory amino acids (e.g., glutamate, GABA), and it is very hard to envision how genetics could mold the detailed wiring of those systems, but easy to imagine having different amounts of such computational networks in different areas of the neocortex. We believe that humans simply have much more of such brain tissue than other animals, without it being uniquely dedicated to any specific inclusive-fitness functions. Admittedly, there are bound to be some preferred modes of information transmission within such cognitive networks (i.e., canalizations or forms of preparedness guided by functionally dedicated systems of deeper parts of the brain), but most of the higher results are bound to be dramatically shaped by the individual experiences of each person.

By comparison, it is easy to envision how various peptide modulatory systems that we share with the other animals are more dedicated for specific evolutionary ends. There is now abundant evidence that neuropeptide systems can modulate very specific emotional and motivational tendencies (e.g., as reviewed by PANKSEPP 1986, 1993, 1998a). Working together with such ingrained emotional and motivational systems, the intensity of whose influence could easily be modified by genetic selection, higher cognitive systems can surely be epigentically molded more readily by differential fitness demands that vary with individual environmental exigencies. This is not to say that emotional systems do not also change with experience; they certainly do (PANKSEPP 2001) and with many cognitive consequences (LIU et al. 2000).

3. General mammalian mechanisms of kin selection

Although a great deal of discussion in sociobiology and evolutionary psychology has been premised on the nature of social relationships (e. g. inclusive fitness, altruism, kin-recognition—see HAMILTON 1964; WILSON 1975), for over a quarter of a century these disciplines have remained silent about how proximal mechanisms in the brain might control such processes. Although there is abundant talk about 'mechanisms', acknowledgement of the abundant ongoing neuroscience work on such issues is rarely offered. Meanwhile, behavioral neuroscientists have been making considerable progress in understanding these processes by studying brain functions related to the emotions of separation distress and attachment (PANKSEPP 1989/1999). Among the neural systems that are especially important in the mediation of social attachments are brain opioids, oxytocin, glutamate, norepinephrine and probably prolactin (INSEL 1997; NELSON/PANKSEPP 1998; PANKSEPP 1998a). Just one of many noteworthy findings is that kin-appreciation in male mice is mediated by the release of endogenous opioids (D'AMATO 1998), and the tolerance that occurs in opioid synapses, and hence social bonds, could easily help explain inclusive-fitness related changes in social strategies (PANKSEPP 1998a). Surprisingly, evolutionary psychologists have exhibited little interest in integrating or advancing such findings, even though there is direct and rather profound implications for their scientific agendas.

4. On the nature of play, power, tickling and friendship

Powerful social control systems for playful engagement exist in subcortical regions of the brain, and their functions are beginning to be systematically dissected (VANDERSCHUREN, et al. 1997; PANKSEPP, et al. 1984; PANKSEPP 1993b). These systems allow higher brain regions, perhaps via recently discovered 'mirror neurons' (RIZZOLATTI et al. 1996; GALLESE/GOLDMAN 1998), to establish and solidify social strategies that have enormous implications for the structuring of animal societies. Adult relations, including alliance-friendship patterns, may arise as a result of the high degree of activity in social engagement systems during juvenile development. The desire to engage in such activities is communicated in various ways that are not well understood, but discoveries such as play vocalizations, that may have more than a passing resemblance to primitive forms of human laughter (KNUTSON et al. 1998; PANKSEPP/BURGDORF 1999 2000), may help open the door to systematic studies of how positive social emotions regulate the construction of personality differences, as well as social systems (PANKSEPP 2000e).

A study of these systems also has implications for seemingly distant concepts such as anticipatory eagerness (BURGDORF et al. 2000), the cravings that accompany drug addictions (PANKSEPP et al. 2001) and the nature of current social problems such as the increasing incidence of attention-deficit, hyperactivity disorders (PANKSEPP 1998c). The underlying evolutionary pressures that have molded such basic mammalian play systems should be of foundational importance for understanding the underlying nature of many inclusive-fitness concepts in evolutionary psychology.

5. Male sexual jealousy and potential neurochemical therapies

The fact that males are defensive over reproductive opportunities has been repeatedly noted in the human and animal behavior literature. We finally have animal models that suggest subcortically situated vasopressin systems are critically involved in such processes (WINSLOW et al. 1993). Male voles become highly aggressive toward intruders after a single sexual experience with a specific female. However, if their brain vasopressin signals were immobilized with receptor antagonists, a maneuver that does not compromise copulatory success, the males do not exhibit offensive behavior toward intruders. Conversely, the mere experience of elevated brain vasopressin in the presence of a female conspecific, but in the absence of sexual contact, is sufficient to establish the aggressive attitude toward 'intruding' males. In contrast, there is a parallel line of investigation that indicates oxytocin can promote maternal care and increase friendly relations among individuals (CARTER 1998; NELSON/PANKSEPP 1998). These types of subcortical processes appear to be essential for the emergence of positive social relations, which may be of foremost importance for the development of a 'mechanistically' sound and clinically productive understanding in evolutionary psychology. For instance, if there are deep evolutionary homologies in these systems, we anticipate that orally-effective vasopressin antagonists might eventually be designed to combat male sexual jealousy and related forms of male irritability.

By comparison, orally-effective oxytocin agonists should promote feelings and attitudes of nurturance, and thereby promote many higher psychological reasons for being less aggressive. Genetic manipulations along these lines have already modulated the social temperament of animals (YOUNG et al. 1999). One relevant finding in modern evolutionary psychology is that females tend to have better social memories than males (SALMON/DALY 1996), and in that context it is worth noting that mice whose oxytocin gene has been disabled exhibit deficient social memories (Ferguson et al. 2000). Both of these findings contribute powerful support for a general mammalian 'brain's eye' view in evolutionary psychology, as females have more oxytocin activity in their brains than males. In this context it is also noteworthy that female rats, just like their human counterparts, exhibit better social memories than males (BLUTHE/DANZER 1990).

6. Beauty, sexual attraction and mate selection

The manner in which facial expressions and bodily gestures help elicit, communicate, and regulate affective states is a central issue of both ethological and evolutionary psychological approaches to animal and human behavior. Although there are many cortical perceptual processes related to the detection of beauty, sexual attractiveness and reproductive fitness (for recent reviews, see JOHNSON 1999; BOROD 2000), a key issue is whether such perceptual mechanisms can operate independently of the socio–sexual circuits that exist subcortically. We simply do not know at present, but we suggest that the ability of various perceptions to regulate such desires and aversion (e.g., PERRETT, et al. 1998) is based very much on how they link up to subcortical socio–sexual circuits. (PFAFF 1999). We would be surprised if any higher perceptual 'modules' for social attraction in humans could operate effectively without connections to the types of basic subcortical emotional systems that have already been clarified in other species (PANKSEPP 1998a).

In this context, we also note the many recent animal studies that have highlighted how mate attraction may be mediated by signals of reproductive fitness (e.g., WELCH et al.,1998; GIL et al 1999). The inclusive-fitness based predictions that have been confirmed by population genetical and behavioral ecological points of view are truly remarkable (e.g., ELLEGREN et al. 1996; KOMDEUR 1997; Johnsen et al. 1998; SHELDON/ELLEGREN 1998). Although these studies highlight the fact that various species-typical perceptions are involved in mate choice within all species, we must again wonder whether these inputs operate independently of the subcortical socio–sexual motivational circuits that share a high similarity in all vertebrates. Although little is known about the brain mechanisms that mediate such adaptive socio–sexual strategies as noted above, we suspect that the primitive, genetically dictated affective brain systems are essential in all mammals (PANKSEPP 1998a).

7. Affect, sociopathy and the primordial self

We believe that studying how affective states are generated by the brain will be critical for understanding how evolution guides various behavioral strategies in humans and animals (PANKSEPP et al. 2000). The existence of various emotional and motivational feelings, along with general-purpose learning systems, can provide practically all the types of adaptive behavioral strategies that have been discussed by evolutionary psychologists (e.g., BUSS 1999). If this viewpoint is correct, then we have no alternative, but to make the study of affect a primary concern in evolutionary psychology. Although this principle appears to be recognized by the majority of investigators, in our estimation this cannot be achieved in any deep sense without assimilating evidence from brain research in related animals.

For instance, subcortical emotional systems are decisive in the way organisms spontaneously comport themselves socially. Sociability can be modulated powerfully by a great number of subcortical brains systems (NELSON/PANKSEPP 1998), with the most detailed analyses having been conducted on the differential expression of oxytocin and vasopressin systems in voles exhibiting very different social temperaments (for review, see INSEL 1997, 1998). It is expected that even human sociopathy (for review see MEALEY 1995), is modulated by similar systems, even though critical evidence for such issues is not yet available. Of course, there are bound to be gradients of sociopathy, some of the milder versions being simply based upon the normal variability of primitive emotional urges in our species coupled with deficient higher forms of consciousness. For instance, rape by human males, thought by some to be an adaptive human reproductive strategy (see THORNHILL/PALMER 2000), is more likely to be based on strong testosterone driven erotic and dominance urges in males, coupled with emotional insensitivity (e.g., deficient frontal lobe functions) and inadequate moral socialization. To the best of our knowledge, higher cerebral 'modules' have not evolved to promote such nefarious behavioral tendencies, even though one can envision how diminished frontal lobe functions (e.g., diminished guilt, shame, empathy and sympathy) could be reproductively advantageous in certain primitive environments. In short, it is easy to construct credible neuronal hypotheses of how tendencies toward social insensitivity, sexual coercion and agonistic tendencies could emerge from the interaction between basic emotional systems and general-purpose neocortical processors that are able to learn a vast number of context-dependent behavioral strategies.

However, to the best of our knowledge, the neocortex is not the source of affective experience (PANKSEPP 1998a, 2000b). Rather, raw emotional feelings emerge from ancient brain systems in ways that are just beginning to be understood in humans (HEATH 1996; DAMASIO, et al. 2000). For instance, a primordial form of 'self-representation' that can generate basic bodily expressions and emotional feelings was probably laid down deep in the brainstem. Of course, the full resolution of affective feelings (cerebral feelings and sentiments) surely emerges through the dynamic interactions of higher brain systems that regulate emotional states (DAMASIO 1999; PANKSEPP 1998a, 1998b). Evolutionary psychology needs to consider such possibilities, if it is to make substantive contributions to its fundamental 'mechanistic' concerns.

The Evolution of Behavior and Developmental Views in Psychology

All of what we have proposed now needs to be grounded in other key intellectual traditions. At somato–phenotypic levels, the power of genetic influences in guiding the construction of bodies and behavioral repertoires has long been affirmed by animal-husbandry practices. The heritability of temperament has also been confirmed by abundant work from behavioral-geneticists (SCOTT/FULLER 1965; FREEDMAN 1979; PLOMIN et al. 1997; HAMER/COPELAND 1998; SEGAL 1999). Furthermore, modern neuroscience has verified the existence of homologous brain mechanisms by its demonstration of abundant structural, neurochemical and functional similarities throughout the brains of all mammals (for overview PANKSEPP 1998a).

However, when we begin to speculate about the functional organization of the higher aspects of mind—the 'natural kinds' of the cortico–cognitive brain—we proceed at our own peril if we do not fully assimilate the *developmentalist challenge*—that structure-function relations do not simply emerge from DNA codes, but as much from the many interactions between genetic information, environmental information and the ontogenetic experiences of individuals (TINBERGEN 1972; OYAMA 1985/2000; GRIFFITHS 1997). All Evolutionary Psychological endeavors should recognize that genes do not directly control mind or behavior, but only the proteins and developmental patterns that help construct specific types of brains. Equally important is the recognition that genes and brains can only operate within environmental constraints (OYAMA 1985/2000). These stipulations will help temper radically reductionistic agendas in evolutionary thinking that simply cannot work. They are also a potential saving 'grace' for our apparent proclivity to misuse genetic knowledge.

Such assertions are not offered to simply acknowledge the inextricable roles of nature and nurture in all aspects of behavior, but to highlight the fact that apparent 'natural kinds' can be produced developmentally, as well as linguistically. Due to the intrinsic ambiguities that result from our impoverished knowledge about the developmental causes of individual differences in animals, the chasm from genes to final psychobiological product cannot be credibly bridged through any form of genetic determinism, especially in humans. At the same time, the power of genetic information as a key arbiter in the construction of specific brain/mind functions (at least within 'normal' environmental constraints) should not be underestimated. In this context, most of the half-truths of previous generations in claiming that there were 'genes for behavior', may in retrospect, be ascribed simply to sloppy, short-hand forms of communication that should have only been advanced with the proviso: the developmental processes operating during ontogenesis must always be given their due (TINBERGEN 1972; OYAMA 1985/2000). This should be repeated as a mantra whenever one begins to fall into the trap of ascribing any form of genetic determinism to psychological matters.

Such intellectual constraints must be heeded consistently in our currently immature brain/mind sciences, where the natural order of psychological processes has to be theoretically inferred rather than directly visualized. Because of such ambiguities, there has been a long-standing and persistent resistance in psychology and behavioral neuroscience to viewing the hidden brain/mind functions as inborn faculties. Learning, guided by still mysterious reinforcement-reward processes, has continued to be accepted as the main agent of behavioral control and plasticity. As a result, the issue of how many reinforcement-reward processes and related world-grasping mechanisms actually exist in the brain, not to mention whether they were internally experienced, remains largely undiscussed in both modern behavioral psychology and neuroscience. Now, partly because of the inroads made by evolutionary psychology (BARKOW et al. 1992; BUSS 1999) and Affective Neuroscience (PANKSEPP 1998a), there is a growing taste for such possibilities.

Since evolutionary psychology established relations with cognitive neuroscience (PINKER,1997), a few attempts at cross fertilization have emerged (e.g., TOOBY/COSMIDES 2000). However, even with the current revolution in functional brain imaging, which seems like a blessing for visualizing intrinsic brain functions (and hence, some believe, the historical record of past adaptations), we should proceed cautiously, continually constrained by converging evidence, rather than accepting direct genetic adaptations behind every consistent phenotypic trend or brain difference that is documented with new imaging technologies (TOGA/MAZZIOTTA 2000). Since evolutionary psychologists are likely to begin using such technology in the near future, they should appreciate the many problems that accompany such approaches.

At present, brain imaging only provides a low-order, ghostly image that something of importance may be transpiring somewhere in the brain. There are, no doubt, many 'false positives' and perhaps even more 'false negatives' due to the many technical constraints that exist in such studies. No form of human brain imaging, by itself, can assure us that what is being observed is a 'natural kind'—a brain/mind process that is strongly dictated by the way genetic factors help direct the function of certain brain circuits. Also, brain-imaging studies only provide neural correlates of psychological functions. Traditional animal studies, where brain systems are directly manipulated, as well as neuropsychological studies following cases of human brain damage and stimulation, provide better causal evidence of the functional characteristics within brain organization. Obviously, causal studies provide the best evidence upon which a scientifically sound evolutionary psychology can be built.

Some 'Prescriptions'

The aim of this paper was to critique some of the prominent views in evolutionary psychology, and to share an alternative perspective that is grounded in the ethological and comparative psychoneurological traditions that have long acknowledged the importance of natural selection in constructing the behavioral and psychological capacities of all animals. The current appeal and danger of evolutionary psychology arises from its unflinching willingness to utilize extreme modes of adaptationist thinking about the human mind, without critically distinguishing genetic adaptations from epigenetically emergent phenomena. Although the HAMILTONIAN (1964) concept of inclusive-fitness has turned out to be the most compelling evolutionary priniciple of the 20th century and an apt guide for ecological research on the social preferences of many species, those types of behavior patterns have not yet been shown to arise from genetically ingrained neural circuits in the higher regions of the human brain/mind.

We should not forget that in most mammals, including humans, social-bonds are learned to a great extent. Although social learning is based upon primitive social-emotional systems that all mammals share (PANKSEPP et al. 1978, 1980, 1984, 1985, 1988, 1994) and is surely under robust genetic control (SCOTT/FULLER 1965; PLOMIN et al. 1997; PANKSEPP/BURGDORF/GORDON 2001), everything that mature humans do is filtered through their higher neural capacity for flexible 'intelligent' action. Perhaps all too many individuals utilize those capacities largely for their own selfish ends, but it is important to emphasize that the *nature* of the higher regulatory systems in humans *does* permit many alternative courses of action. Although our psychological tendencies are tethered to ancient emotional concerns, we can individually entertain options of our own making. Also, there are group-dynamics operating in our lives and in the ongoing procession of evolution that have barely been envisioned (LUMSDEN/ WILSON 1981; SOBER/WILSON 1998). Animals exhibit non-genetic modes of inheritance (e.g., LIU et al. 2000), and such trans-generational effects provide a changing social fabric that has enormous implications for the way we conceptualize evolutionary canalizations. Surely, differential survival of groups may lead to differential survival of brain mechanisms that only operate efficiently in groups. Such effects may emerge more rapidly via group selection than by individual selection. At the very least, these issues remain more open, especially in humans, than some selfish gene advocates would like to believe.

With the emergence of high culture and high intelligence in our species, group selection may operate in human societies in ways that few have considered in the past few decades. We must remember that in humans, within the context of cultural evolution, biological evolution can move along much more rapidly than it does in animals who do not cognitively conceptualize their place in nature (CHAGNON/IRONS 1978; NEEL/WARD 1970). As Dan FREEDMAN, one of the first modern evolutionary psychologists, put it, "Sewall WRIGHT (1940)... consistently made the point that the same gene varies in its effect and action depending on the genes in its company... Wright has demonstrated mathematically that competition between individuals can be insignificant when compared with competition between families or competition between populations" (1979, p5). Although the classic 'group-selection' views of WYNN-EDWARDS (1962) have been widely held in disrepute (e.g., DAWKINS 1976), ever since HAMILTON (1964) introduced the compelling concept of inclusive-fitness, human mental and cultural evolution has in fact created extraordinary opportunities for alternative modes of evolutionary change. Modern genetic engineering is only the most blatant recent example of this claim. It could be argued that such possibilities are only little waves on the great sea of evolutionary progression, but perhaps it is a bit more in creatures like ourselves that have been endowed with a massive random-access type of general-purpose intelligence. To some yet unfathomed extent, we have been liberated from the crucible of a mindless biological emergence. This point is commonly acknowledged by many humanistically oriented scholars, but all too rarely by evolutionary psychologists.

At least for the human species, the excesses of selfish-gene type inclusive-fitness heuristics now need to be tempered by what we already know about mammalian and human brain/minds. Evolutionary psychology, along with the other variants of the discipline, have all too often neglected causal scientific understanding of the underlying brain processes shared by all mammals by continuing to prefer descriptions of surface processes which they believe are directly related to presumably recent human evolutionary issues. However, can we credibly utilize phases of recent human brain evolution that have left essentially no historical traces as a basis for our theorizing? Can imaginary evolutionary scenarios be used as a crystal ball to the real pre-historic EEA's that molded our brain/minds? To what extent should we consider tenuous 'modular' visions of human brain/mind evolution when more credible non-modular alternatives are already available? How do we adjudicate among the contending views? How do we incorporate the strikingly relevant evidence culled from our fellow animals into the exceedingly anthropocentric modes of thought that characterize so much of present day psychology? How shall we accept our animal heritage without demeaning our vast intelligence?

In our estimation, the type of psychological functions that evolutionary psychologists speak of, arise largely from the utilization of very old emotional capacities working in concert with newly evolved inductive abilities supported by the vast general purpose neocortical association areas. Although there are bound to be certain manifestations of emotional and motivational tendencies within these newly evolved regions of the human brain/mind, the massive modularity thesis entertained by evolutionary psychologists remains, except for certain well-accepted sensory-perceptual processes, far fetched and inconsistent with what we presently know about the higher reaches of the human brain/mind. It will rapidly become a circular discipline if it comes to rely exclusively on fMRI-type brain-imaging data for its neurological conjectures. What types of evolutionary engravings have, in fact, occurred in higher cortical areas remain anyone's guess. Processes such as the social use of our eyes, facial expressions, rhythmic gestures and prosodic intonations for instrumental purposes—perhaps even musical ability—are bound to be prominent dimensions of higher brain functions (e.g., EKMAN 1998; BOROD 2000; EMERY 2000), but such general issues are too rarely considered in mainstream evolutionary psychology.

It is good to see the discipline moving toward an eager confrontation with the brain (TOOBY/COSMIDES 2000), but it should pay attention to both higher and lower brain/mind functions, cultivating a research tradition that considers matters at all relevant levels. It is among the higher cortical functions, that learning, plasticity, and epigenetically emergent software functions prevail. Investigators should be especially cautious about concluding what was or was not constructed by natural selection during recent brain evolution within the human species. Likewise, it is at the cortical level that we cannot extrapolate readily from the animal data. Among those recently evolved brain areas there was abundant opportunity for massive evolutionary divergence (DEACON 1997). However, we should remember that those higher systems simply cannot work without the support of the basic subcortical systems we share with our fellow animals. If critical neuro–cognitive experiments eventually demonstrate the existence of genetically-dictated, special-purpose cognitive/affective modules in the higher reaches of the human brain-mind, all humans committed to an accurate description of our species' nature should be delighted to accept them into the pantheon of scientific evidence. However, as long as no compelling neurobehavioral evidence is provided, we should continue to regard distal adaptationist stories of human evolution, even when supported by some preliminary correlative evidence, as simply another aspect of the wondrous carnival of human fancy.

In our estimation, animal brain research will be more decisive in giving a clear scientific picture of how human social behaviors are guided by past evolutionary forces. What those vast cerebral expansions that emerged during the Pleistocene probably provided was a vast symbolic capacity that enabled foresight, hindsight, and the brain-power to peer into other minds and to entertain alternate courses of action, thereby allowing humans to create the cultures that dominate our modern world.

Conclusions

Why was serious scrutiny of mental processes based on the DARWINIAN framework so greatly delayed in the brain/mind sciences? Historically, it was partly due to the embarrassment of phrenology—where mental faculties were related, with the most trifling evidence, to the shape of cerebral structures inferred from craniometry. In a similar vein, we must now wonder whether modern brain imaging is also giving us many false conclusions (especially abundant false negatives) based on modest biophysical changes in cerebral blood flow and oxygenation. The visually impressive statistical maps of brain-imaging, with their islands of color suspended on ghostly MRI images of the brain, do not do justice to the actual neurodynamics that create mind. They provide only one approximation that must be supplemented by many other techniques.

With the possibility of real neuropsychological understanding, we may also need to question the potentially misleading 'saving grace' of the computational-computer revolution that has allowed investigators to readily create distorted digital views of psychological processes which are fundamentally analog in nature. As the search for heuristic computational algorithms came to be deemed a more worthy scientific enterprise among cognitive psychologists (CUMMINS/CUMMINS 2000), the possibility of a mature and fully substantive evolutionary psychology diminished. To discover what types of special-purpose adaptive functions evolution truly created within our brain/minds, there is no substitute for integrative cross-species brain research. We are only at the beginning of that grand intellectual journey.

A discipline that is as exceedingly anthropocentric as modern evolutionary psychology is likely to make a variety of preventable mistakes if it does not fully immerse itself in neurobiological, phylogenetic and ontogenetic issues. Our impression is that the most visible form of modern evolutionary psychology is currently pursuing a rash course of over-interpretation of the human condition, simply because the inclusive-fitness idea is so tremendously compelling.

Although we subscribe fully to the idea that natural selection gave rise to the rudiments of our fundamental neuro–mental apparatus, we are disappointed by the fact that current evolutionary psychological thought (not to mention that in most of psychology) is not guided by what we already know about the neurology of emotional and motivational processes in the brains of related animals. It is equally sad to see clear disregard for the likelihood that much of our neocortex is based upon general-purpose engineering principles rather than genetically guided modules. A top-down cognitive view may be deluded by the many emergent epigenetic specializations of the *adult* neocortex. The study of primitive emotional systems in other mammals provides a variety of parsimonious 'bottom up' tools for decoding the developing neural landscapes of higher mind functions in humans. Since the evolution of subcortical systems was probably guided by straightforward inclusive fitness issues, we anticipate that a great deal of our higher cognitive apparatus is still affected by similar influences, but certainly not in any inherently modularized ways.

What makes humans unique, perhaps more than anything else, is that we are a linguistically adept story-telling species. That is why so many different forms of mythology have captivated our cultural imaginations since the dawn of recorded history. Evolutionary psychologists also have many intriguing stories to tell, but if we are committed to a deep evolutionary view, their current speculations should not be accepted as credible foundations for our fundamental nature. The only massive cortical modules we should be convinced of at the present time are our vast linguistically based foresight and hindsight abilities, which mediate our compulsion to tell tales to each other. Incidentally, the basic urge to speak to each other may be closely linked to anterior cingulate and adjacent frontal lobe tissues which appear to mediate certain types of pain, feelings of separation distress and thereby social sensitivities (MACLEAN 1990; PANKSEPP 1998a; MAYBERG/MCGINNIS 2000; VOGT/DEVINSKY 2000; VOGT/SIKES 2000). Clearly, something very interesting is happening to the social-emotional realm in these brain tissues that are of great importance in controlling our motivation to communicate with each other.

Evolutionary psychology and sociobiology are such attractive scientific views (e.g., FREEDMAN 1979; SCOTT 1989; SEGAL et al. 1997) that they need to be carefully cultivated and constructed as accurately as possible, continually constrained by genetic and cross-species brain evidence from our fellow animals rather than by the sea-swell of imaginary neuropsychological possibilities in humans. If we continue to proceed without considering *all* the available evidence, we will only produce more of the polarized views that have been endemic to this troubled corner of evolutionary thought. Now that we have a real chance of bringing serious evolutionary views to the study human mind and behavior, we should proceed in as disciplined a manner as possible. If we do not pursue such reasonable courses of action, we may become mired in myth making rather than remaining on the shores of sound scientific inquiry.

Authors' address

Jaak Panksepp and Jules B. Panksepp
J. P. Scott Center for Neuroscience Mind and Behavior, Departments of Psychology and Biology, Bowling Green State University, Bowling Green, OH 43403.
Email: jpankse@bgnet.bgsu.edu
All correspondence to the first author.

Note

We dedicate this paper to the memory of our esteemed colleague John Paul SCOTT (1909–2000) who first coined the term 'sociobiology' for a series of joint sessions of the American Society of Zoologists and the Ecological Society of America held initially at the AAAS meeting of 1946—a series that eventually led to the organization of the American Animal Behavior Society. Paul's research was the first to systematically analyze the heritability of behavioral tendencies in a mammalian species, specifically, temperamental differences among different strains of domesticated dogs (see SCOTT/FULLER 1965). We appreciate the realistic tradition of rigorous and respectfully complex evolutionary scholarship he promoted at BGSU and around the world.

References

Barkow, J./Cosmides, L./Tooby, J. (eds) (1992) The Adapted Mind. Oxford Univiversity Press: New York.

Betzig, L. (ed) (1997) Human Nature: A Critical Reader. Oxford University Press: New York.

Bittar, E. A./Bittar, N. (2000) Biological psychiatry. JAI Press: Stamford CT.

Bluthe, R. M./Danzer, R. (1990) Social recognition does not involve vasopressinergic neurotransmission in female rats. Brain Research 535: 301–304.

Borod, J. C. (2000) The Neuropsychology of Emotion. Oxford University Press: New York.

Brown, D. E. (1991) Human Universals. New York: McGraw-Hill.

Buck, R. (1999) The biological affects: A typology. Psychological Review 106: 301–336.

Budiansky, S. (1998) If a Lion Could Talk: Animal Intelligence and the Evolution of Consciousness. The Free Press: New York.

Burgdorf, J./Knutson, B./Panksepp, J. (2000) Anticipation of rewarding brain stimulation evokes ultrasonic vocalizations in rats. Behavioral Neuroscience 114: 1–8.

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

Buss, D. M. (2000) The evolution of happiness. American Psychologist 55: 15–23.

Buss, D. M./Haselton, M. G./Shackefortd, T. K./Bleske, A. L./Wakefield, J. C. (1998) Adaptations, exaptations, and spandrels. American Psychologist 53: 533–548.

Campbell, J. (1982) Grammatical Man: Information, Entropy, Language and Life. Simon and Schuster: New York.

Carter, C. S. (1998) Neuroendocrine perspectives on social attachment and love. Psychoneuroendocrinology 23: 779–818.

Carter, C. S./Lederhendler, I./Kirkpatrick, B. (eds) (1999) The Integrative Neurobiology of Affiliation. MIT Press: Cambridge MA.

Chagnon, N. A./Irons, W. (eds) (1978) Evolutionary Biology and Human Social Behavior. Duxbury: North Scituate MA.

Chen, C./Rannie, D. G./Greene, R. W./Tonegawa, S. (1994) Abnormal fear response and aggressive behavior in mutant mice deficient for alpha-calcium-calmodulin kinase II. Science 266: 291–294.

Clark, A. (1997) Being There: Putting Brain, Body, and World Together Again. MIT Press: Cambridge MA.

Cosmides, L./Tooby, J. (2000) Evolutionary Psychology and the Emotions. In: Lewis, M./Haviland, J. (eds) The Handbook of Emotions, 2nd edition. New York: Guilford, pp. 91–116.

Cummins, R./Cummins, D. D. (eds) (2000) Minds, Brains, and Computers: The Foundations of Cognitive Science. Blackwell Publishers: Malden MA.

D'Amato, F. R. (1998) Kin interaction enhances morphine analgesia in male mice. Behavioral Pharmacology 9: 369–373.

Damasio, A. R. (1999) The Feeling of What Happens: Body and Emotion in the Making of Consciousness. Harcourt Brace: New York.

Damasio, A. R./Grabowski, T. J./Bechara, A./Damasio, H./ Ponto, L. B./Parvizi, J./Hichwa, R. D. (2000) Subcortical and cortical brain activity during the feeling of self-generated emotions. Nature Neuroscience 3: 1049–1056.

Darwin, C. (1872/1998) The Expression of the Emotions in Man and Animals, 3rd edition. Oxford University Press: New York.

Dawkins, R. (1976) The Selfish Gene. Oxford University Press: Oxford.

Deacon, T. W. (1997) The Symbolic Species: The Co-Evolution of Language and the Brain. Norton: New York.

Deacon, T. W. (1990) Rethinking mammalian brain evolution. American Zoologist 30: 629–705.

Ekman, P. (1998) Universality of emotional expression? A personal history of the dispute. In: Darwin, C. (1998) The Expression of the Emotions in Man and Animals, 3rd editon. Oxford University Press: New York, pp. 363–393.

Ellegren, H./Gustafsson, L./Sheldon, B. C. (1996) Sex ratio adjustment to paternal attractiveness in a wild bird population. Proceedings of the National Academy of Science USA 93: 11723–11728.

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

Ferguson, J. N./Young, L. J./Hearn, E. F./Matzuk, M. M./Insel, T. R./Winslow, J. T. (2000) Social amnesia in mice lacking the oxytocin gene. Nature Genetics 25: 284 – 288.

Finlay, B. L./Darlington, R. B. (1995) Linked regularities in the development and evolution of mammalian brains. Science 268: 1578–1584.

Fischer, H./Andersson, J. L. R./Furmark, T./Fredrikson, M. (2000) Fear conditioning and brain activity: A positron emission tomography study, Behavioral Neuroscience 114: 671–680.

Fodor, J. (1983) The Modularity of Mind. MIT Press: Cambridge MA.

Freedman, D. J. (1979) Human Sociobiology: A Holistic Approach. The Free Press: New York.

Freeman, W. J. (1999) How the Brain Makes Up its Mind. Blackwell: London.

Fridja, N. H. (1986) The Emotions. Cambridge University Press: Cambridge UK.

Gallese, V./Goldman, A. (1998) Mirror neurons and the simulation theory of mind-reading. Trends in Cognitive Sciences 2: 493–501.

Gil, D./Graves, J./Hazon, N./Wells, A. (1999) Male attractiveness and differential testosterone investment in zebra finch eggs. Science 286: 126–128.

Godfrey-Smith, P. (1996) Complexity and the Function of Mind in Nature. Cambridge University Press: New York.

Gould, S. J. (1991) Exaptation: a crucial tool for evolutionary psychology. Journal of Social Issues 47: 43–65.

Gould, S. J./Lewontin, R. C. (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proceedings of the Royal Society of London B 205: 581–598.

Graeff, F. G. (1994) Neuroanatomy and neurotransmitter regulation of defensive behaviors and related emotions in mammals. Brazilian Journal of Medical and Biological Research 27: 811–829.

Griffiths, P. E. (1997) What Emotions Really Are: The Problem of Psychological Categories. University of Chicago Press: Chicago.

Hamer, D. H. (1998) Living with our genes: Why They Matter More Than You Think. Doubleday: New York.

Hamilton, W. D. (1964) The genetical evolution of social behaviour (I and II). Journal of Theoretical Biology 7: 1–52.

Hauser, M. D. (2000) Wild Minds: What Animals Really Think. Henry Holt & Co: New York.

Heath, R. G. (1996) Exploring the mind-brain relationship. Moran Printing Inc: Baton Rouge LA.

Heinrich, R, Cromarty S. I./Horner, M./Edwards, D. H/Kravitz, E. A. (1999) Autoinhibition of serotonin cells: an intrinsic regulator sensitive to the pattern of usage of the cells. Proceedings of the National Academy of Sciences USA 96: 2473–2478.

Huber, R./Delago, A. (1998) Serotonin alters decisions to withdraw in fighting crayfish, Astacus astacus: the motivational concept revisited. Journal of Comparative Physiology A 182: 573–583.

Ikemoto, S./Panksepp, J. (1999) The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. Brain Research Reviews 31: 6–41.

Insel, T. (1997) The neurobiology of social attachment. American Journal of Psychiatry 154: 726–735.

Johnsen, A./Lifjeld, J. T./Rohde, P. A./Primmer, C. R./Ellegren, H. (1998) Sexual conflict over fertilizations: female bluethroats escape male paternity guards. Behav. Ecol. Sociobiol. 43:401–408.

Johnsen, A./Andersson, S./Örnborg, J./Lifjeld, J. T. (1998) Ultraviolet plumage ornamentation affects social mate choice and sperm competition in the bluethroat (Luscinia s.svecica): a field experiment. Proceedings of the Royal Society of London B 265: 1313–1318.

Johnson, V. S. (1999) Why We Feel: The Science of Human Emotions. Perseus Book Group: New York.

Kanwisher, N. (2000) Domain specificity in face perception. Nature Neuroscience 3: 759–763.

Keverne, E.B./Fundele, R./Narasimha, M./Barton, S. C./Surani, M. A. (1996) Genomic imprinting and the differential roles of parental genomes in brain development. Developmental Brain Research 92: 91–100.

Kim, J. (1993) Supervenience and Mind: Selected Philosophical Essays. Cambridge University Press: New York.

Knutson, B./Wolkowitz, O. M./Cole, S. W./Chan, T./Moore, E. A./Johnson, R. C./Terpstra, J./Turner, R. A./Reus, V. I. (1998) Selective alteration of personality and social behavior by serotonergic intervention. American Journal of Psychiatry 155: 373–379.

Komdeur (1997) Extreme adaptive modification in sex ratio of the Seychelles warbler's. Nature 385: 522–525.

Kraemer, P. D. (1993) Listening to Prozac. Viking: New York.

Kravitz, E. A. (1988) Hormonal control of behavior: amines and the biasing of behavioral output in lobsters. Science 241: 1775–1781.

Kravitz, E. A. (2000) Serotonin and aggression: insights gained from a lobster model system and speculations on the role of amine neurons in a complex behavior. Journal of Comparative Physiology A 186: 221–238.

LeDoux, J. (1996) The Emotional Brain. Simon & Schuster: New York.

Lewis, M. D./Granic, I. (eds) (2000) Emotion, Self-Organization, and Development. New York: Cambridge Univ. Press.

Liu, D./Diorio, J./Day, J. C./Francis, D. D./Meaney, M. J. (2000) Maternal care, hippocampal synaptogenesis and cognitive development in rats. Nature Neuroscience 3: 799–806.

Lumsden, C. J./Wilson, E. O. (1981) Genes, Mind, and Culture: The Coevolutionary Process. Harvard University Press: Cambridge MA.

MacLean, P. D. (1990) The Triune Brain in Evolution: Role in Paleocerebral Functions. Plenum Press: New York.

Mauk, M. D. (2000) The potential effectiveness of simulations versus phenomenological models. Nature Neuroscience 3: 649–651,.

Mayberg, H. S./McGinnis, S. (2000) Brain mapping: the application, mood and emotions. In: Toga, A. W./Mazziotta, J. C. (eds) Brain Mapping :The Systems. Academic Press: San Diego CA, pp. 491–522.

Mealey, L. (1995) The sociobiology of sociopathy: an integrated evolutionary model. Behavioral and Brain Sciences 18: 523–599.

McGuire, M. T./Troisi, A. (1998) Darwinian Psychiatry. Oxford University Press: New York.

Neel, J. V./Ward, R. H. (1970) Village and tribal genetic distances among american indians, and the possible implications for human evolution. Proceedings of the National Academy of Sciences USA 65: 323–330.

Nelson, E./Panksepp, J. (1998) Brain substrates of infant-mother attachment: Contributions of opioids, oxytocin, and norepinepherine. Neuroscience & Biobehavioral Reviews 22: 437–452.

Nesse, R. M./Williams, G. C. (1974) Why We Get Sick. Times Books Random House: New York.

Nesse, R. M. (1990) Evolutionary explanations of emotions. Human Nature 1: 261–289.

Olmstead, M. C./Franklin, K. B. (1997) The development of a conditioned place preference to morphine: effects of microinjections into various CNS sites. Behavioral Neuroscience 111: 1324–1334.

Oyama, S. (1985/2000) The Ontogeny of Information: Developmental Systems and Evolution. Cambridge, UK: Cambridge University Press (reprinted in 2000 by Duke University Press).

Pani, L. (2000) Is there an evolutionary mismatch between the normal physiology of the human dopaminergic system and current environmental conditions in industrialized countries? Molecular Psychiatry 5: 467–475.

Panksepp, J. (1981) Hypothalamic integration of behavior: rewards, punishments, and related psychobiological process. In: Morgane, P. J./Panksepp, J. (eds) Handbook of The Hypothalamus, Vol. 3, Part A. Behavioral Studies of the Hypothalamus. Marcel Dekker: New York, pp. 289–487.

Panksepp, J. (1982) Toward a general psychobiological theory of emotions. The Behavioral and Brain Sciences 5: 407–467.

Panksepp, J. (1986a) The anatomy of emotions. In: Plutchik, R. (ed) Emotion: Theory, Research and Experience Vol. III. Biological Foundations of Emotions. Academic Press: Orlando FL, pp. 91–124.

Panksepp, J. (1986b) The neurochemistry of behavior. Annual Review of Psychology 37: 77–107.

Panksepp, J. (1989) The psychobiology of emotions: the animal side of human feelings. In: Gainotti, G./Caltagirone, C. (eds) Emotions and the Dual Brain. Experimental Brain Research, Series 18, pp. 31–55. Berlin: Springer-Verlag.

Panksepp, J. (1989/1999) Altruism and helping behaviors, neurobiology. In: Edelman, G. (ed) The Encyclopedia of Neuroscience. Boston: Birkhauser, pp. 55–57.

Panksepp, J. (1990a) The psychoneurology of fear: evolutionary perspectives and the role of animal models in understanding anxiety. In: Burrows, G. D./Roth, M./Noyes Jr., R. (eds) Handbook of Anxiety, Vol. 3: The Neurobiology of Anxiety. Elsevier: Amsterdam, pp. 3–58.

Panksepp, J. (1990b) Can mind and behavior be understood without understanding the brain?: a response to Bunge. New Ideas in Psychology 8: 139–149.

Panksepp, J. (1990c) Gray zones at the emotion-cognition interface: a commentary. Cognition and Emotion 4: 289–302.

Panksepp, J. (1992) A critical role for ‘affective neuroscience’ in resolving what is basic about basic emotions. Psychological Review 99: 554–560.

Panksepp, J. (1993a) Neurochemical control of moods and emotions: amino acids to neuropeptides. In: Lewis, M./Haviland, J. (eds) The Handbook of Emotions. Guilford: New York, pp. 87–107.

Panksepp, J. (1993b) Rough-and-tumble play: a fundamental brain process. In: MacDonald, K. B. (ed) Parents and Children Playing. SUNY Press: Albany NY, pp. 147–184.

Panksepp, J. (1998a) Affective Neuroscience: The Foundations of Human and Animal Emotion. Oxford University Press: New York.

Panksepp, J. (1998b) The periconscious substrates of consciousness: affective states and the evolutionary origins of the SELF. Journal of Consciousness Studies 5: 566–582.

Panksepp, J. (1998c) Attention deficit hyperactivity disorders, psychostimulants, and inteolerance of childhood playfulness: a tragedy in the making? Current Directions in Psychological Sciences 7: 91–98.

Panksepp, J. (2000a) Emotions as natural kinds within the mammalian brain. In: Lewis, M./Haviland, J. (eds) The Handbook of Emotions, 2nd edition. Guilford: New York, pp. 137–156.

Panksepp, J. (2000b) Affective consciousness and the instinctual motor system: the neural sources of sadness and joy. In: Ellis, R./Newton, N. (eds) The Caldron of Consciousness: Motivation, Affect and Self-organization, Advances in Consciousness Research. Amsterdam: John Benjamins Pub. Co, pp. 27–54.

Panksepp, J. (2000c) The neurodynamics of emotions: an evolutionary-neurodevelopmental view. In: Lewis, M. D./Granic, I. (eds) Emotion, Self-Organization, and Development, pp. 236–264. New York: Cambridge Univ. Press.

Panksepp, J. (2000d) The neuro-evolutionary cusp between emotions and cognitions: implications for understand consciousness and the emergence of a unified mind science. Consciousness & Emotion 1: 17–56.

Panksepp, J. (2000e) The riddle of laughter: neural and psychoevolutionary underpinnings of joy. Current Directions in Psychological Sciences 9: 183–186.

Panksepp, J. (2001) The long-term psychobiological consequences of infant emotions: prescriptions for the 21st century. Infant Mental Health Journal (in press).

Panksepp, J./Burgdorf, J. (1999) Laughing rats? Playful tickling arouses high frequency ultrasonic chirping in young rodents. In: Hameroff, S./Chalmers, D./Kazniak, A., Toward a Science of Consciousness III. Cambridge MA: MIT Press, pp. 231–244.

Panksepp, J./Burgdorf, J. (2000) 50k-Hz chirping (laughter?) in response to conditioned and unconditioned tickle-induced reward in rats: effects of social housing and genetic variables. Behavioral Brain Research 115: 25–38.

Panksepp, J./Burgdorf, J./Gordon, N. (2001) Toward a genetics of joy: Breeding rats for 'laughter'. In: Kazniak, A. (ed) Emotions, Qualia, and Consciousness. World Scientific: London (in press).

Panksepp, J./Herman, B./Conner, R./Bishop, P./Scott, J. P. (1978) The biology of social attachments: opiates alleviate separation distress. Biological Psychiatry 9: 213–220.

Panksepp, J./Herman, B. H./Villberg, T./Bishop, P./DeEskinazi, F. G. (1980) Endogenous opioids and social behavior. Neuroscience and Biobehavioral Reviews 4: 473–487.

Panksepp, J./Knuston, B./Burgdorf, J. (2001) The role of emotional brain systems in addictions: A neuro-evolutionary perspective. Addiction (in Press).

Panksepp, J./Nelson, E./Siviy, S. (1994) Brain opioids and mother-infant social motivation. Acta Paediatrica 397: 40–46.

Panksepp, J./Normansell, L. A./Herman, B. Bishop, P./Crepeau, L. (1988) Neural and neurochemical control of the separation distress call. In: Newman, J. D. (ed) The Physiological Control of Mammalian Vocalization. Plenum Press: New York, pp. 263–299.

Panksepp, J./Siviy, S./Normansell, L. (1984) The psychobiology of play: theoretical and methodological perspectives. Neuroscience and Biobehavioral Reviews 8: 465–492.

Panksepp, J./Siviy, S. M./Normansell, L. A. (1985) Brain opioids and social emotions. In: Reite, M./Fields, T. (eds) The Psychobiology of Attachment and Separation. Academic Press: New York, pp. 3–49.

Parrot, W. G./Schulkin, J. (1993) Neuropsychology and the cognitive nature of emotions. Cognition and Emotion 7: 43–59.

Penton-Voak, I. S./Perrett, D. I. (2000) Female preference for male faces changes cyclically: Further evidence. Evolution and Human Behavior 21: 39–48.

Perrett, D. I./Lee, K. J./Penton-Voak, I./Rowland, D./Yoshikawa, S./Burt, D. M./Henzl, S. P./Castles, D. L./Akamatsu, S. (1998) Effects of sexual dimorphism on facial attractiveness. Nature 394: 884–887.

Pfaff, D. W. (1999) Drive: Neurobiological and Molecular Mechanisms of Sexual Motivation. MIT Press: Cambridge MA.

Pinker, S. (1994) The Language Instinct. Morrow: New York.

Pinker, S. (1997) How the Mind Works. Norton: New York.

Plomin, R. (1999) Genetics and general cognitive ability. Nature, Vol. 402 (Suppl): C25–29.

Plomin, R./DeFries, J. C./McClearn, G. E./Rutter, M. (1997) Behavior Genetics, 3rd edition. New York: W.H. Freeman and Co.

Rizzolatti, G./Fadiga, L./Gallese, V./Fogassi, L. (1996) Premotor cortex and the recognition of motor actions. Brain Research and Cognition 3: 131–141.

Robins, R. W./Gosling, S. D./Craik, K. H. (1999) An empirical analysis of trends in psychology. American Psychologist 54: 117–128.

Robinson, T. E./Berridge, K. (1993) The neural basis of drug craving: an incentive-sensitization theory of addiction. Brain Research Reviews 18: 247–291.

Rolls, E. T. (1999) The Brain and Emotion. Oxford University Press: Oxford UK.

Rosen, J. B./Schulkin, J. (1998) From normal fear to pathological anxiety. Psychological Review 104: 325–350.

Salmon, C./Daly, M. (1996) On the importance of kin relations to Canadian women and men. Ethology and Sociobiology 17: 289–297.

Samuels, R. (1998) Evolutionary psychology and the massive modularity hypothseis. The British Journal for the Philosophy of Science 49: 575–602.

Schaffner, K. F. (1998) Genes, behavior, and developmental emergentism: one process, indivisible? Philosophy of Science 65: 209–252.

Scheibel, A. B./Schopf, J. W. (eds) (1997) The Origin and Evolution of Intelligence. Jones and Bartlett: Boston.

Scott, J. P./Fuller, J. L. (1965) Genetics and the Social Behavior of the Dog. University of Chicago Press: Chicago.

Scott, J. P. (1989) The Evolution of Social Systems. Gordon and Breach Science Publishers: New York.

Segal, N. L./Weisfeld, G. E./Weisfeld, C. C. (eds) (1997) Uniting Psychology and Biology: Integrative Perspectives on Human Development. American Psychological Association: Washington DC.

Segal, N. L. (1999) Entwined Lives: Twins and What They Tell Us About Human Behavior. Dutton: New York.

Sheldon, B. C./Ellegren, H. (1998) Paternal effort related to experimentally manipulated paternity of male collared flycatchers. Proceedings of the Royal Society of London 265: 1737–1742.

Shand, A. F. (1920) The Foundations of Character: Being a Study of the Tendencies of the Emotions and Sentiments. MacMillan and Co: London.

Silverman, I./Choi, J./Mackewn, A./Fisher, M. Moro, J./Olshansky, E. (2000) Evolved mechanisms underlying wayfinding: further studies on the hunter-gatherer theory of spatial sex differences. Evolution and Human Behavior 21: 201–213.

Sober, E./Wilson, D. S. (1998) Unto Others. Harvard University Press: Cambridge MA.

Spence, S. A./Frith, C. D. (1999) Towards a functional anatomy of volition. Journal of Consciousness Studies 6: 11–30.

Stevens, A./Price, J. (1996) Evolutionary Psychiatry. Routledge: London.

Thornhill, R./Palmer, C. T. (2000) A Natural History of Rape: Biological Bases of Sexual Coercion. MIT Press: Cambridge MA.

Tinbergen, N. (1972) Functional ethology and the human sciences. Proceedings of the Royal Society of London 182 B: 385–410.

Toga, A. W./Mazziotta, J. C. (eds) (2000) Brain Mapping: The Systems. Academic Press: San Diego CA.

Tomasello, M. (1999) The Cultural Origins of Human Cognition. Harvard University Press: Cambridge MA.

Tooby, J./Cosmides, L. (2000) Toward mapping the evolved functional organization of mind and brain. In: Gazzaniga, M. S. (ed) The New Cognitive Neuroscience (2nd edition). MIT Press: Cambridge MA, pp. 1167–1178.

Vanderhaeghen, P./Lu, Q./Prakash, N./Frisen, J./Walsh, C. A./Frostig, R. D./Flanagan, J. G. (2000) A mapping label requaired for normal scale of body representation in the cortex. Nature Neuroscience 3: 358–365.

Vogt, B. A./Devinsky, O. (2000) Topography and relationships of mind and brain. Progress in Brain Research 122: 11–22.

Vogt, B. A./Sikes, R. W. (2000) The medial pain system, cingulate cortex, and parallel processing of nociceptive information. Progress in Brain Research 122: 223–235.

Winslow, J. T./Hastings, N./Carter, C. S./Harbaugh, C. R./Insel, T. R. (1993) A role for central vasopressin in pair bonding in monagamous prairie voles. Nature 365: 544–548.

Welch, A. M./Semlitsch, R. D./Gerhardt, H. C. (1998) Call duration as an indicator of genetic quality in male gray tree frogs. Science 280: 1928–1930.

Weisfeld, G. E. (1997) Discrete emotions theory with specific reference to pride and shame. In: Segal, N.L./Weisfeld, G.E./Weisfeld, C.C. (eds) Uniting Psychology and Biology: Integrative Perspectives on Human Development, pp. 419–443. American Psychological Association: Washington DC.

Wilson, E. O. (1975) Sociobiology. Cambridge MA: Harvard Univ. Press.

Wimmer, M. (1995) Evolutionary roots of emotions. Evolution and Cognition 1: 38–50.

Wimmer, M./Ciompi, L. (1995) Evolutionary aspects of affective-cognitive interactions in the light of Ciompi's concept of 'Affect-Logic'. Evolution and Cognition 2: 37–58.

Winslow, J. T./Hastings, N./Carter, C. S./Harbaugh, C. R./Insel, T. R. (1993) A role for central vasopressin in pair bonding in monogamous praerie voles. Nature 365: 544–548.

Wright, R. (1994) The Moral Animal. Academic Press: New York.

Wright, S. (1940) The statistical consequences of Mendelian heredity in relation to speciation. In: Huxley, J. (ed) The New Systematics. Oxford University Press: New York.

Wynn-Edwards, V. C. (1962) Animal Dispersion in Relation to Social Behavior. Edinburg: Oliver/Boyd.

Young, L. J./Nilsen, R./Waymire, K. G./MacGregor, G. R./Insel, T. R. (1999) Increased affiliative response to vasopressin in mice expressing the V1a receptor from a monogamous vole. Nature 400: 766–768.

Zajonc, R. B. (2000) Feeling and thinking: closing the debate over the independence of affect. In: Forgas, J. P. (ed) Feeling and Thinking: The Role of Affect in Social Cognition. Cambridge University Press: New York, pp. 31–58.

Gordon G. Brittan

The Inevitability of Anthropomorphism

Abstract

Students of animal behavior are told to shun 'anthropomorphism'. This stern warning is not unambiguous. I will understand 'anthropomorphism' as the policy of attributing beliefs and desires to other animals, in the attempt to explain their behavior. This particular policy is to be avoided, so we are told, not simply because there are no reasonable grounds on which to attribute beliefs and desires to other animals, but also because such attribution is not, even in our own case, predictive. 'Anthropomorphism' is thus not 'scientific'.

Philosophers have a twofold burden, to demonstrate both the possibility and the necessity of the concepts they want to apply. I argue that it is possible reasonably to attribute beliefs and desires (and 'rationality') to other animals, that behavioral and functional accounts (which dispense with intentionality) of animal behavior are not entirely adequate (since they cannot explain failures in particular routines), and that there are underlying reasons why this should be so, i.e., that there are reasons why we cannot dispense with 'folk psychology' in our attempts to understand animals, which is to say that animals are not machines, but in fact very much like us.

Key words

Anthropomorphism, intentionality, animal behavior, functional description.

Serious students of animal behavior are invariably told, in no uncertain terms, to shun 'anthropomorphism'. As a first approximation, 'anthropomorphism' may be characterized as the application to animals of a vocabulary typically and primarily applied to human beings. To shun its use is, in particular, to avoid attributing anything like human thoughts, purposes, feelings, and customs to other animals in the description and explanation of their behavior.

Two considerations led me, early on, to be skeptical about this advice. One is that those who offer it almost always revert within several pages to talk of aggression, altruism, courtship and other thoroughly anthropomorphic ideas.[1] Indeed, their habitual use of cost-benefit analysis and similar notions to describe and explain animal behavior betrays not merely a kind of anthropomorphism, the postulation of (ostensibly unconscious) 'fitness calculators' and optimizers, but a degree of cultural chauvinism as well, an apparently unquestioned belief in the efficacy of free-market mechanisms.

The other consideration is that successful animal trainers are more often than not those who are frankly 'anthropomorphic', understanding dogs and horses in something very like human terms. (COREN 1995, chapter 4) Those who shun this way of working with animals tend to make slow progress.

Of course, neither of these considerations is conclusive. It could be held that reversion to anthropomorphic language is simply a function of the fact that most people are very familiar with it. If one wants to be understood, particularly by a lay audience, there really is no option to its use. But in principle, and among professionals, it is dispensable, a mere *facon de parler*. Similarly, the apparent success of 'anthropomorphic' animal trainers cannot stand close scrutiny. For it is highly probable that those who understand animals in human terms treat them more considerately as well, reinforcing desired behavior in subtler and more effective ways than those who do not.

Still, these considerations should lead us to look more closely at the reasons why anthropomorphism is supposed to be bad policy. We can begin with two. In fact, I don't think that they are very good reasons.

One reason is that anthropomorphism is ill-founded; the kinds of claims that it involves go well beyond the evidence for them. We know the kinds of claims: horses that can do arithmetic, parrots that chide other birds with "you naughty Polly" for getting out of their cages and eating apples on the kitchen table, elephants that "take delight in love and honor, and also possess a rare thing amongst men—honesty, self-control, and a sense of justice".[2] Whenever I scold one of our dogs, my wife invariably reports that it feels guilty and that I should be able to tell by its expression that it won't do the same thing again.

These claims are difficult to take seriously. But our reluctance is not based so much on the fact that they are anthropomorphic in character as that they are anecdotal, unreproducible in anything like controlled conditions, and more often than not sentimental. Most importantly, they do not rest on anything like a *pattern* of behavior. They are in these respects exactly like the claims of grandparents concerning their grand-children; my mother continues to insist that our son waved her good-bye with a cheery "see you later" at the age of six months, his tone hinting that she shouldn't tell us just yet that he already understood everything we were saying. But grandparents are not the only ones who 'read in' a great deal more than the evidence warrants. Max RAPHAEL, in his book *Prehistoric Cave Paintings* (1945), sees in the ice age images he describes "the first conception of the *Liebestod*", "the first idea of catharsis", and something like "EINSTEIN's discovery of the dependence of mathematics upon the electromagnetic and gravitational fields". (GOMBRICH 1996, p10) In short, it is not anthropomorphic claims *per se*, but a common and (as we shall see) explicable tendency to make them and others like them in under-supported ways that is at fault.

A second reason why anthropomorphism is alleged to be bad policy is that it is misleading; the kinds of claims that it makes lead to faulty inferences. In *Vaulting Ambition* (1985), Philip KITCHER aligns himself with those critics who view "anthropomorphism as the original sin of pop sociobiology" (pp184–185). KITCHER's argument is that when we apply words like 'monogamy' and 'rape' to animal behavior we are apt to draw the wrong conclusions. There is undoubtedly something to this, particularly when we find cases of 'rape' or 'polygamy' among other species and conclude without further apology that it is 'natural' or somehow conducive to 'fitness', and start looking for its genetic basis. But his discussion suggests that the problem has to do with the often ill-defined and emotive character of these terms more than it has to do with their application to other animals. Exactly the same sorts of wrong conclusions are apt to be drawn when we apply such terms (think of 'homosexuality') to human behavior. The lesson is that these terms should be handled very carefully, not that they cannot or should not be applied either to other animals or to ourselves, even when we are doing serious science.[3]

The point needs to be stressed. Many of the reasons given for shunning the application of an 'anthropomorphic' vocabulary to animal behavior tell equally against its application to human behavior. Certainly this seems to be true of the reasons discussed so far. One could infer either that a *reductio* of the anti-anthropomorphic position had been demonstrated or that even in the case of human behavior 'anthropomorphism' must go.

The latter horn of the dilemma is, in fact, the one that some contemporary students of human and animal behavior would like to embrace. Thus the first page of MCFARLAND/BOSSER (1993):

"Our reasoning is derived partly from current thinking in biology and partly from a distaste for anthropomorphism (the tendency to assume that our minds work the way we think they work and to attribute such workings to other agents, both animate and inanimate). In some respects, anthropomorphism is an incurable disease, because we are probably incapable—for good evolutionary reasons—of resisting the temptation to interpret the behavior of other agents in terms of our own subjective experiences. Nevertheless, as scientists we should recognize that our introspection is probably not a good guide to reality, and we should make some attempt to break the spell."

At least the following claims are made or implied in this passage.

First, anthropomorphism is the view that behavior, our own and others, is to be described and explained in terms of the beliefs and desires of agents, a practice sometimes referred to as 'folk psychology'. We can call such description and explanation 'intentional'. 'Anthropomorphism' thus defined involves the use of an intentional vocabulary.

Second, it is a mistake to use an intentional vocabulary in the description of either human or animal behavior. 'Folk psychology' provides us with an adequate theory of neither. The reason given here is that the attribution of intentions is based on introspection, and introspection is unreliable. For one thing, it is impossible to verify the introspective reports of others; they are not, in the usual phrase, 'inter-subjectively testable'. Even our own introspective awareness is not infallible; it must often be revised in the light of our own subsequent behavior. For another thing, introspective reports, directly verifiable or not, are not predictive with respect to behavior. There are no *laws* connecting the beliefs and desires they report with subsequent actions. But inter-subjective testability and predictability are the hallmarks of science. Hence, 'folk psychology', and the anthropomorphism with which it is closely linked, are not scientific. They must therefore be shunned.

Third, there is a much better, 'biological' way in which to describe and explain human behavior, in terms of selection pressures and adaptation. Anthropomorphism is what KANT might have called a 'natural disposition', like other such dispositions to be given an evolutionary explanation, but it is not the inter-subjectively testable and predictive 'science' that biology is.

In brief, MCFARLAND and BOSSER argue a double thesis: anthropomorphism (as characterized) is not possible, at least not as a scientifically credible practice, and it is not necessary, since a perfectly adequate (biological) alternative is available. My task is to show that this thesis is false.

The first point to note is that we unreflectively attribute beliefs and desires in order to make sense of certain kinds of behavior. Such attribution has nothing to do, initially, with introspection or, in the usual meaning of the expression, 'subjective experiences'. It is enough that certain observable conditions are met.

Grizzlies, unique among bears in this respect, routinely cache food. We explain an individual's returning to its cache and eating what was stored there by saying that it was hungry, that is, desired food, and believed, on the basis of what it remembered, that it would find it at just that spot. This sort of explanation is appropriate just in case the behavior is not coerced or prompted, not accidental (the bear did not simply stumble across the cache), not automatic (the bear does not return to this particular place at regular intervals, whether or not food is there). But we can tell, in a perfectly objective way, whether or not the behavior is coerced or prompted, that is, we can tell whether or not the bear was pushed or cued to that spot, just as we can tell whether it returns merely coincidentally or as a function of habit. Occasionally, as in the case of unmasking Clever Hans, the latter requires setting up sophisticated experiments. It is precisely the failure to screen out, or attempt to screen out, the coerced or prompted, the accidental, and the automatic, that renders what might be called 'anecdotal anthropomorphism' (the arithmetical horses and honest elephants) so suspect.

It is simply gratuitous to ask, but are the animal's beliefs and desires just like mine, 'subjectively', when I head for the refrigerator, for there is no more to beliefs and desires, at this stage, than the explanatory roles they play. These roles do not require any 'felt' or phenomenological quality, nor does the animal have to be aware, self-consciously, that it desires food or believes that it has cached some nearby. For attributing beliefs and desires does no more, initially, than indicate that the behavior, so far as we are able to determine, is purposeful, and we are surrounded by examples of apparently purposeful behavior. All that is necessary is that the behavior have a particular pattern. But whether or not it has this pattern is certainly a matter for inter-subjective verification, however difficult it might be to carry out such verification in individual cases.

At the same time, there is something problematic about intentional explanations. It follows from the fact that the attribution of beliefs and desires allows us to understand why an agent, human or animal, acts as it does only if we assume its rationality, that is, only if we assume that, given its beliefs and its circumstances, it acts in the way most likely to satisfy its desires. But, so one could argue, the assumption of rationality, at least in the case of other animals, is impossible to justify. Confronted with this difficulty, we simply ignore it, 'reading into' their behavior a conception of how a rational *human* would respond to a similar situation. The tendency to model the behavior of other species on our own is unavoidable if we are to explain this behavior in intentional terms. It is also entirely natural. But it is, finally, or so the critics of 'folk psychology' argue, indefensible. Therefore, the attempt to explain the behavior of other species in intentional terms should be given up. It presupposes what is in fact false if not also arrogant, that we can put ourselves in their place.

The crucial premise in this argument, evidently, is that the assumption of rationality, at least in the case of other animals, is impossible to justify. Indeed, I don't think that this assumption can be 'justified', in anything like the usual meaning of the word. But it does not follow that the behavior of animals cannot, justifiably, be explained in intentional terms. For my purposes, it suffices to make three points.

The first point is that this same line of argument is used, increasingly, in connection with human behavior, and has nothing specifically to do with animals. Rationality is held to be race, gender, or culture-relative in the same way and to the same extent that it is held to be species-relative (cf. HEARNE 1986, pp99–100, who makes the parallel explicit).

Now what could this claim of relativity amount to? Only, I should think, that behavior 'irrational' from one point of view is 'rational' from another. But there is no way in which one could establish that irreducibly different conceptions of rationality were at stake. For it would always be possible, instead, to claim that heretofore unnoticed features of the situation were responsible for the seemingly 'irrational' behavior and that once they were factored in the behavior could be seen to be, from the usual point of view, entirely rational.

This is, in fact, where intentional explanations get their bite. We are confronted with some behavior, say the careless way in which the Ik people treat their offspring, and react "this is crazy, why would anyone do such a thing?" Only gradually, as we learn more about the desperate conditions in which the Ik find themselves, do we begin to understand their behavior (cf. KITCHER 1985, pp284–285). Often it is very difficult to obtain this information, and there is, of course, no *a priori* guarantee that it exists. The point is that it is in principle always possible to obtain it, and hence always possible to construe any apparently puzzling behavior as rational. The claim that 'rationality' varies with race, gender, culture, or species is empirically indistinguishable from the claim that we have not understood the situation in which 'others' find themselves in enough detail.[4] *All* of the pieces of the puzzle have to be found and assembled before they fit into a coherent whole.

This point applies *a fortiori* to the understanding of animal behavior. When we come across erratic or otherwise 'irrational' reactions, it is always possible to discover information available to the animal, but not (at least not immediately) to us, which makes these reactions entirely comprehensible. Those who have lived in close proximity to animals come to appreciate the range and variety of this information. It allows them to put themselves in the place of the animals, not insofar as they share their feelings (whatever this might mean), but to the extent that they have available to them the same knowledge of context and circumstance.

As I have indicated, 'imposing one's own conception of rationality' is experimentally equivalent to being insensitive to the factors which inform an animal's response. It is easy to surmise, then, why it is natural to 'impose', the result of which is the 'anecdotal anthropomorphism' already dismissed. Presumably the time spent in carefully observing animals and grounding oneself in the details of their background, would not usually be repaid, particularly since, white, male, or western, we are a dominant group. The same would not be true of Native American hunters.

It needs to be added, finally, that coming to understand the reasons for which an animal acted, and thus to construe its behavior as rational, is not necessarily to understand these reasons as 'good'. Whether we think the reasons good or bad is very possibly relative to our own background. But this is irrelevant to the fact that, good or bad, these same reasons allow us to make sense of the behavior in question.

The second point to be made in connection with the assumption of rationality is that rationality, like causality (and eventually optimality), functions as a kind of methodological imperative, enjoining us to keep looking for motives, however opaque the behavior might seem initially. It is for this reason primarily that the 'assumption' cannot be justified, if by 'justify' we mean support with evidence. It plays a different sort of role, the kind KANT characterized as 'regulative'.

Now this claim might seem to conflict with the fact that we often say of particular actions and even particular individuals that they are not rational. For this fact suggests that we *can* determine, on the basis of the available evidence, whether claims of rationality are true or false, and therefore that the assumption of rationality can, like any other empirical hypothesis, be tested. But the conflict is merely apparent. When we say of actions or individuals that they are 'irrational', what we mean is that they do not take account of the information available to them. Beliefs are not based on the evidence, desires are incapable of being realized. But to say this is not to say that, given these beliefs and desires, the actions and individuals are not rational. For 'rationality' in the latter case serves only to indicate that the beliefs and desires fit together in the right sort of way, not that the beliefs and desires are themselves ones that a rational agent should hold. By this same token, any animal whose behavior fits the intentional pattern behaves 'rationally', however 'stupid' or 'intelligent' it might otherwise be.[5]

The third point to be made in connection with the assumption of rationality follows naturally on the first two. It is that this 'assumption' is a conceptual necessity.

A great deal of 18th century moral philosophy turns on the notion of sympathy. Sympathy, in turn, involves 'amiable substitution', unprejudicially putting oneself in the place of another. Our own age has it, rather, that such substitution is never amiable, but always constitutes an usurpation. As we have seen, this sometimes happens, for reasons that have survival value. But it does not always happen, or have to happen. When we come to understand an agent's behavior on discovery of some information formerly available to it but not to us, we have to that extent 'taken its place'. Although we have no criterion of what completely taking its place might mean (and no clue as to whether the possibility is even coherent), surely we can determine something like the extent to which we have, on occasion, done so, even to the point where we can begin to 'read its thoughts'.

At the same time, the notion of 'usurpation' makes sense only if we can characterize the agent's point of view independently of our own. I have argued that this is possible. But it is possible only insofar as we can make sense of its behavior, and this requires assuming from the outset that its behavior is rational. Put the other way around, it would be incoherent to claim that it operated with another concept of rationality than ours, for the only evidence for such a claim would have to be behavioral, and to say that it operated with another concept of rationality would be to admit that we didn't understand the behavior in the first place.[6]

If there were situations in which we might be tempted to say that an animal (or a human for that matter) possessed a special and inaccessible kind of rationality, beyond our comprehension, then I think that we would suspect that the behavior in question was 'instinctual', a hard-wired short-cut to a generally adaptive result (DENNETT 1987, p256), in which case its intentional explanation would be inappropriate.

To the extent indicated, then, anthropomorphism as I have characterized it is possible. When a particular pattern is present in the behavior of an animal it is natural for us to describe and explain it in intentional terms. Although application of these terms is to be justified only on the basis of close observation of the animal and its context, it cannot also require, what is otherwise impossible, that we 'introspect' its mind.

But the possibility of anthropomorphism to this point must be understood cautiously. Nothing I have said so far should be taken to imply that animals really have beliefs and desires, or that other explanations, which do not refer to beliefs and desires, cannot be given of the same behavior, or that these other explanations are not in some sense more 'insightful' or 'fruitful'. All I have claimed is that at least some animal behavior, contextually situated and carefully observed, can plausibly be described and explained in intentional terms. It remains to see whether we *must* do so.

We can contrast three different types of descriptions that might be given of animal and human behavior. They can be called 'behavioral', 'functional', and 'intentional' respectively.

The 'behavioral' description has to do with what is often termed the *form* of the sequence of movements and postures characteristic of the behavior of various animals. In fact, it is difficult to make the notion of a behavioral description precise.[7]

The main problem has to do with what should be taken as the fundamental *unit* of behavior; how should the members in a series of movements and postures be counted? A paradigm description of egg-rolling by Greylag geese was supplied by Lorenz. "When the goose has approached the egg sufficiently, the egg is first touched with the tip of the beak at the point on its surface nearest to the goose…" and so on. (LORENZ 1970, p330)

Three considerations motivate such behavioral descriptions. One is that the isolation of the simplest elements in the sequence of movements and postures allows us to distinguish the specific sensory and motor mechanisms involved, and hence allows us to determine *how* the sequence was brought about. Eventually it is thought possible, and highly desirable, to link these mechanisms in turn with specific structures and processes in the central nervous system.[8]

A second consideration is that behavioral descriptions involve little or no inference, to either the psychological state of the agent or the function of the activity, and in this sense do not involve interpretation or risk error.[9]

This is particularly important when a functional description, in terms of the consequences of the behavior, cannot readily be given, either because the consequences are not obvious or because they are so numerous and complex (see HUNTINGFORD 1984, pp12–13). To provide a functional description in these circumstances would be to speculate.

The word 'functional' here is doubly ambiguous. Any behavioral sequence or routine has a number of consequences, some beneficial, others not. In this very general sense, that it has consequences, all behavior is functional. But in a narrower sense, to say that a behavioral routine is functional is to claim that its origin and maintenance is to be accounted for by the processes of natural selection working with respect to some critical consequence. In this second sense, 'function' has much the same meaning as 'adaptation'. We also need to distinguish in certain cases between the functions behavioral routines now serve and the functions originally responsible for their evolution. In the former case, a routine is said to be adaptive, in the latter case an adaptation. This way of characterizing functions implies, moreover, that they apply to groups or populations. Individuals adapt to their environment by modifying their behavior; we very generally call such adapted behavior 'learning'. But when an evolutionary explanation can be supplied for a behavioral routine common to a group or population, it is a matter of 'instinct'.

Now to provide a functional description of behavior, as I understand it here, is to describe it with respect to some critical consequence, in terms of which its origin can be accounted for by the processes of natural selection. Thus characterized, functional descriptions differ from behavioral descriptions in at least two ways: behavioral descriptions make no reference to consequences and behavioral descriptions do not lend themselves to evolutionary explanations. We can, of course, explain the particular movements and postures in a behavioral sequence in terms of stimuli to which an animal is responding and the underlying physiological mechanisms involved. These are sometimes called proximate causes. But, trivially, it is not until we have a functional description of the behavior that we can begin to talk in terms of selection pressures and adaptation.[10] The behavior of the Greylag goose described by LORENZ and TINBERGEN has a number of different consequences, disturbing the air in the vicinity of its head among them. But it is only when we have characterized the goose's behavior in terms of its critical consequence, the retrieval of an egg that had fallen out of the nest, that we can begin to ask questions concerning its role in differential reproduction and survival. To put it very broadly, we first have to know *what* the animal is doing before we can determine *why* it is doing it.

There is, of course, a great deal of controversy among biologists and others concerning the character and extent of functional descriptions and adaptive explanations of human and animal behavior, some of it fueled by extra-scientific considerations. But I see no reason in principle why responsible functional descriptions and adaptive explanations cannot be given of all types of behavior, in just the same way that they have been given of morphological developments. Whether or not a type of behavior can be given such a description and explanation is an empirical question, to be decided on a case by case basis, although the successes of the theory of natural selection encourage us to keep looking for 'functions' even when they are not readily apparent.

Finally, an 'intentional description' of behavior is in terms of the intention with which it is performed. The paradigm examples, not surprisingly, have to do with human behavior. "They are performing a rain dance" is such a paradigm. That is, they desire rain and they believe that dancing in just this way will at the very least raise the probability of its raining.

It is important to note two things at once about the 'rain dance' description. One is that it does not preclude the possibility of giving a functional description of the same activity.[11] Indeed, in the case of complex human behavior we may describe the activity in terms of its cultural or social function (which in the 'rain dance' case as in many others might have to do with ritual or solidarity), or in terms of its biological function (indicating the ways in which this type of 'bonding' activity provides adaptive benefits to groups who practice it). So far as I can see, giving an intentional description of an activity is always compatible with giving a functional description of it.

The claim of compatibility is reinforced by the following consideration. Intentional explanations, like those in terms of stimuli and physiology, provide us with proximate causes of actions, functional explanations with what are sometimes called 'ultimate' causes. They operate on different levels. Indeed, there is no reason why a functional explanation of having certain types of beliefs and desires could not be given, and intentional explanations put in a larger, biological framework.

The other thing to notice about the 'rain dance' description is that it differs from a functional description in a crucial respect: it is in terms not of the consequences of the behavior (since presumably dancing is not correlated in a statistically significant way with raining), but in terms of the *intended* consequences of the behavior. When we say that a tribe of individuals is engaged in a series of collective bodily movements, nothing is said or implied about what, if anything, is going on 'inside their heads'. But when we say that these individuals are trying to make it rain, it is. For 'trying' is an irreducibly intentional notion.

The point here can be elaborated. Although our evidence for whether or not someone is trying to do anything is in large part behavioral, among other things whether its behavior is both persistent and flexible, 'trying' cannot be *identified* with a behavioral routine. For to say that the dancers are trying is also to say that they are engaging in the routine with a particular intention, that something is going on 'inside their heads'. It is at this point that we must advance beyond the interpretation of a behavioral pattern to the postulation of beliefs and desires as something over and above that pattern. We can determine that the dancers are trying to make it rain;

determining this in no way requires that we can 'see' inside their heads. But if they are trying, then they must also hold certain beliefs and desires.

A related line of argument leads to the same conclusion. If the dancers are trying to make it rain, it can only be with respect to a state of affairs, or goal, not now present. As such, this state of affairs must be represented internally in order to be a factor explanatory of their behavior. That is, in this and other cases like expectation, we must posit the existence of internal representations, although we don't have direct access to them. But beliefs and desires are just such species of internal representations explanatory of behavior.[12]

We can, and in certain respects must, distinguish between intentional and functional descriptions of human behavior. The intentional description is particularly appropriate when the consequences of a behavioral routine have not been, perhaps cannot be, realized. Can the same be said of animal behavior, that is, are there cases in which we have to say that an animal is *trying* to obtain a particular result? If so, then it follows that anthropomorphism as I have characterized it here is inevitable.

Before we answer this question, however, the ways in which intentional and functional descriptions are alike needs to be made clearer.

To begin, functional descriptions are often offered in the same careless way that intentional descriptions are, and for much the same reason: they cut effort short. What might be called 'speculative adaptationism' is as rampant as 'anecdotal anthropomorphism'. It is this sort of adaptationism run wild that characterizes too much work in sociobiology and undermines its credibility.[13]

The difficulty, of course, with both intentional and functional descriptions of behavior is that they are so easy to provide. We can invent functions and intentions (our own or the gods') on little evidence in order to understand and explain virtually everything. What is not so easy is to determine, first, whether the phenomena in question *can* be given a functional or intentional description,[14] that is, whether the criteria on the basis of which they are allowable have been satisfied, and then, second, what the *correct* functional or intentional description is. It is no more and no less difficult to generate and apply the correct intentional descriptions than it is to generate and apply the correct functional descriptions. As we have already seen, the former, in particular, do not require a mysterious ability to 'introspect' the minds of others. In fact, our attempt to generate and apply descriptions in both sorts of case is often frustrated; there are times when we can discover no good grounds for isolating either the intention with which the behavioral routine is performed or the function that it serves.

The difficulty about the plethora of intentional and functional descriptions is connected to the fact that they have to do with a behavioral *pattern*. When we have a purely 'behavioral description' of an activity, we are not thereby given its pattern. That is, two behavioral descriptions describe, within small limits of tolerance, two activities. But very different activities, behaviorally described, can instantiate the same intentional or functional pattern. Indeed, rain-dancing and drought-dancing can both serve the same social function, and presumably there is no limit to the ways in which we can intend to do something. Moreover, the same activity, behaviorally described, can instantiate any number of different intentional or functional patterns. There will always be more than enough patterns to discover in a particular behavioral routine. But once again, whether we have the correct pattern, though it involves interpretation,[15] is perfectly objective; there are criteria on the basis of which we can accept some patterns and rule out others. Which is to say that it is a matter of fact whether behavioral routines have certain functions or whether they are performed with particular intentions.

The pattern-like character of intentional and functional descriptions is linked, in turn, to the fact that neither is, as it stands, predictive. Of course, we can predict that a person with given desires and beliefs will, if he is rational, act in a certain way, and that, if a behavioral routine is functional, then, on certain optimality assumptions, an animal will engage in it. But in neither case are there (non-trivial, empirical) *laws* linking beliefs and desires to actions or functions to behavior,[16] although there may be laws linking each as described in neurophysiological or molecular terms to actions and behavior.

There is, I think, a dilemma at this point for anyone who wants to insist on lawful predictability as a hallmark of science, and *a fortiori* as a hallmark of a science of human or animal behavior. Presumably there are descriptions of behavior which are predictive in the sense that they can be used, together with laws, to entail descriptions of further behavior. Such descriptions are not functional in character. But as I remarked earlier (without much argument), only functional descriptions allow us, at least directly, to enlist selection pressures and adaptations as explanatory of the behavior. To give them up would be to give up a very great deal indeed. The plausible alternative, in the case of both intentional and functional accounts, is to give up the insistence on lawful predictability.

We can, at last, return to the central question: are there cases in which we not only *can*, but also *must*, describe animal behavior in intentional terms, cases in which a degree of anthropomorphism is inevitable?

One sort of case has to do with unrealized consequences.[17] If a goose engages in all of the requisite 'egg-retrieval' behavior, but doesn't in fact manage to retrieve the egg (perhaps because we're conducting an experiment and have rigged things so that she will fail), it is natural to say that she is *trying* to retrieve the egg.

But although it is natural to say so, it is not quite correct. Persistent and perhaps also suitably flexible behavior in the fact of unrealized consequences is a necessary but not also sufficient condition of genuine trying.

The first thing to notice about the idea of trying is that it applies to individuals. The functional concepts typical of biology apply to behavioral *routines*. A routine, as such, is functional, it is a group, or phylogenetic, trait. But intentional concepts are ontogenetic in character, they apply to, in fact serve to individuate, particular agents. Put still another way, 'the goose' whose behavior LORENZ and TINBERGEN describe in functional terms is stereotypical, representative of all (Greylag) geese. An intentional vocabulary is necessary where the behavior is no longer stereotypical, where there are otherwise inexplicable breaks in or departures from the routine. As such, the goose's frustrated egg-retrieving cannot, for all her effort, be correctly described as trying.

The second thing to notice about intentional descriptions generally is that they are particularly appropriate when the behavior in question has a 'cultural' character. This sort of character is obviously at stake in our 'rain dance' case. We both describe what the dancers are doing and explain why they are doing it in the same breath and in the same cultural terms.

'Culture' carries a carload of connotations, and its application to groups of animals is controversial. We might talk instead of 'local traditions', as when the Japanese macaques at Koshima wash sweet potatoes. Cases of local animal traditions are difficult to document,[18] particularly when it is alleged that animals transmit certain traditions from one generation to the next by imitation. Still it is clear enough that animals can learn, and if so they can (sometimes with our help) create 'local traditions' of their own, certain idiosyncratic ways of behaving. However 'cultural' it might not be, a man with his dogs can create a 'local tradition'; the dogs learn that jumping in the back of the pick-up truck pleases the man, and as a result try to do it as often as circumstances permit and are appropriate. At a minimum, then, we require of any animal that can be said to be trying that its behavior can be described and explained in terms of a local tradition, where this may mean no more than that the behavioral routine has been learned. But our goose did not learn (how) to retrieve eggs.

If the behavior of an animal is frustrated, individual, and 'cultural', in the senses indicated, as well as persistent and flexible, then it is correct to say that it involves trying, in which case it is irreducibly intentional.

These points can be put in perspective by considering the likely reply of the critics of anthropomorphism.[19] It is that devices like homing rockets can with equal justification be said to 'try' (to hit their plane or missile targets), yet these devices lack beliefs and desires. However much we colloquially slip into talking about them *as if* they had intentions, it is always possible, in fact much more appropriate, to explain their behavior in terms of the working of their mechanical parts.

Whether or not *any* mechanical device *could* have intentions, homing rockets and their ilk do not. By the same token, they cannot be said to 'try' to hit anything.

For one thing, there is no need to posit 'internal representations' in the homing rocket, not in its heat sensors, not in its processing and guidance systems. The behavior of the homing rocket is simply determined by the behavior of its heat source target. When the heat source target changes flight path, the homing rocket, processing this information, re-directs its flight path. No 'internal representation' (no "map by means of which it steers") is necessary, or for that matter possible, conclusions forced on us when we consider what the homing rocket's behavior might be if all heat sources were eliminated.

For another thing, the homing rocket is not an 'individual' in the same sense that animals that truly try are, nor are they part of any local tradition. It is difficult to make this point without begging important questions. But there is a clear difference between the way in which at least some animals have acquired certain of their intentions, however functional they might be, as a result of learning, and the way in which the homing rocket has acquired 'its' intentions, as the result of being designed and programmed by engineers to perform as they intended.[20] An animal's intentions are its own, the homing rocket's are not. But it is only if the intentions are the individual agent's that we can describe its behavior more than metaphorically using an intentional vocabulary. We typically (i.e., where we cannot find a physical malfunction) invoke beliefs and desires, once again, where there are breaks in or departures from the behavioral *routines* of humans and animals. If there are breaks in or departures from the 'behavioral routine' of a homing rocket, the explanation is always in terms of poor design or mechanical failure.

If we have reason to believe that an individual animal tries to please its owner or expects food at its cache site, then, since 'trying' and 'expecting' are irreducibly intentional, we must also describe and explain its behavior in an anthropomorphic vocabulary. But to argue for this is no more than to rationalize at great length a belief already evident to those who hang around with animals, that they are very much like us.[21]

Author's address

Gordon G. Brittan, Jr., Department of History and Philosophy, Montana State University, Bozeman, Montana 59717.
Email: uhigb@montana.edu

Notes

1 ELLIS (1985) is typical. We are first warned, p29, that "There are three kinds of 'understanding' of animal behavior which are not very helpful or not experimentally verifiable. In practice, they often lead to fallacies which are downright wrong. The first is anthropomorphism: 'understanding' in terms of human emotions and subjectivity". Then the warning is forgotten, as, for example, on p. 88: "Thus angry or amorous octopuses flash red and pale colors and dark stripes in various patterns".

2 From PLINY's Natural History, quoted by DRICKAMER/VESSEY (1982), p11. The first example is Clever Hans, the famous counting horse, the second is taken from DARWIN (1901), p130.

3 SOBER (1993) argues that the apparently verbal dispute between sociobiologists and their critics about the use of certain words masks substantive questions concerning which sorts of similarities between species are to be taken as critical.

4 Of course, considerations of race, culture, gender, or species may make us blind to certain sorts of information, and not others. There is no doubt, for example, that humans, with their late-developing emphasis on sight and touch, tend to play down the importance of the olfactory information available to most mammals. But perceptual biases of this sort have little to do with rationality, and they can be overcome.

5 The last three paragraphs owe a great deal to DAVIDSON (1980), pp268–75, although he would undoubtedly disavow them.

6 The point is indebted to DAVIDSON (1984), essay 13.

7 Some might say impossible. For a detailed, and optimistic, account of how ethologists proceed, see HUNTINGFORD (1984), chapter 2.

8 LORENZ and TINBERGEN were able to demonstrate, for example, that different stimuli were responsible for different movements of the goose's head in rolling an egg back to the nest, some of them instinctive, others under the control of visual feedback. There is lurking here nonetheless an important problem of *scale*. Behavior of even the simplest kinds, e.g., the different movements of the goose's head, are comparatively gross, and clearly perceptible, whereas the events in the central nervous system with which we would like to correlate them are molecular (at best) in character and require instrumentation to detect.

9 Although emphasis on the *visual* form of behavior already betrays a human perceptual bias; as indicted earlier, for most mammals, the *olfactory* 'form' of behavior seems to be paramount.

10 It is not that the stickleback fish (much studied by TINBERGEN and his followers) has a red spot that is adaptive, but that the stickleback has a spot that alerts other males.

11 The word 'activity' is used here in an attempt to skirt difficulties. There are problems with saying 'same action', 'same thing done', 'same event', etc. See DAVIDSON (1980), p4, footnote 2.

12 David ARMSTRONG, following RAMSEY, suggestively characterizes beliefs as (cognitive) maps by means of which we steer (see DRETSKE 1988, p79). In (1988), DRETSKE works out a position somewhat like mine in great and persuasive detail.

13 Just as there are other sources of evolutionary change than adaptation, not all human behavior can be given an intentional explanation.

14 See TINBERGEN (1951), "It is no use denying that there are behavioural elements that may be non-adaptive..." (p153).

15 The same kind of 'interpretation' involved everywhere in science when we see a pattern in the data.

16 For an argument against the possibility of laws in intentional explanations, see DAVIDSON (1980), pp207–225. For an argument against the possibility of laws in functional explanations, see SMART (1963).

17 A related sort of case has to do with anticipated consequences. Although there are flaws in his analysis, and an ill-advised attempt to integrate it with behaviorism, TOLMAN (1959), was apparently the first to see clearly both that concepts like 'expectation' were indispensable in the description and explanation of animal behavior and that they were irreducibly intentional.

18 See GALEF (1996) for some healthy (and very knowledgeable) scepticism.

19 I am not going to consider another line of reply, that the 'beliefs' and 'desires' implicit in trying are not *really* beliefs and desires, 'in the full sense of these terms', since beliefs and desires are propositional attitudes, i.e., require that she who has them must be competently linguistically, and many animals who might otherwise be said to 'try' are not competent linguistically. Beliefs and desires, even in the 'more limited' sense at stake here, have characteristic intentionality (i.e., 'aboutness') and are explanatory of behavior, the point on which McFARLAND and BOSSER's dismissal of 'anthropomorphism' turns. In earlier papers (BRITTAN 1999a,1999b), I deal with problems concerning propositional attitudes.

20 So long as we are careful, it makes some sense to talk about *nature* 'designing' organisms with respect to their structures and behavior, but it shouldn't be taken to follow that nature therefore has certain intentions (the linchpin of the design argument for the existence of God) or that the organisms so 'designed' do not (the linchpin of contemporary arguments for the non-existence of beliefs and desires).

21 An earlier version of this paper was read at Stanford University. Conversations with James ALLARD, Prasanta Bandyopadhyay, James BOWER, Daniel DENNETT, Fred DRETSKE, Sanford LEVY, and Patrick SUPPES have deepened my understanding of the issues involved.

References

Brittan, G. (1999a) The secrets of antelope. Erkenntnis 51: 59–77.

Brittan, G. (1999b) The mental life of other animals. Evolution and Cognition 5: 105–113.

Coren, S. (1995) The intelligence of dogs. Bantam Books: New York.

Darwin, C. (1901) The descent of man and selection in relation to sex. John Murray:

Davidson, D. (1980) Essays on actions and events. Oxford University Press: Oxford.

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

Dennett, D. (1987) The intentional stance. MIT Press: Cambridge MA.

Dretske, F. (1988) Explaining behavior: Reasons in a world of causes. MIT Press: Cambridge, Mass.

Drickamer, L. C./Vessey, S. H. (1982) Animal behavior: Concepts, processes, and methods. Willard Grant Press: Boston.

Ellis, D. V. (1985) Animal behavior and its applications. Lewis Publishers: Chelsea, MI.

Gombrich, E. H. (1996) The miracle at Chauvet. The New York Review of Books, November 14.

Galef, B. G. (1996) Tradition in animals: Field observations and laboratory analyses. In: Bekoff, M./Jamieson, D. (eds) Readings in animal cognition. MIT Press: Cambridge MA, pp. 91–105.

Hearne, V. (1986) Adam's task. Vintage Books: New York.

Huntingford, F. (1984) The study of animal behavior. Chapman and Hall: London.

Kitcher, P. (1985) Vaulting ambition. MIT Press: Cambridge MA.

Lorenz, K. (1970) Studies in animal behavior. Harvard University Press: Cambridge MA.

McFarland, D./Bosser, T. (1993) Intelligent behavior in animals and robots. MIT Press: Cambridge MA.

Raphael, M. (1945) Prehistoric Cave Paintings. Pantheon: New York.

Smart, J. J. C. (1963) Philosophy and scientific realism. Routledge and Kegan Paul: London.

Sober, E. (1993) Philosophy of biology. Westview: Boulder CO.

Tinbergen, N. (1951) The study of instinct. Oxford University Press: Oxford.

Tolman, E. C. (1959) Principles of purposive behavior. In: Koch, S. (ed), Psychology: A study of a science. Vol 2. General systematic formulations, learning, and special processes. McGraw Hill: New York, pp. 92–157.

Jeanette Emt and Sverre Sjölander

Utopia in Mind

An Inevitable Consequence of Human Cognition?

Daydreaming got us where we are today; early on in our evolution we learned to let our minds wander so well that they started coming back with souvenirs. — Terry PRATCHETT

Abstract

The dream of a better or perfect world, Utopia, seems to be inescapable for humans. We have the cognitive capacity to imagine any conceivable future. This capacity, needed for complex social interaction, hunting, and gathering, has as a byproduct given us the everpresent notion of a better state, internally or externally. With this notion in mind, we seek ways of improving the inner world (e.g., through meditative or ascetic self-control, use of mind-expanding drugs, etc.) and the external world (via science, engineering, artistic practice, political action, etc.). Another byproduct is the multitude of prophets and political leaders with their minds set on realizing their Utopias at any cost. The evolutionary history of our cognition is also mirrored both in our willingness to follow such leaders and in their will to power.

Key words

Evolutionary biology, cognition, cognitive ethology, psychobiology, Utopia, human internal and external improvement, scientific world-view, artistic perfection, political action, leadership.

1. Utopia in mind

Utopianism, in its most common signification, designates the habits of mind of a spiritual, social, or political reformer, given to visionary dreams and schemes of human improvement. The term originates with St. Thomas MORE's dialogue *Utopia* (1516).[1] In this work, MORE presents a social critique of the political and economic conditions of his day. Tyranny and corruption were ubiquitous, the fundamental evil being seen as the misuse of private property with resultant eviction of the poor from the land, unemployment, and waves of crime to which the repressive state retaliated with brutal laws.[2] MORE's remedy is a contrasting vision of an ideal society on an imaginary island, where tyranny and private property have been abrogated and luxury is unknown.[3] However, MORE was by no means the first to entertain the thought of a much better world than the one at hand; nor will he be the last.[4] It seems that Utopia, whatever form it may take, is an inescapable dream in humanity, an essential aspect of the human condition, as it were. Why is this so?

A dog can obviously have some notion of the future. When its master says, 'Let's go for a walk', it will run to the door, wagging its tail in anticipation. A chimpanzee may cry 'Snake', to make everyone else leave the area, in hopes of having a hidden bunch of bananas all for itself. But as far as we know, man is the only animal capable of making up any conceivable future he likes. He may, for instance, imagine an entire future hunt, from the choice of hunting grounds and what routes to take, through the implements needed, the division of labor within the group, the appropriate steps to take and arrangements to make, to the ultimate possible result and the division and distribution of the spoils.[5] If evolution has provided us with this basic capacity, it follows that we can also imagine other worlds and alternative ways of living, different in degree or kind from the favors apportioned by fate. Thus, the notion of a better world—or the capacity to represent a counterworld—free of starvation, disease, harsh weather, shameful indolence, bad sex, and boring dinner parties appears to be an irrevocable consequence of the evolutionary construction of human cognition.

The ability to think of another reality than the one allotted is not what evolution was geared at, but it is a byproduct of the essential capacity to understand the world, to handle reality, to analyze the environment, to predict possible outcomes of our agency, and to plan ahead for the future. It is a consequence of our cognitive capacity to envision probable future states of affairs, which in turn is a prerequisite for hunting, for storing food supplies, for building huts, etc. But this very capacity allows us to think up entirely novel realities and ways of being as well, to envisage a brave new world where all evil has been eliminated and everything is hunky-dory. With our basic disposition, the possibility of imagining a better or even perfect world follows suit, and—most important—the urge to make it come true.

To be sure, the probability of succeeding in that very endeavor must have been almost trifling in a stone-age culture. You could very well indulge in sweet dreams of health, ample food supplies, and satisfying personal relationships, but there was very little you could do about it. The dreams of Utopia were virtually bound to remain individual and impotent; they rarely sparked off any change in society at large. Perhaps a tribe with a charismatic leader could incidentally advocate and carry into effect a total communism, with regular changes of sexual partners and a communal upbringing of all children. However, this would in all likelihood only be a period of a most transitory nature in the history of the tribe. Soon enough things would return to normal, that is, to an organization structured by connubial, kinship, and friendship bonds.[6]

The feasibility of imposing Utopia upon one's fellow human beings radically increased with the onset of agriculture and husbandry in the great flood cultures, i.e., in India, Mesopotamia, China, and Egypt. For the first time in human history, the tribal system was superseded by hierarchic systems, based not on personal bonds but on roles. In such systems, the practical possibilities of making dreams come true multiply. A pharao, like IKHNATON, would no longer be obliged to resort to his own personality in persuading his cohorts to obey him. The system would, for example, enable him to decide that henceforth everybody shall worship only one god instead of a plethora of small gods—or the other way around, for that matter. The upshot of this is that the new, comprehensive hierarchically structured models made it possible for single leaders to impose their own belief-systems and agendas on very large numbers of people at the same time, independently of contingent personality traits or other idiosyncratic factors.[7]

This turnstone in history brought forth countless of prophets with their minds set on founding new sects or, like MOSES, on having already established leaders embrace the idea of a new better world way over yonder. Reading the Bible will yield sufficient proof of this, but evidence of the same development can be found in all cultures that pass on to agriculture and husbandry and, as a consequence of this, develop large organized societies. Inevitably, these societies all seem to set up a caste system with dejected classes (slaves, workers, or what have you) and dominating classes feeding on them (priests, masters, sages, rulers, warriors, etc.). Without fail, a tantalizing dream gradually comes alive in such a caste system, to wit, the notion of an egalitarian society, where all men are born equal and should treat each other as equals. The fact that very early visions of Utopia (for instance, the parable of the Good Samaritan) stress an artificial kinship—the brotherhood of man—is pleasing to sociobiologists, since it supports the biological point that being closely related is the key to helping and supporting other human beings. Undoubtedly, JESUS was not the first reformer to use terms such as 'father', 'son', 'mother', 'brother', and 'sister.' As soon as an organized society with a more or less rigid caste system is in place, the seed of the egalitarian dream has naturally been sown: intimations of a better world where everything is shared, where all men are equals and worthy of equal respect, and where no individual is disproportionately happy in comparison to the rest of us.

The general point here is that the practical consequences of utopian thinking are dependent on the social structure within which the individual lives. In a tribal society, these consequences will be short-term and restricted to a small social sphere; in a large-scale hierarchic society, the effects of utopianism may be revolutionary, long-term, and very often terrifying.

2. Human Cognition and Alternative Worlds

We share almost 99 percent of our genes with our closest relative in the animal kingdom, the chimpanzee. Still, we are obviously distinctive in a marked degree, being able to accumulate knowledge and to create cultures. But what are those salient features, more specifically, which differentiate us from other species, by enabling us to master our environment and to fashion worlds and ways of living that no other mammal or primate could ever dream of?

Different times and traditions have suggested various answers to this question. For instance, many religions attribute a soul to man, some kind of spiritual quality or substance that only members of our species are endowed with. This notion commonly entails that, equipped with this soul, humans also have privileged access to the Tree of Knowledge, or that we are the unique bearers of reason. In some extreme cases, the assumption is not just that we differ from nonhuman animals by having reason and the capacity for symbolic thought, but also that only we have emotions. Oftentimes the practical consequences of this thinking, which places other animals on an equal footing with mere automatons, have been atrocious, in the form of excruciating animal suffering.

Since the Enlightenment era modern science has searched for explanatory alternatives to the soul hypothesis, pointing to man's ability to make tools, his capacity to plan ahead in several steps, general intelligence, and—last but not least—language. However, recent primate research has shown that these supposedly exclusive human traits to some degree also are prevalent in our animal cousins. Yet the fact still remains that man is the only animal that has left nature's unrelenting thralldom behind, by evolving into the only species capable of deciding—in a great measure, at least—how it wishes to conduct its own life (e.g., SJÖLANDER 1997).

Are there biological behavioral traits unique to human beings? Indeed, there are several, but in this context one is of particular interest. Humans display a considerable amount of altruistic behavior. In other social species, interactions that seem altruistic at the first blush can well-nigh always be explained in terms of kinship, or else as a rudimentary form of reciprocal altruism (e.g., TRIVERS 1971). But the human species has added a highly developed form of reciprocal altruism, namely, contract-based mutual cooperation, which proceeds tit for tat. Many small favors call for a big one in return, or vice versa. Metaphorically speaking, each party to the agreement has a bank account in the other party, where deposits and withdrawals can be made. Balancing the checkbook, one may either owe or have justifiable claims to make on favors, help, or compassion to or from the banker. In order to succeed in this enterprise, we must know what the other person knows and does not know, differentiating between our own and his knowledge. This is a prerequisite not only for cooperation but also for fraudulent proceedings, or the ability to cheat. Apart from humans, only the great apes are capable of cheating, though in a rather limited degree (e.g., DE WAAL 1982, 1988). Noteworthy in this context is also the ontogenetic fact that the ability to readily distinguish between what I know about myself and what others know about me does not occur until the age of five or six in the mental development of the human child (e.g., PIAGET 1992).

On the face of it, the most plausible explanation for the evolution of this mental capacity in our species is that it is required by a fully developed social interaction and cooperation in a 'normal' human tribe (probably, around thirty to fifty individuals). For knowing what everybody else knows, more or less, is a prerequisite for efficient and economic linguistic practice. In communicating we do not transmit complete information; rather, we just send the critical bits that fill in the gaps in the recipient's knowledge, or we suggest a rearrangement or reconstruction of information already existent in his mind. To be able to know what other people know one must possess the capacity to imagine another inner world than one's own, alternative ways of thinking, and different ways of understanding or explaining the world. And one must be able to entertain such notions abreast of one's own beliefs, to differentiate between alternative inner worlds, and—whenever this is appropriate—to seize upon them, exchanging an inner world for a new and better one. After all, this bartering and comparison of conceptions is what humans in all cultures spend a very considerable time doing; the importance of gossip can hardly be overestimated.[8]

An interesting consequence of the ability to understand how other people see the world is that it also enables us to empathize, even with strangers, since we can understand that the other individual suffers or is in pain.[9] This capacity for empathy in turn is a necessary condition for a moral obligation: Only if we *can* understand that another person suffers, there exists a moral duty to the effect that we *ought* to try to alleviate the suffering (as well as the prohibition that we must not be the cause of the person's pain).[10] A genuine or pure altruism—one that seeketh not its own—thus emerges, made possible by our capacity to understand that other people may suffer the same pains as we ourselves do, or would be as happy about aid and care as we would be in a similar situation.

So the proposal made here, then, is that our ability to think of other worlds than the one at hand evolved as a necessary means for interacting with other individuals in a complex social setting, i.e., the human tribe. But as so often in evolution, once a trait that has evolved for a specific reason is in place, it may be used for new, unexpected purposes as well. Flight in birds may be used not only to escape enemies, but also to catch prey in midair or on the surface of the ocean, to nest in inaccessible places, to migrate, etc. Although it originally evolved as a means of speeding up the decomposition of prey swallowed whole, snake poison may also be used as a defense against predators. The ability to imagine how a fellow human being actually sees the world may also be used to dream about a nicer world, independently of occurrent stimuli.

3. Improving the Inner World

There are two ways of coming to grips with a toothache: either by controlling your mind through meditative techniques to the effect that you can ignore the toothache, or by having a dentist repair or extract the tooth so that the pain is eliminated. In a similar vein, the dream of a nicer world may either emphasize inner change, or lay stress on manipulation of the external world.

During most of the history of mankind, which spans at least six million years, we have been obliged to resort to the inner world and its improvement. As the opportunities to exert control over the external world were highly limited, we were thrown upon our own resources of mind, so to speak. Instead of experiencing frustration and helplessness, or venting one's anger on innocent people or inanimate objects, one should strive for a better state of affairs by bearding the lion in its den; that is, personal betterment through self-mastery. According to Sir Isaiah BERLIN, 'being one's own master' is the content of the notion of 'positive' freedom (as contrasted with 'negative' freedom, which consists in noninterference from other people):

"Have not men had the experience of liberating themselves from spiritual slavery, or slavery to nature, and do they not in the course of it become aware, on the one hand, of a self which dominates, and, on the other, of something in them which is brought to heel? This dominant self is then variously identified with reason, with my 'higher nature', with the self which calculates and aims at what will satisfy it in the long run, with my 'real', or 'ideal' or 'autonomous' self, or with my self 'at its best'; which is then contrasted with irrational impulse, uncontrolled desires, my 'lower' nature, the pursuit of immediate pleasures, my 'empirical' or 'heteronomous' self, swept by every gust of desire and passion, needing to be rigidly disciplined if it ever is to rise to the full height of its 'real' nature." (BERLIN 1967, p150)

The goal of self-control is to master the inner world of desires, emotions, feelings, impulses, pangs, affects, preoccupation and rumination.[11] This ideal, in different guises, can be found in all major religions and life-philosophies.

Classical Greece virtually abounded with examples of it. Stoicism, for instance, preached the virtue of self-control in the form of Stoic resignation or *apathia*, which encourages one to accept one's situation in the world, and to view this as a reflection of the ultimate reason of things. To live according to reason means to simplify one's life. Simplicity as a route to the austere independence of the will was also cherished by the Cynics. DIOGENES of Sinope is a case in point. According to legend, he lived in a tub at the temple of Cybele. On seeing a slave boy drink from his hands, DIOGENES destroyed the single wooden bowl he owned. ALEXANDER the Great offered to fulfill any desire he had, and DIOGENES requested that ALEXANDER not stand between him and the sun. Being an advocate of virtuous self-control, he held that morality implies a return to natural simplicity. And virtue requires the avoidance of physical pleasure, both pain and hunger being positive aids in the attainment of virtue. To the EPICUREANS, on the other hand, the ideal state of mind and feeling was *ataraxia*, a state of calm, untroubled pleasure, enjoyed in tranquillity and free of mental or physical disturbance. *Ataraxia* must not be mistaken for any lustful kind of pleasure on the sensational model; what it amounts to, essentially, is peace of mind. The EPICUREAN communities provided a peaceful escape from ordinary society and also a substitute for it, employing pastoral techniques in guiding their adherents to a serene and happy life (on these elements of Greek philosophy, see LONG 1974, O'CONNOR 1985, RIST 1969, and SAYRE 1948).

Self-knowledge always lies at the heart of the ideal of self-control or self-restraint; the famous maxim of the Delphic Oracle, 'Know thyself', is by no means an accidental slip of the tongue. Evidently, many modern therapies of the depth-psychological variety focus on self-knowledge via the 'talking cure' toward inner health. In psychoanalytic therapy, for instance, the patient is supposed to utilize the method of free association in order to reach insight about the unconscious cause of his neurosis by connecting it to traumatic childhood experiences, perhaps using the material of his dreams. Once its origin has been revealed in the clear light of consciousness, the neurosis will disappear, and in that way the patient will gain control over his inner world of emotional attachments and reactions (cf. FREUD 1955).

Self-control in the form of self-denial is the dominating theme of asceticism: the body is to be denied, possibly mortified, in order to make possible the purification of the soul in its progress toward a better or perfect state. This point can be exemplified very widely in religion. In Hinduism, the third and fourth stages of life are expected to embody renunciation, separation from family, and a mendicant life as a means to purification. In yoga, which has its roots in Buddhism, the techniques for disciplining the body are often quite rigorous, especially in *Hatha* yoga (ZIMMER 1951). The point of disciplining the body is to make oneself endure such pain and discomfort that ordinary mortals normally cannot bear, e.g., in the kind of trials that Tibetan monks have to stand when they go out in ice-cold weather, wrapped up in a soaking wet sheet to let it dry slowly on the surface of the body. In Christianity too, the examples of asceticism range from monasticism to hermitage. Be it through withdrawal from society or self-torture, "the law of sin that dwells in my members", as ST. PAUL has it, is to be overcome (*Romans* 7.23).[12] The penitential movement of the Flagellants in the Middle Ages was characterized by processions of men scourging themselves with leather thongs and iron whips—a powerful reminder of the central role of suffering in the religion of charity.

One might well ask why pain and suffering have been so much in focus in these movements aimed at self-control. After all, the main concern in our present-day society in the Western world seems to be to maximize happiness and escape boredom. No doubt the ability to increase one's happiness by direct manipulation of the inner world would be extremely attractive to any human being. But earlier in history, it was a far more important concern to control, endure, and/or remove pain and suffering. The reason for this is, of course, close at hand. Until recent times pain and suffering were present more or less on a daily basis, and in large quantities at that. Under such circumstances any method promising to lighten or remove this burden would be of great interest. Undoubtedly, a person's ability to withstand pain, discomfort, hunger, thirst, etc., would then be considered a highly admirable and impressive quality. If you wanted people to regard you as a wise and knowledgeable spiritual leader, what would be more persuasive than displaying your control over the worst and least controllable aspects of your inner world—pain and suffering—through tedious and interminable prayers? What could be more convincing than a preponderance of evidence showing that, after days and nights of insomnia, with your hunger and thirst still unslaked, you were even able to master your insatiable desire for sex?

Did the Flagellants whip themselves in order to practice the control of pain, to show off their ability to control it, or to offer this control as a promise of a better world? Whatever the answer, their existence is just one of many examples of the important role of suffering in the inner world of people during the lion's share of human history when the available technology could do but little to relieve the pain. For want of external technology, we were compelled to utilize any internal techniques for voluntary control over the psychological experience of pain, whether they focused on numbing our reactions or dimming our attention to pain.

The evolutionary function of pain is to prevent permanent damage to the organism, prompting responses that aid recuperation and lead to greater precaution in the future. A direct, voluntary control would be counterproductive, since the ability to shut out the sensation of pain at will would, in the long run, mean a decreased biological fitness. The capacity to forget about the pain in a sprained ankle would be all too tempting to make use of, with the effect that the healing of the ankle was delayed or offset. Evolutionary biology has little difficulty explaining why it would be a drawback to have voluntary control over the sensation of pain.

At the same time, however, we all know that the body has the capacity to block the sensation of pain, if necessary. It is not a law of nature that we must feel pain. Under special circumstances—extreme stress, panic, defending one's offspring, etc.—it is quite possible that even excruciating pain gets blocked. This, of course, also has a survival value, since defending your children despite your own agony will increase your biological fitness. Likewise, it is more adaptive to escape on a sprained ankle, which will be damaged for life, than to perish and lose all chances of reproduction. Accepting pain as inevitable would be far easier if we knew that, on grounds of principle, it cannot be blocked. But we all know that it happens sometimes. Consequently, it should be possible to do the very same thing by voluntary control that a good fright or anger can do. The idea that it has to be practically possible to block pain at will—by way of some mental technique or other—must be as old as mankind.

The same line of argument would also hold true for the sensation of pleasure; it too, after all, is an internal state, triggered by certain external factors. At first sight, it may seem odd that, in order to experience pleasure, we have to make the effort to create an external situation which includes factors that will produce pleasure. For example, you work hard to make the down payment for a yacht, in order to take pleasure in the activity of sailing it. There ought to be some mental switch by which we could make ourselves feel pleasure directly, without having to travel the cumbersome route through the recalcitrant physical world. Many of us would no doubt find it most convenient, and much cheaper, if we could have pleasure simply by chanting a mantra with our eyes closed, in lieu of buying a yacht.

But, as in the case of pain, this would be highly counterproductive in evolutionary terms. If I could have the full pleasure, in my mind, of tasting a juicy mammoth steak, without having to hunt, kill and slaughter the animal, why on earth should I take the pains doing all that? Or if, by sheer imagination, I could make a sour apple taste exquisitely sweet, why not swallow the bitter pill, instead of making the exertion to find a better-tasting apple? If sexual fantasies could give the same pleasure as sexual intercourse, why bother about the potentially embarassing debacle of making contact—in courting, foreplay, and taxing physical exercises that soak the sheets? Why is masturbation just a poor substitute for copulation?

The evolutionary answer is not hard to find. Pleasure is the inner reward system of our body. If we satisfy our bodily needs, doing what is good for us and what increases our biological fitness, our body will reward us by providing pleasure. Obviously, this reward must be withheld until we have done our fair share of the deal. There must be no back door to the candy store where lazybones can sneak in and steel the sweets; they have to go through the motions, using the front door and putting up the dough like the rest of us.

Suppose there was a mutation by which an individual gained control over all the different kinds of pleasurable sensations in the normal human repertoire. This mutation, however desirable for its bearer, would clearly be an evolutionary disadvantage. It is not very likely that a bearer of it would have been competitive and a harbinger of biological fitness in a toilsome stone-age environment, where you had to pay at the price of exposure, hunger, thirst, exhaustion, and boredom for the few evanescent pleasures offered south of the Ice Rim. A more plausible assumption is that this indolent, blissful person would very soon be obliterated from our ancestry. The people who became our ancestors were manifestly those willing to pay the price for survival, rather than the ones who created their own dainty titbits in their heads. But even though this evolutionary consideration is perfectly sound, it seems to have difficulties competing with the grand old idea of a mental technique which allows us to sit down in lotus position and, by an act of will, have great sex, unmitigated fun, and general amusement all day long without moving.

A quite interesting point is that, while it is exceedingly difficult to control, by volition, one's reactions to simple bodily sensations, it is even harder with socially determined responses. For example, we are more hard-pressed to neglect the sorrow and discomfort caused by a fellow human being than the pain of a simple toothache. And just as ANTONY's CLEOPATRA awakens appetites where most she satisfies, the social bed of torment may hold us in equal uncompromising thralldom. Whereas sages in many cultures have practiced strict regimens of bodily control, and surprisingly often have been successful in their endeavor, they all seem to underwrite the shibboleth that the only way to escape the pain that other people may inflict on you is to withdraw from ordinary society entirely. The total solitude of hermits is a phenomenon which is known from Greenland to Tasmania, in all recorded history. And many people in our present-day society every now and then feel that it might not be a bad idea to become a recluse, living on the top of a high pillar in the desert.

Here, as before, the evolutionary background is not hard to make out. More important still than personal pain and suffering is membership in a tribe, and to be accepted by other people. For the major part of the time our species has existed, a solitary human was a dead human. As for the unfeasibility of living a secluded life, one may point to our notorious incapacity to solve our problems by our own accord. It seems that we are wired to deal with our preoccupations in a social context, i.e., by talking to somebody, rather than coming to grips with our difficulties through private ruminations. Unless I discuss my problems with someone, I am not likely to become aware of what they really are—and, hopefully, of their triviality. It is significant that in a society like ours, where it is possible to lead your life as a self-invented figure which, in fact, may be nothing but an empty facade, many people run into trouble they cannot divulge in their immediate social surroundings, because that would mean that they no longer could keep up appearances but would have to unmask. A whole flourishing industry has emerged on this market, where it is current to charge a substantial fee for the service of being an understanding friend or the perfect stranger in the cocktail bar, who listens with a therapeutic mien to our ramblings about our petty misfortunes. In this prosperous cadre of professional helpers, the cure for our sorry condition is never in short supply: ingesting substances that disrupt our nervous functions, new ways of breathing or disturbing the bowel movements, crystals to wear under the waxing moon, or sticking needles into protruding parts of the body.

An expansive path into the inner world that most cultures have explored is the use of hallucinogens, such as, for instance, mescaline and LSD. Significantly, these substances are often labeled mind-expanding drugs, being seen as ways to apprehend the outermost aspects of reality, to acquire knowledge of what is otherwise hidden or unknowable, or to gain a deeper insight into the ultimate reason of things. An interesting phenomenon that clearly has a biological, cognitive foundation is the fact that the person under the influence in general assumes that the hallucination is an experience of something that really exists out there, albeit as an aspect of reality which is normally indiscernible, looming behind an epistemic veil, as it were. Rarely the assumption is that the hallucination is just a state in the drugged mind. However, the belief that what we see is what there is must overall be a vital assumption. As the constructivists emphasize, all perception is in the head, to the effect that the difference between a hallucinating mind and a person in full possession of his senses is not that the former sees a mirage or an illusion, while the latter directly perceives reality as it is; it rather lies in the fact that, in the sober mind, the experience so closely resonates with reality that an uninfluenced person can interact with the external world in a useful way, whereas in the hallucinating mind the construction is in faulty resonance with reality or none at all (for a further discussion and references, see SJÖLANDER 1997). Evolution has favored individuals in whom the resonance is close and reliable. For such individuals it is a most serviceable notion that the things they experience actually exist. Constructing a lion in the mind in close resonance with the appearance of a hungry lion in the external world is very adaptive, as is acting on that construction as if it were real, by dodging the lion. Those of our human predecessors who did not exhibit a sufficiently close and precise resonance between mind and reality simply did not live long enough to become our ancestors.[13]

If the smoke from the herbs tossed on the fire by the Holy Man makes the young apprentice see strange things and hear voices, his first choice will be to gather that he has made a breakthrough to the other side and is being furnished with knowledge about the secret aspects of reality. Far less attractive is the conclusion that it is only his mind running amuck due to poisoning of the brain. To be sure, this inference would fly in the face of the evolutionarily installed resonance between mind and reality, as hallucinogenic substances hardly constituted a selective factor in human evolutionary history, thus making the expectation that we are prepared for their effects unwarranted. Furthermore, by drawing this conclusion the young man would also be prevented from conceiving of himself as a chosen one, an initiate in a clandestine fraternity hoarding ancient wisdom; on the up–and–up, he is being transformed into a holy man who has visions indicating which actions he ought to take or avoid, an awe-inspiring person who is wiser than the rest of us and should be treated to food, shelter, and other mundane necessities of life so that he can devote himself entirely to the mystical rites by the holy well, offering up newborns and a sacrificial lamb or two. So, should Mother Mary come to you at nightfall, speaking words of wisdom, it would no doubt be propitious to assume that this occurrence was for real, thus making you a venerable person who belongs to the select few, cherished and holy. In contrast, by making the inference that your vision is an effect of fatigue, hunger, hysteria, or your condition of being stoned out of your skull, and that it therefore has a reality only in your mind, you will end up with a rather unproductive theory, leading as it does to few attractive repercussions, if any at all. Sometimes, perhaps often, the adoption of a vital lie may turn out to be a successful strategy in improving the inner world of self-perception and self-esteem, even if what it boils down to is self-deception (e.g., GOLEMAN 1998).

The above should not be interpreted so as to say that there is a sharp division to be drawn between a die-hard, matter–of–fact realist, and a superman *manqué*, soaring in mystical dreams. All things considered, it must be a rather advantageous aspect of our cognition to assume that we are smarter, sexier, and more attractive than we actually are, and that our experiences (including the slides from our last vacation) are valuable to other people as well. These exaggerations are apt to increase our self-confidence, which in turn is likely to improve our social standing and ranking in the tribe, on condition that the exaggeration is at least half-way credible. Perhaps we should all lie to ourselves, but to an optimal degree, since lying too much will lead to unacceptable losses in terms of social status and friendship, not to mention financially, when it comes to footing the bill for psychotherapy.

4. Improving the External World

Shifting our attention from the amelioration of the inner world to the external world and its improvement, we find that a precondition for any major progress in the latter enterprise is the availability of successful technologies for exerting control over the various processes and events taking place in our environment. This, in turn, calls for the development of a reliable body of empirical data and testable explanatory hypotheses, based on the general paradigm of natural science. Among all the cultures that have seen the light of day in history, only the modern Western civilization has fully developed what might be termed a scientific world-view, which shuns the disorderly whimsicality of accounts invoking phenomena that occur by a freak of fate, preferring instead simple, systematically naturalistic explanations. Unscientific or religious world-views, consisting of a mishmash of either compatible or incompatible though rarely comprehensively related beliefs, ultimately rest on an epistemic canon or bedrock, a set of fundamental truths or dogmas that are not be disputed under any circumstances. By contrast, natural science may be said to comprise the first general outlook with the radically critical tenet that, given certain conditions, anything can be doubted. Our knowledge of the external world is only provisional and bound to change little by little, as it keeps in step with our observations and experiments. Reality, as construed in those observations and experiments, takes precedence over any favored, inveterate notions we uncritically might embrace about the world, not the other way around (e.g., RIEDL 1987).

Despite the simplicity of the scientific world-view, it yet seems utterly hard for human beings to truly take its inherently critical stance to heart, as well as its naturalizing tendency. Vying for our stakings of faith, and with a head start of several million years, two other human attitudes remain seemingly unfazed: the tendency to uncritically believe in authorities, even if they capriciously set forth a confused, self-contradictory mixture of dogmas; and a disposition for animistic thinking. As for the latter of these cognitive limitations, a lot of people, even in secularized Western societies, still appear to inhabitate some kind of animated universe, where astrology, reincarnation, ghosts, poltergeists, soothsayers, sorcery, and all kinds of antiquated new-age mumbo-jumbo play a significant role, as is evidenced by virtually any double-page spread produced by the tabloid press.

It is not hard to see why an animistic outlook on the world is so close at hand for members of our species. During millions of years, the most important items by far in our environment were other human beings and the animals we hunted, that is, living creatures with affects, emotions, volition, memory, and knowledge of some kind. Against this background, the notion that even plants, in virtue of belonging to the living world, have affective, volitive, and cognitive faculties appears to be a natural and sensible assumption. Why should not a tree have feelings, experience pain when felled, and be prone to retaliate by turning the canoe I carved out from its trunk into a lousy vehicle? On the whole, it seems a good idea to placate the tree before cutting it down, by prayers, sacrificial offerings, or whatever action the wise man of the tribe deems appropriate in this context. And if plants are ascribed a mind or spirit, why not brooks, clouds, and rocks as well? After all, lightning is a very strong indicator of the considerable power possessed by clouds. On the assumption that this tendency toward animistic thinking is a basic aspect of human cognition, as inherited from prehistorical times, it becomes clear why people, even in high tech societies, still respond to their milieu as if the inanimate objects in it were alive and could have a grudge against them or be appeased by prayers to fulfill their desires. Every now and then, the computer will be screamed at, or disciplined with a rapid succession of biffs and blows against the keyboard. And on a frosty morning, the car will be urgently requested to start, and then kicked on the wheels when it refuses.

As for the uncritical belief and trust in the statements issued by authorities, it must have been an exceedingly prudent and advantageous response nearly throughout human history to unquestioningly believe what older, experienced people said. The stone-age world hardly changed at all from generation to generation, and most of the orally transmitted knowledge was reliable. If your grandmother told you that cobras were venomous and mortally dangerous, it paid just to believe her rather than showing healthy curiosity or applying independent thinking in creative explorations of your own. If the elders claimed that wily, large predators lurked in the dark African night, abhorrent creatures that undoubtedly would regard a little boy as a tasty snack, it was a very good idea to take this at face value without any inquisitive experiments as to the truth of the statement. If your mother ruled that meat from carnivorous animals is inedible, she did not have to get herself entangled in wordy explanations referring to trichinosis. *It just is not done* is still a very persuasive reason for adherence to social custom in most parts of the world. And blind obedience, in spite of the fundamentally uncritical mind-set involved, is often an instrumentally rational strategy to adopt.

Thus, one may safely assume that uncritical belief in authorities has a genetically determined foundation in our cognition, going far back in prehistory. This tendency is not only apparent in ordinary everyday life, but has also been pronounced even in the history of science. The maneuver of adjusting one's theories about the world in light of the results of controlled experiments and observations may seem obvious to us now, but it only represents a fairly recent development. Most cultures have never given the idea of it much thought, and our own historical record shows that an allegiance to various authorities by and large was cock of the roost in the development of science until the Enlightenment era. Before that it was a general rule that if authorities like ARISTOTLE stated that dolphins are fish, then so be it—despite overwhelming evidence to the contrary, apparent to any whale hunter. If the Papal States promulgated that the Earth is flat, then it was inappropriate to brood over the awkward fact that ships sailing away gradually disappear below the skyline.

The rise of a full-fledged natural science, governed by the norm that theories be adjusted to the results of observations and experiments, was a revolutionary turn for our way of thinking about the world and our role in it.

The questioning of the ancient wisdom that matter ultimately is reducible to four basic elements, accompanied by piecemeal alterations of the theory and the number of elements until the resonance with reality was so good that experiments produced the predicted results, paved the way for a science of chemistry which has produced substances far more valuable than the alchemist's gold that never materialized, as well as less desirable inventions.

Following systematic observations of animals raised in captivity or the laboratory, the untenability of behavioristic explanations for behavior, solely in terms of reflexes and learned responses, crystallized itself. Having thus ridded ourselves of an image of the brain as an amorphous, unstructured neuronal mush, we were able to understand that the brain instead is a well-adapted, evolutionarily shaped organ, laden with species-specific behavioral programs.

By viewing the human body as a machine to be taken apart, in order to methodically examine its basic construction and various functions in dissections and experiments, with a view to developing effective treatments for the illnesses that human beings may be stricken with, the foundation for the success story of modern medicine had been laid. Its sophisticated cures are drastic improvements when it comes to therapeutic efficacy, in comparison with the medicine man's remedies of tapping large quantities of blood, ingesting long shreds of linen, eating toad turds, or being pulled through a hollow tree trunk. In fact, the medical treatments offered today are so highly developed and subtle that we can look forward to a protracted process of dying of old age, hooked to buzzing machines and flashing appliances, bedridden but in terminal care.

It is outside the scope of present paper to explain, in historical terms, why the Western civilization took this particular course at the outset of the modern age, adopting the method of changing the map when it does not fit the terrain, instead of insisting on the correctness of the map (which is the primary reaction, as service in any army will reveal). However, some general aspects of this question may be of interest here, since they directly involve cognition. We are prone to believe that the modern Western way of thinking is basic and inevitable. But as even modest crosscultural experiences will show, this is by no means the case. For instance, many elements of logical thinking are extremely hard to understand and apply for people in general, and people (even trained statisticians) have a very poor intuitive grasp of statistical principles, something which indicates that these subtleties may have compatibility problems with the stone-age programming.

Self-referential paradoxes only seem paradoxical to people oriented toward the Western way of thinking. In our culture, the liar paradox, in which EPIMENIDES, a Cretan, claims, "All Cretans are liars", typically evokes the following line of thought: "Is he telling the truth or not? One or other must be the case, but if it indeed is the case that all Cretans, including Epimenides, are liars, then both cases lead to contradictions." In many other cultures, however, people have profound difficulties when it comes to dealing with the literal import of universal propositions, maybe because a more particularist kind of thought-pattern has been activated and cultivated in their context (RIEDL 1992). Particularist thinking proceeds in the absence of the notion of a regularity or *ceteris paribus*–law that brings together separate instances under the same umbrella. The following anecdote may be illuminating. Having meticulously trained his pupils to take apart and repair the carburetor of a certain car, the British teacher of a Nigerian mechanics class discovered, to his great astonishment, that the pupils thought that they had learned how to repair the particular carburetor of that particular car. The teacher had taken it for granted that they would universalize their hard-earned knowhow to carburetors in general, or at least to carburetors for the same make of car. But to his pupils that was a new and very surprising strategy, every carburetor being different in some respect from the other; the subsumption of individual carburetors under a general law of mechanics, in virtue of their similar construction, simply was not in the culturally acquired repertoire of these students.[14]

Another interesting cognitive aspect is that we must learn to question obvious clashes in our worldview. It is very easy to see that humans ordinarily are fully prepared to affirm self-contradictory statements and to accept glaring inconsistencies in their general outlook. Most mythologies and sacred texts are full of logical impossibilities, mutually exclusive explanations, and incompatible sayings. However, this does not at all prevent a very large segment of the human population from believing in the truth of the word. Seemingly untroubled, many people simultaneously embrace contradictory theories of the world, e.g., by assenting to the truth of modern astronomy at the same time as believing in astrology. This is perhaps understandable in view of the fact that man's knowledge of the world and its processes did not have to be coherent in prehistorical times. A scattered mosaic of well-founded beliefs was sufficient for responding adaptively in many different situations. There must have been little need to make a unified, systematic picture out of this mosaic, as long as each piece of the tessellation fulfilled its function. A need for a more coherent picture of the world probably did not emerge until agriculture commenced in a larger scale, making much greater demands on longterm planning and organization. Repetitious life in a small tribe does not require a comprehensive, logically sound system of beliefs.

The scientific stance does not only include the basic norm that theories be testable, but also the desiderata that a theory must be internally consistent, that each theory has to be compatible with the other, and ideally, that the same fundamental principles be applicable in every single domain, be it ship-building, the construction of bridges or aircraft, or the technical design of dishwashers. The modern Western culture dreams of a great unifying theory in science, but it is thought-provoking that other cultures and civilizations have shown little of this aspiration, and that humans, even in the Western sphere, likewise seem able to live happily with muddled, logically incoherent, and mutually exclusive ideas about the world they inhabitate.

So far, natural science has yielded singularily efficient methods for improving the external world and making it more fit for human needs and desires. Now we are about to enter a new era. The computerized systems we develop for monitoring and controlling various realms of reality, from manufacture through industrial design to climatic conditions, are getting more and more complex. It would be naive to think that, in the future, we will not build information systems so complex, fast, and comprehensive that no single human brain will be able to monitor or completely understand everything that goes on, or why. It is a crucial question whether we are really prepared to trust such systems, especially information about the consequences of different courses of action, not least political ones. Already there are scary indications that even well documented and scientifically founded predictions about, e.g., the effect of our activities on the atmosphere or maritime ecology are, if not neglected, then at least not at all countered by the necessary political action. A cornerstone of our faith in science and the power of reason has been that, once we have reached a good understanding of some significant part of reality, we will act in a way advantageous for us. Once science has shown, in a convincing way, the risks involved in drug abuse, smoking, drinking alcohol, etc., people will refrain from these actions. Once it is unequivocally clear what greasy cholesterol-distended junk food might do to you and your children, people will turn to healthier food. But this cornerstone of modernity's progressive rationalism is rapidly turning into a colossus with feet of clay.

The assumption that we are rational beings, and that—invoking a somewhat SOCRATIC identification of reason and virtue—once we know what is good for us and our society, we will also be motivated so as to do that, has very little, if any, foundation in history or the development in present-day societies around the world. On the contrary, the examples are legion as to our steady march into sheer folly, as individuals, groups, nations, or supranational federations. By and large, the cherished notion that the foolish deeds by earlier generations were due to the fact that they did not know any better is untenable. In many cases, it can be demonstrated that people actually understood perfectly well where their community or culture was heading, but that they lacked the inclination, the courage, the necessary support from the public opinion, or the political means to do anything about it. The fall of the Roman empire would constitute a fairly convincing example in this regard (GIBBON 1799).

We already have information systems that enable us to make better decisions than the ones we make by our own accord. Doubtlessly, we will see a major development in this sector, systems with a potential of putting us in a position to foresee a great many problems that the future may hold as well as their possible solutions. The question is whether we will be prepared to act on the computations and predictions emanating from these systems, or if we will rashly charge into a vale of tears, if not the valley of death. The notion of a unified view of reality—in tandem with the idea of adjusting the map to the terrain, or harnessing the horses to the carriage (rather than the other way around)—has lead to the greatest revolution in our relation to the external world that humanity has ever known, for better and for worse. On the face of it, there is no limit to this development. We might well ask if it is not the case that we have started a number of self-structuring systems, e. g., in economy, where we have little influence over the upcoming development. Despite the fact that we have constructed and designed these systems, they may develop structures that we have not foreseen and give rise to consequences that we cannot handle. We have let a jinn out of its bottle, and so far we can only guess—at best—what further developments are in store. Human cognition is the same as it was ten thousand years ago, but it has to cope with this new, everchanging world that is so hard to survey. Manipulation of the human genome will perhaps eventually let us change the basic characteristics of our species to order, to the effect that evolution by natural selection could be replaced with evolution by human intervention. But until then, we have to get along with the mental equipment we have inherited.

5. Artistic Perfection

The propensity to delight in artistic creation, or at least in ornament and decoration, is universally observable among human cultures. However variable the admired objects, there is reason to believe that certain basic aesthetic preferences for color, form, pattern, sound, and proportion can be accounted for in evolutionary terms and traced back, for instance, to mating preferences. As for proportion, the cross-cultural male preference for a female waist–to–hip ratio of .70 is a case in point (e.g., SINGH 1993). This phenomenon, or the 'sense of beauty' as DARWIN called it, is ubiquitous in the animal kingdom, where certain kinds of color, body-form, movement, plumage, song, and smell are found attractive. If such aesthetic preferences emerge in a species, they will always play a role in natural selection, thus perpetuating the relevant qualities and behaviors.[15] It is a vexed question whether these preferences in nonhuman animals, too, are accompanied by positive internal states of enjoyment and appreciation. However, at least in mammals it seems a sound assumption to posit some kind of inner experience. Why should our species be so different from closely related ones? Chimpanzees at least seem to enjoy themselves when they get a chance to paint. And if mammals have the capacity for aesthetic pleasure, why not vertebrates in general? There is a striking example in birds. Bower birds decorate their lekking sites by applying blue color (from fruit) to the twigs of the arena with a brush, and by collecting ornamental shells and other conspicuous objects. It makes you wonder if it is really credible that the bird does not take pleasure in this activity, nor appreciates the result thereof. In any case, the burden of proof lies on those who maintain that aesthetic pleasure is an exclusively human affair (e.g., HAECKEL 1899, who was among the first naturalists to argue this point).

Our capacity for aesthetic appreciation, coupled with our ability to imagine yet better aesthetic stimuli than those readily found in nature, paves the way for efforts at artistic perfection. As is well evidenced in ethological research, animals may react more strongly to supernormal stimuli than to normal ones (e.g., TINBERGEN 1951), but man is uniquely capable of creating such stimuli by himself and to his own taste. For example, using a formula reminiscent of DISNEY cartoons, we can create an image which artifactually epitomizes baby cuteness: snub nose, a pair of huge, lustrous eyes, high forehead, small chin, chubby cheeks, a pouting expression, etc. (e.g., LORENZ 1943). In a dialogue on truth and probability in art, GOETHE poses the question of why a perfect work of art also appears like a work of nature. The reply is, 'Because it harmonizes with your better nature. Because it is above natural, yet not unnatural' (GOETHE 1921, p57). The movement toward aesthetic idealization of nature, broadly speaking, has indeed been very powerful in the history of the arts.

Consider, for instance, the urge to represent the perfect human body. According to details of measurement first specified by Jean COUSIN in *l'Art de* *dessiner* (1685), the ideal proportions of the human body are set by the basic unit of a head's length, that is, the distance from the crown to the point of the chin. Based on a golden section or two and the finest works of sculpture from classical antiquity, such as POLYCLITUS' *Spear Bearer*, it is determined that this distance should equal one eighth of the entire length of the human body, though Nature is more prone to display a ratio of 1:7.5. Then the rest of the body is divided into seven more parts, each equalling the head in length: from the point of the chin to the nipples, from the nipples to the navel, from the navel to the lower end of the pubic bone, and so on and so forth, till you have *Aphrodite of the Cnidians* before you. There are ideal measures for each and every part of the body as well: the head, for instance, is divided into four equal perpendicular parts, each of which having the same length as the nose.

In real life, one might very well encounter a hero with an unimpressive physique, and a noble-minded woman may look rather emaciated and have a lop-sided face. But in classical sculpture there is no inner perfection without corresponding external forms; no hero with weak muscles, nor feminine graciousness in want of symmetric voluptuousness. Through direct idealization, the represented figure in a sculpture often becomes a symbol of its own kind; for instance, a hero will be the universal hero, a man will signify manhood in general, or a winner in the Olympic Games will be transformed into The Winner. Even in portraits of real people it is not unusual with some degree of artistic perfection or idealization of the real looks of the historic characters—to render them the way they *ought to* have looked had their appearance completely been in character with their inner selves. This can be seen in portraits ranging from the classical renderings of SOPHOCLES and DEMOSTHENES to THORVALDSEN's Schiller and SERGEL's Gustavus III; these figures all had to be represented as monumental in order to convey their perceived personal qualities. And there are also examples of historic portraits where the idealization primarily functions in a compensatory, indirect manner—e.g., in DA VOLTERRA's bronzes of Michelangelo, whose ugliness, aggravated by his broken nose, had to be set aside so as to enhance the expressive qualities of sensitivity and drama necessitated by his artistic genius. In the end, the idealized portrait is likely to survive, as the 'true' version of its subject. Speaking of MICHELANGELO, it was pointed out that his statue of Giuliano de Medici in the Medici Chapel bore no resemblance at all to its subject, being an idealized Prince in military garb, young, muscular, and pensive. Allegedly, MICHELANGELO's only comment was, 'A thousand years from now, nobody will know what he actually looked like' (JANSON 1986, p455).

Nature may be artistically perfected in so many different ways. The aesthetic effect of a scenic view in a Flemish landscape, for instance, may be enhanced by adding a high mountain to the picture, though mountains are rare things in those parts of Europe (cf. *The Return of the Hunters* by BREUGHEL THE ELDER). The addition may be justified either in compositional terms, or by reference to a sense of drama. In the first case, the rugged rocks of the mountain may function as a contrast to the soft shrubbery and the billowing fields; in the second case, the role of the mountain would be to set a mood to the painting, e.g., by indicating the grandeur of nature. This simple example introduces two general principles of direct idealization in landscape painting: the plastic and the picturesque style, respectively. In a picturesque landscape, the local physiognomy is brought out very sharply, but permeated by a depth of mood through disharmonic forms and a pallet of unruly, romantic, and emotionally tinged color shades. A plastic landscape, on the other hand, launches a nature that is imprinted by a calm ideality, using pure lines in an interplay between curved and rectilinear motifs, distances softly shaded off, etc. The former style is exemplified in works by Jacob VAN RUISDAEL (e.g., *The Jewish Cemetery*); the latter is found in the stylized, heroic landscapes by Nicolas POUSSIN (e.g., *Landscape with the Burial of Phocion*) and Claude LORRAINE (e.g., *A Pastoral)*.

Symmetry plays a vital role in many forms of artistic perfection. The preference for bodily symmetry is well documented in higher vertebrates, including humans, and is thus in all likelihood innate. The evolutionary explanation for this preference is that symmetry is a most reliable indicator of physical vigor and health as well as an upbringing under good, nourishing conditions (e.g., GANGESTAD/THORNHILL 1993). Our inclination toward symmetrical shapes and patterns in art and craft is, plausibly, an offshoot to the basic genetic preference. Why else bother about depicting human faces in such a way that one side mirrors the other? Or representing horses with symmetric front or hind legs? As *perfect* symmetry is seldom, if ever, found in nature, it is not strange that so much artistic effort during the ages has been directed at creating it—in decorating vases, ornamenting knife handles or sheaths, tiling floors, etc. A proximate explanation for the aesthetic appeal of symmetry might perhaps refer to the pleasant impression a symmetric structure is likely to afford us, namely, that of a uniform solidity conforming to law and the ability to carry its own weight with equipoise. This is particularly evident in architecture, where it is also plain that the question of symmetry is related to matters of great practical importance; a building may collapse, for instance, if the pillars supporting the roof are much thinner on one side than the other. In painting, compositional symmetry is often achieved through visual emphasis and subordination; the compositional pattern may then either be based on the pyramid shape, with a central figure or group flanked by two other ones (e.g., in RAPHAEL's *Madonna di San Sisto*), or it may be a horisontal composition which is gathered around a focal point, as in *The Last Supper* by LEONARDO.

Broadly speaking, all music is idealized sound: a rhythmic, characteristically structured progression of tones, or melody. In a normative sense, a melody is a perfected succession of tones that carries an autonomous, complete meaning in itself, and which, therefore, is comprehensible (and perhaps pleasing as well) even when heard abstracted from its larger context and without accompaniment. In a way, the melody is the fundamental musical form, but as simple or homophonic melodious progression it also constitutes a kind of its own among other more plentiful musical kinds. Historically, the transition to polyphonic forms, fuelled by the passionate joy of combination, was a development toward the ideal of pure musical construction, an alluring musical Utopia that hove in sight to the mind's eye: the grand musical cycle with a closed structure, perfectly uniform yet exceedingly rich, exploiting both thematic unity and structural and formal variety to the maximum in a way that seems perfectly natural and facile. The fugue, with its strictly ordered imitation and contrapuntal plaiting of the different parts, is arguably the perfection of polyphony—or the ideal ensemble. In a fugue, the unity of the basic subjects is joined to a variable manifold and wealth of harmonic combinations, as well as an agility in the different parts as they subdue and relieve one another; and this in a manner, too, which exhausts all the possibilities of polyphonic music. What BACH created with the complex polyphony of *Die Kunst der Fuge* was more or less the perfect jam session—and with a one-man band at that! It is just an accidental circumstance that this great work in full, with its vast cycles, can only be performed live by a whole ensemble of instruments, as it would otherwise require some kind of superinstrument that does not as yet exist.

By that we have brushed against a characteristic that has been cultivated in isolation in Western art music, namely, structure; that is, the construction of the succession of notes, the variations on a theme, the interplay between different parts of the work, the rules of what is permissible and what is not, etc. The perfection of musical structure requires, first, the presence of a notational system, and second, musical professionalism. Structure is, of course, already important in improvisatory folk music as well, but in that context it is for the most part a question of how the soloist varies a more or less fixed theme, in a structured extempore improvisation to a relatively simple, repetitious accompaniment in the background. In this connection, the simple but persistent bass rhythm of the boogie-woogie comes to mind, as well as the medieval basso continuo (or thorough bass) which could be designated by numerals to indicate the proper harmony, pretty much in the same way as that of the chords for guitar accompaniment. However, much more than this cannot be accomplished if you are an amateur playing at occasional barn dances, and in the absence of a proper notational system. No doubt a band of talented musicians would be able to learn a multitude of improvisations by heart, but they would clearly meet with resistance in trying to produce two different parts simultaneously without reducing the one to being an accompaniment to the other. It is so much easier to play in unison, creating the illusion of polyphony by brandishing triplets and quadruplets, etc. And a larger structure can also be introduced, designed as a pattern of rules for repeats, with the effect that a couple of rather simple phrases may be extended into a whole dance.

However, when a society has developed to the point that it parades professional musicians, for instance, owing to the fact that the religious institutions or the courts provide opportunities for people to work with music full time, then the need to transcend extempore improvisation follows suit. At this point, the invention of a manageable notational system will also be close at hand, so that you can write down unusually ingenious improvisations by gifted professionals. In that way, other people can play these outstanding pieces too. Furthermore, you can develop the accompaniment beyond simple bass harmonies, for instance, by constructing separate parts that correspond to the solo part in extending tonally and rhythmically harmonious progressions. And then you can pore over the music sheet, polishing it all until every single note tallies with the rest. The result is on the analogy of a jazzband, where all the players simultaneously are soloists, each performing his part, and where all the parts are in perfect accord and harmony. And not only that, for this is indeed a remarkable band: in the final analysis, the players are one and the same person. The entertainment provided stems from the composer, playing his own minutely constructed piece on the different instruments. In this way, the occasional executors are reduced to craftsmen carrying out somebody else's ingenious plan.

The perfect plan is something that fascinates us in many areas of life, from musical composition through chess and crime to technology and social engineering. The individual who thought of it all, covered his tracks and got away unimpeded, perhaps even with murder, is likely to engage our interest and may even, unless he is morally too reprehensible, be an object of admiration. In narratives, literary or otherwise, the perfection of plan or *plot* has been a major structural concern as long as the practice of storytelling has existed, a period which probably equals the time our species has been around. You want to hold the attention of your listeners while telling a story, and you do not want them to be disappointed when all is told. There must be no anticlimax; you cannot have your audience think that the narrated events all came to nothing or just were too far out. By the same token, ARISTOTLE states in his *Poetics* that the events of a plot "make the strongest impact when they occur unexpectedly and at the same time in consequence of one another", whereas in a bad plot "the episodes do not follow one upon another in accordance with the probability or necessity" (1451b–1452a). By and large, the terms of probability and necessity are set internally by the work itself and/or determined by the institution of fiction and its genres, rather than being imported wholesale from the ordinary world. In fiction, we generally allow for a certain degree of condensation and simplification of plot: for instance, it does not seem incredible that everything that transpires in an episode of *Friends* should happen to a small group of people during a very short period of time, or that a die-hard fighter of Bruce WILLIS's ilk could keep on doing business around the clock without taking a leak.

To be really captivating, the successful plot, which is both coherent and complex, must also keep the audience in suspense, and this requires that the storyteller closely monitors the mechanics of the dramatic structure; that is, the way the story is told, the building and relaxation of tension. When it all comes out nicely we experience heightened control, a sense of closure, or even catharsis. Apart from its obvious connections to a mechanistic view of the universe as a great clockwork, the strong appeal of the perfect plot maybe also is parasitic upon the kind of superstitious beliefs to which even mature human beings are prone to revert when unchecked: our existence makes sense because there is a meaning or purpose behind everything that happens, and we are part of the demiurge's grand plan.

Tapping into this sense of meaningfulness, a propagandistic narrative work may be overwhelming to the recipients, especially if it uses powerful aesthetic means in the construction of plot and drama. Here Leni RIEFENSTAHL's film *Triumph of the Will* comes to mind, a documentary of the 1934 Nuremberg rally of the Nazi Party. Its undisputed success as a piece of propaganda must owe in part to how skillfully it casts the German people in a meaningful story, by weaving the ideas of *ein Volk, ein Reich, ein Führer* into a persuasive plot. The vision it projects is that HITLER is the hero of a grand narrative; both leader and savior, he has come to restore a defeated Germany to its ancient splendor. As for historical truth and consequence, the real event of the 1934 Nazi rally did not just unfold; it was constructed in part to be the subject of RIEFENSTAHL's movie. And part of the future this film promised is now a horrifying past it helped to create.

6. Perfecting the World Politically

Every political ideology or movement has a more or less clearly formulated visionary end-state, a desirable goal which is the ultimate objective of all political activity.[16] The element of ideality in the utopian vision may be construed in different ways: (i) as *possible* to carry out *in practice*, by using a particular strategy which is proposed; (ii) as *possible* to realize *in principle*, though we (presently) lack knowledge of or access to the necessary means for doing that; (iii) as a *regulative ideal* to which we ought to strive as best we can, though it will never be fully realized; or (iv) as an *ideal type* or a theoretical abstraction, encapsulating the valuable properties of social or political systems.[17] There are many interesting philosophical issues concerning the nature and justification of political Utopias (for a critical survey of the contemporary debate, see KYMLICKA 1990). Here, however, we shall focus more on their practical dimensions, and on some of the mental factors involved in the construction, acceptance, and application of political ideologies.

Successful as the scientific world-view may have been in indicating workable strategies for changing man's external world into a human zoo, clear of dirt, hunger, pain, disease, toilsome labor, and dangerous animals, it yet has had little effect on political thinking and action. In the realm of politics, the basic assumption still seems to be that the Word is the beginning; ideas of paradise take precedence over reality. Ideologies are seldom based on scientific analyses of actual social conditions and/or human needs and capacities; the analyses involved are often value-laden and question-begging, in that the outcome is already, by and large, tailored to fit Utopia. Whether they are realistic or not, in terms of the present situation, the economic conditions, man's psychological constraints, or the laws of nature, seems to be an issue of minor importance. The worst enemy of politics is reality, and wishful thinking appears to be the main mechanism in adopting goals and strategies. It seems that people with firm political convictions often adhere to the view that the desire to realize Utopia also will make it come true, one way or the other. Purpose and perseverance are everything; and if the political will is strong enough it shall subjugate reality, as faith can remove mountains. The former Swedish prime minister Olof PALME expressed the gist of this pragmatic voluntarism by saying that politics is will, but Benito MUSSOLINI's motto *Potere è volere* was perhaps more dashing.

However, scientifically inspired political Utopias are by no means unknown. Francis BACON, for instance, projected in *The New Atlantis* (1660) an ideal society in keeping with his scientific principles. In that society, science is regarded as the key to universal happiness and is fostered under state guidance and control. Ostensibly, the MARXIST ideology is based on materialism and science, dubbed a 'scientific socialism' as it was by MARX and ENGELS, in contrast to the 'utopian socialism' of PROUDHON, SAINT SIMON, FOURIER, and others. Unfortunately, it must be judged an abysmal failure in this respect, even though it has provided many valuable insights and analyses. Despite its scientific ambitions, the ideology already at the outset was distorted because the theoreticians chose theories and beliefs that fit their visions of the projected end-state and the path leading up to it. The paradigm of adjusting the theory to the outcome of experiments and observations was belied from the beginning.

A pertinent example is that MARXISM immediately seized upon a behavioristic view of human and animal behavior,[18] since the image of the individual as a *tabula rasa* encompassed a highly attractive notion, necessitated by the general upshot of the strategy toward the desirable state of society: namely, that man can be endlessly indoctrinated and molded to fit Utopia. Due to this commitment, the only biology tolerated by the MARXIST–LENINIST ideologues and party officials in the Soviet Union was science that supported the *tabula rasa* idea, or the theory that the environment fully determines behavior and development. For instance, PAVLOV tried to give this general idea a scientific basis in his theory that reflexes are the sole foundation of behavior; and LYSENKO, denying the existence of genes and heredity, worked on teaching potatoes how to resist cold weather, instead of using selective breeding to improve the strains (cf. JORAVSKY 1970). When modern behavioral biology and genetics gained momentum and steadily developed well-founded theories, thus reducing behaviorism and Soviet LAMARCKIANISM to historical anecdotes, MARXISTS nonetheless insisted on the truth of these doctrines—and still do, to a surprising degree.[19]

Nazism similarly claimed to have scientific principles at its core. A huge amount of resources, not least in academic research, was spent in trying to prove the superiority of the master race and the inferiority of others, and to give the extensive sterilization program a 'scientific' foundation. Here, as in many other cases, the leaders had no difficulty finding an abundance of scholars and researchers who were willing to pursue their scientific work in accordance with the desires and directives of the political establishment. In the case of Nazism, two things became crystal clear: first, the enormous respect that even the highest officials had for natural science and the legitimacy it could provide; and second, the ease with which the scientific practice can be corrupted through insidious political gerrymandering, a fact indicating that science involves a way of thinking which is exceedingly hard to achieve and sustain.

Political ideologies seldom, if ever, prosper in the absence of some kind of leader. Like religions, they seem to need a personification, a human face, to flourish. As for penetrative power, the doctrine in itself never seems enough, and so prophets or executive do-gooders become indispensable. This need for personification may well be a reflection of our biological heritage, since the willingness to follow, even to adore, an imposing leader is evident in all group-living primates. The advantages of letting one experienced and ready-witted individual make quick decisions, whom every member of the tribe unquestioningly follows, are by no means hard to identify. However, it may be a bit harder to explain why we strive for power at all, and why so many of us are more than willing to make enormous sacrifices, personally and otherwise, in order to transcend our competence level.

When asked about his literary motivity, author William STYRON allegedly replied that he wrote books because he wanted to get rich so that he could fuck starlets. Similarly, a common opinion is that politicians are in it for the dough, or for all the pleasures that money, prestige, and power can buy. There are, of course, many examples of this, but at the same time we can point to a whole lot of historical cases of an entirely different kind, especially among the very powerful. Many of them lived an austere life, more or less devoid of such petty pleasures as sex, good food, family, children, and friends. The personal happiness of HITLER, STALIN, DE GAULLE, or NAPOLEON does not seem to be worth a life of enormous personal sacrifices, with far-reaching responsibilities, around–the–clock work, stress, and intrigues.

Consequently, the question arises as to the motive or driving force behind such leaders. What could possibly account for their will to power? If it is not the pleasures that the Oval Office can give, we must look for other reasons that can out-lewinsky any strictly self-centered concerns. It is not to be denied that power as such may provide great satisfaction (and increase the attractiveness in men who otherwise would have a hard time), but it seems that a more probable, and perhaps terrifying, explanation is that these people often believe that they have a mission, and that they take on the work and accept all the stress, not for the rewards, nor for power, but in order to create a better world. Today, it may seem ludicrous to say that a man of HITLER's caliber regarded himself as a well-intentioned savior, who was to restore Germany's rightful greatness. Yet even a slight aquaintance with his writings and speeches will substantiate this claim, as well as give the clear impression that he fervently believed in his skewed ideals.

If each generation produces a number of people who are convinced that they have been appointed to create a better world, by leading people out of or into deserts, and who are confident that all nonbelievers and dissenters have to be eliminated, it is an even stranger phenomenon that, no matter how abstruse or evil their political goal is, there will always be people willing to follow the leaders. A common explanation is that the leader is followed because of the good his followers expect from this:

"So even the immoral leader, insofar as he is a *leader*, advances the good of his followers. HITLER claims to be giving back to Germans what the Versailles peace settlement deprived them of. NERO claims to be giving to Rome the cultural values of Greece. STALIN claims to be carrying forward the workers' interests around the world. The evil man is followed for the good he convinces followers he can lead them to. Nothing could better prove the dialectical nature of leadership—the *structural* importance of the followers' will. They must see their own stake in the goal to which they are being mobilized" (WILLS 1994, p228).

This explanation may, however, overstate the consciously rational element involved, i.e., that of 'seeing one's own stake.' In this connection, blind trust in perceived authorities also seems germane. Evolutionary biology may again give us a hint, it being adaptive to follow an experienced and persuasive leader and trust his advice, even at the cost of complying to demands or judgments that initially may seem repulsive or absurd.[20] After all, very few people who were slow on the uptake got old and experienced in prehistory, and those who perchance did were hardly equipped with eyes burning brightly and a honied or thundering voice, filled with persuasiveness. In modern society, self-confidence is not necessarily based on experiences or actions indicating that the person is highly competent in the relevant domain. In prehistory, however, quite the opposite was in all likelihood the case: a charismatic, silver-haired prophet could reasonably and safely be relied upon (e.g., DE WAAL 1982, 1988).

Charismatic figures, like SIMEON Stylites, ST. FRANCIS, or the BAAL SHEM TOV, more often than not are legendary even to their contemporaries. It is of considerable interest to note that two things are often mentioned about the great charismatic leaders in history: their eyes and their voice. The latter is rather obvious, since charisma can hardly be carried without good rhetorical qualities. The eyes, burning like fire, as attributed to many such leaders, seem a bit harder to explain, unless one submits to the simplistic idea that these people just had a confident, steady stare. Everyday experience tells us that outstaring the enemy is a very efficient way of subduing him, and the skeptic may find evidence of this in any old western movie. HITLER is a case in point when it comes to the importance of charisma, since there are many contemporaries bearing witness to his 'steely gaze', and—perhaps somewhat surprisingly to us latecomers, who have only seen him as a raging and screaming lunatic—to him being a very avuncular, charming, and persuasive person, with a warm and mellow baritone (e.g., KERSHAW 1987). The image of HITLER in today's media may even be perilous, giving us the false impression that future leaders of the same dangerousness would be easily spotted as bellowing maniacs.

Without denying that the actual ideas preached by the leader or the prophet are of fundamental importance, it seems clear that personal persuasiveness is a crucial factor. However, one should not forget that some leaders in history have mainly based their position, not on charisma, intelligence, or profound knowledge, but rather on their high ability as plotters in subtle, complicated schemes. Acting in the background as it were, by letting the flashy, charming, and creative colleagues who aspire for power eliminate one another, the grey eminence just waits until none is left and he stands alone—pathetically devoid of any charismatic qualities or persuasive grand ideas, but resting his power on the ability to mislead everyone into thinking that he or she is in particularly high favor with him, while threatening with terror if that favor is lost. One such historical figure that comes to mind is STALIN, but there are many others, even though this kind of tactic seems to be more efficient in smaller institutions, like companies and universities. The fact that this MACHIVELLIAN intelligence clearly has roots going back to our primate forefathers is evidenced by the complicated workings of chimpanzee politics (BYRNE/WHITEN 1988).

As for the birth of new ideologies and new Utopias, the Western civilization seems to have been a very fertile ground. Whereas religious movements and their leaders dominated in the Middle Ages, one chasing after the other, the political ideologies became the vogue in the Enlightenment era. The growing success of the scientific view of the world, the improving infrastructure of the Western civilization, and the discovery of the rest of our planet gave an impetus to dreams of a better or even perfect world, since in all this we found reason to believe in the feasibility of making profound changes in the old society. Liberalism is a good example, with its vision of a world that not only has been materially improved, but spiritually and intellectually as well, a society where the prejudices responsible for inequality and discrimination have been extirpated. The last fivehundred years have presented a steady oscillation back and forth between egalitarian, liberal dreams of a new society and a more conservative adherence to old customs and traditions, and sometimes to more or less Fascistic ideals. But it is hard to see that the different Utopias have undergone any drastic development during this time, since the main strivings basically remain the same, just as the gullibility of the population when it comes to the promise that this time it will all be different.

It is a sobering and terrifying thought that we may well be doomed, through our biology and our cognition, to repeat the same errors forever and ever, always being prepared to try to force our new ideas on the rest of humanity, or to sheepishly follow an adored leader, regardless of the repulsiveness of his preachings. If hope really springs eternal in the human breast, it is hard to see why the future should be devoid of charismatic leaders, striving for high ideals, and eager acolytes prepared to do their bidding.

7. Concluding Remarks

If the dream of Utopia is an inescapable byproduct of human cognition, we must expect it to accompany humanity in the future as much as in the past. The idea of a better or perfect world, with beautiful things and shiny, happy people, is certainly not objectionable in itself, and there is no denying that the permanent striving to improve human existence in many cases has yielded results so devoid of some of the major discomforts of life that few people in modern Western societies, for instance, would swap places even with a prosperous prince in the Middle Ages.

But the undeniable progress on a broad front has also engendered the misleading belief that man has radically changed as well, that an intrinsic mental transformation of our species has taken place. This notion is not altogether far-fetched if reinterpreted in terms of powerful cultural constraints imposed on us through nurture, since it would be hard to find many people advocating the use of torture, death penalty for petty crimes, public floggings or decapitations, and other similar atrocities as means to correct crime and lawlessness in our society. However, the idea that modern man, literally speaking, is mentally different from earlier human models lacks a biological foundation, tempting and flattering though it may be. Merely fivehundred generations of agriculture, husbandry, and large-scale civilization simply cannot have exercised any profound influence on human cognition and behavioral genetics—and even much less so, when it comes to the last twohundred years of relative freedom in the Western sphere from some of the heinous deeds standardly perpetrated by earlier cultures. As Nazism has taught us, it is still all too easy to cast this veneer of humanity aside, even in old nations priding themselves of the advances made throughout history as regards the respect for human lives and rights. The assumption that we still essentially are the same kind of humans as our ancestors does indeed seem necessitated by reality and history, however disagreeable it may be.

If we still are the same in all essential respects, it may well be that Utopia as an idea will always be with us, but that the inherited restrictions in our cognition make it utterly hard for us to judge which utopian visions are feasible, which ones will lead to bliss or to terror and bloodshed, and which flamboyant leader we should follow. Without disputing the need for visions of societal improvement and their potential usefulness, it is vital to remember that virtually all wars, genocides, ethnical cleansings, witch-burnings, etc., have been caused by people trying to realize some blueprint of Utopia, which soon turns into Dystopia. Speaking for ourselves, we cannot but imagine that this world would be a far better place if people at least occasionally stopped imagining better worlds.

Authors' address

Jeanette Emt, Department of Philosophy, Lund University, S-22222 Lund, Sweden. Email: jeanette.emt@fil.lu.se
Sverre Sjölander, Department of Biology, Linköping University, S-58183 Linköping, Sweden. Email: svesj@ifm.liu.se

Notes

1 Utopia is a fictive island, named after its conqueror and founding father Utopus. As MORE was a gem of the first water when it comes to playing with Greek words, 'Utopia' is a coinage which can mean either 'nowhere' or a 'good place.'

2 In the first book of *Utopia*, the character Raphael Hythloday, a travelled mariner, voices his opinions about Europe's various societal shortcomings in the Renaissance: e.g., the enclosure movement, turning farmland into pastures for sheep, has driven peasants out of their fields and on to the highroads as beggars; thieves are being hanged when society instead, since it forces thievery upon those who cannot otherwise support themselves, should provide work for them to do, etc.

3 Utopia is a republican commonwealth, with a communist economy. The island contains fifty-four cities, of which only one is explicitly described, namely, Amaurote, where the national parliament meets. (For the significance of Utopia's republican institutions, see AMES 1949.) The Utopians lead a life characterized by asceticism, discipline, community, and piety. They work six hours a day and spend the rest of their time cultivating virtue.

4 For a survey of sources and precursors of *Utopia*, especially those from classical antiquity (e.g., PLATO'S *Republic*), see LOGAN (1983). Later utopian visions of ideal societies include, for instance, *La cittá del sole* (1602) by CAMPANELLA, *The New Atlantis* (1660) by BACON, *De la législation ou principes des lois* (1776) by MABLY, *Walden Two* (1948) by SKINNER, and many other works. For an overview of Western utopian thought, see MANUEL/MANUEL (1979).

5 As to the breakthroughs necessary for the evolution of a cognition using detached representations instead of cued ones, not only to resonate with the outer environment but to make up whole chains of possible future events, there is a rich literature in the cognitive sciences, starting with pioneers like VON UEXKÜLL (1985), CRAIK (1943), LORENZ (1941, 1943), CAMPBELL (1974), and others. See, e.g., SJÖLANDER (1993, 1997) for further references.

6 The notion of personal property clearly exists already in chimpanzees, as well as jealousy and preferential treatment of friends and relatives. At least in nonhuman primates, such genetically dispositioned factors obviously make any social organization designed to counteract the behaviors associated with these dispositions utterly hard to sustain; see, e.g., BYRNE/WHITEN (1988). Pending evidence to the contrary, it seems a fair assumption that these traits are also part of the basic human biological constitution.

7 NOZICK (1974) distinguishes between 'imperialist', 'missionary', and 'existential' utopianism. The kind made practicable in a large-scale, hierarchically structured society is imperialist utopianism, which forces a certain social pattern upon all members of a population. Missionary utopianism, by contrast, is aimed at persuading people to voluntarily accept a certain pattern. Finally, in existential utopianism one dreams about a world, where a particular way of life is viable and available to those who desire it.

8 According to DUNBAR (1997), two thirds of human conversation consist of items of gossip, revealing what we think about other people, our feelings towards them, the stories and rumors we have heard about their sexual and friendship alliances, etc.

9 The presence of a 'theory of mind' (e.g., HEYES 1998) is thus a necessary condition for empathy, albeit not a sufficient one, as is evidenced by individuals with sociopathic disorders. Empathy is here taken to involve three main components: (i) the *cognitive* state of understanding what it is like to be in another individual's situation or position; (ii) the *affective* state of feeling some degree of concern for that individual; and (iii) the *conative* state of being motivated to act (e.g., to end the suffering, if the person is in pain) (HARE 1981, Ch. 5). In this sense, empathy is to be distinguished from sympathy. To sympathize is to share another's feelings. This does not require a 'theory of mind.' A dog may react to the signs of distress in a crying baby by displaying signs of the corresponding state, e.g., by whimpering. Empathizing, on the other hand, is to 'comprehend viscerally the [other's] inner state' without being 'vicariously possessed by it' (PIPER 1991, pp735–37).

10 Thus, the dictum 'Ought implies can' holds true in more than one respect; 'ought' implies 'can *do*', but also 'can *understand*.'

11 The notion of an inner world, as used in this paper, encompasses all kinds of 'effects evoked in the nervous system' (VON UEXKÜLL 1985, p223).

12 Protestantism, however, especially Pietism, involves a more 'worldly' asceticism, which does not conclude from the devaluation of the things of this world that the individual should 'surpass morality in monastic asceticism', but instead seek spiritual perfection 'solely through the fulfilment of the obligations imposed upon the individual by his position in the world' (WEBER 1930, p80).

13 G. G. SIMPSON'S (1963) classical commentary is worth quoting: "The monkey that had no realistic perception of the branch he was jumping for was soon a dead monkey—and did not belong to our ancestors" (p128).

14 As told by a teacher of mechanics at the University of Ife to S. SJÖLANDER, in 1965.

15 An aesthetic preference need not be an experience which presents itself 'under an aesthetic concept' (RAILTON 1998, p89, on the 'phenomenological thinness' of aesthetic value). Such prefererences can exist and thus shape behavior before the emergence of distinctively aesthetic concepts, both ontogenetically and phylogenetically.

16 This is even true of libertarianism, though some proponents may deny it. The Utopia of libertarianism is a minimal state, protecting the rights of individuals (e.g., property-rights), and a free market. Within this framework, individuals are free to pursue their own visions of the good life as long as they respect the rights of everyone else. Cf. NOZICK (1974), ch. 10.

17 Examples of (i) are legion; e.g., MARXISM (at least according to its self-image), though POPPER (1966) would rather see it as an instance of (ii), because it presupposes an upheaval of all known social conditions. HABERMAS'S communicative ideal of non-domination exemplifies (iii), whereas (iv) is exemplified by PLATO'S *Republic* as well as MORE'S *Utopia*.

18 That MARX held that all human behavior (eating, sleeping, mating, etc.) is learned, and thus socially determined, is evident from several passages in *The Capital*—see, e.g, MARX (1963), pp66, 77, 91–92, 251. The striking similarities between the MARXIST view of man and the SKINNERIAN behavioristic outlook are made very clear by the general exposition in STEVENSON 1974, Ch. 5 and Ch. 8.

19 Interestingly, some adherents of the so-called 'New Left' have propounded the view that early social learning will irrevocably transform human biology. For instance, MARCUSE argues that socially established behavioral norms become 'organic' in, and 'second nature' to, the socialized individual, thus permeating and changing his or her structure of biological instincts (MARCUSE 1964). See also LEWONTIN/ROSE/KAMIN (1984).

20 People's tendency to conform their judgments to absurd, perceptually unconscionable judgments by other individuals is documented in, e.g., ASCH (1956). The MILGRAM (1974) experiments showed a high rate of compliance to demands to which people even had a strong disposition to rebel but no social support for so doing.

References

Ames, R. (1949) Citizen Thomas More and His Utopia. Princeton University Press: Princeton.

Asch, S. E. (1956) Studies of independence and conformity: A minority of one against a unanimous majority. Psychological Monographs 70, vol. 9.

Berlin, I. (1967) Two concepts of liberty. In: Quinton, A. (ed) Political Philosophy. Oxford University Press: Oxford.

Byrne, R. W./Whiten, A. (eds) (1988) Machiavellian Intelligence: Social Expertise and the Development of Intellect in Monkeys, Apes, and Humans. Clarendon Press: Oxford.

Campbell, D. (1974) Evolutionary Epistemology. (The Library of Living Philosophers, Vol. 14.). Open Court: Lasalle.

Craik, K. (1943) The Nature of Explanation. Cambridge University Press: Cambridge.

Dunbar, R. (1997) Grooming, Gossip, and the Evolution of Language. Faber and Faber: London.

Freud, S. (1955) Studies on Hysteria. Standard edition, 2. Edited by J. Strachey. Hogarth Press: London.

Gangestad, S. W./Thornhill, R. (1993) Sexual selection and physical attractiveness: Implications for mating dynamics. Human Nature 4(3):205–234.

Gibbon, E. (1799) The Decline and Fall of the Roman Empire. Benton: London.

Goethe, J. W. (1921) Goethe's Literary Essays. Edited by J. E. Spingarn. Holt: New York.

Goleman, D. (1998) Vital Lies, Simple Truths: The Psychology of Self-Deception. Bloomsbury: London.

Haeckel, E. (1899) Kunstformen in der Natur. Bibligraphisches Institut: Leipzig.

Hare, R. M. (1981) Moral Thinking: Its Levels, Methods, and Point. Oxford University Press: Oxford.

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

Janson, H. W. (1986) History of Art. 3rd edition. Revised and expanded by F. Janson. Thames and Hudson: London.

Joravsky, D. (1970) The Lysenko Affair. Harvard University Press: Cambridge.

Kershaw, I. (1987) The Hitler Myth. Oxford University Press: Oxford.

Kymlicka, W. (1990) Contemporary Political Philosophy. Oxford University Press: Oxford.

Lewontin, R./Rose, S./Kamin, L.(1984) Not in Our Genes: Biology, Ideology, and Human Nature. Pantheon Books: New York.

Logan, G. M. (1983) The Meaning of More's 'Utopia'. Princeton University Press: Princeton.

Long, A. A. (1974) Hellenistic Philosophy. Routledge and Kegan Paul: London.

Lorenz, K. (1941) Kants Lehre vom Apriorischen im Lichte gegenwärtiger Biologie. Blätter für deutsche Philosophie 15:94–125.

Lorenz, K. (1943) Die angeborene Formen möglicher Erfahrung. Z. Tierpsychol. 5:235–409.

Manuel, F. E./Manuel, F. P. (1979) Utopian Thought in the Western World. Oxford University Press: Oxford.

Marcuse, H. (1964) One Dimensional Man. Routledge and Kegan Paul: London.

Marx, K. (1963) Karl Marx: Selected Writings in Sociology and Social Philosophy. Edited by T. B. Bottomore and M. Rubel. Penguin: London.

Milgram, S. (1974) Obedience to Authority: An Experimental View. Harper & Row: New York.

Nozick, R. (1974) Anarchy, State, and Utopia. Basic Books: New York.

O'Connor, D. J. (ed) (1985) A Critical History of Western Philosophy. Macmillan: Basingstoke.

Piaget, J. (1992) The psychogenesis of knowledge and its epistemological significance. In: Beakley, B./Ludlow, P. (eds) The Philosophy of Mind, pp. 377–84. MIT Press: Cambridge.

Piper, A. (1991) Impartiality, compassion, and modal imagination. Ethics 101:726–757.

Popper, K. R. (1966) The Open Society and Its Enemies. Vol. II. Hegel and Marx. Routledge and Kegan Paul: London.

Railton, P. (1998) Aesthetic value, moral value, and naturalism. In: Levinson, J. (ed) Aesthetics and Ethics: Essays at the Intersection. Cambridge University Press: Cambridge.

Riedl, R. (1987) Begriff und Welt. Biologische Grundlagen des Erkennens und Begreifens. Berlin/Hamburg: Parey.

Riedl, R. (1992) Bewußtsein, systemisch und evolutionär gesehen. In: Guttmann, G./Langer, G. (eds) Bewußtsein: Multidimensionale Entwürfe, pp. 135–61. Springer Verlag: Wien.

Rist, J. M. (1969) Stoic Philosophy. Cambridge University Press: Cambridge.

Sayre, F. (1948) The Greek Cynics. Johns Hopkins University Press: Baltimore.

Simpson, G. G. (1963) This View of Life: The World of an Evolutionist. New York: Harcourt, Brace and World.

Singh, D. (1993) Adaptive significance of waist–to–hip ratio and female attractiveness. Journal of Personality and Social Psychology 65: 293–307.

Sjölander, S. (1993) Some cognitive breakthroughs in the evolution of cognition and consciousness, and their impact on the biology of language. Evolution and Cognition 3:1–10.

Sjölander, S. (1997) On the evolution of reality: Some biological prerequisites and evolutionary stages. Journal of Theoretical Biology 187: 595–600.

Stevenson, L. (1974) Seven Theories of Human Nature. Oxford University Press: Oxford.

Tinbergen, N. (1951) The Study of Instinct. Clarendon Press: Oxford.

Trivers, R. L. (1971) The evolution of recoprocal altruism. Quarterly Review of Biology 46:35–57.

von Uexküll, J. (1985) Environment and inner world of animals. In: Burghardt, G. M. (ed) Foundations of Comparative Ethology. New York: Van Nostrand Reinhold Company. German original: (1937) Umwelt und Innenwelt der Tiere. Zeitschrift für Tierpsychologie 1: 33–34.

de Waal, F. (1982) Chimpanzee Politics: Power and Sex among Apes. Johns Hopkins University Press: Baltimore.

de Waal, F. (1988) Chimpanzee politics. In: Byrne, R. W./ Whiten, A. (eds) Machiavellian Intelligence: Social Expertise and the Development of Intellect in Monkeys, Apes, and Humans. Clarendon Press: Oxford.

Weber, M. (1930) The Protestant Ethic and the Spirit of Capitalism. Allen and Unwin: London.

Wills, G. (1994) Certain Trumpets: The Call of Leaders. Simon & Schuster: New York.

Zimmer, H. (1951) Philosophies of India. Edited by J. Campbell. Bollingen Series XXVI: New York.

C. U. M. Smith

Evolutionary Biology and the "Hard Problem"

Spirit… demands from philosophy not so much knowledge *of what it is, as recovery through its agency of that lost sense of solid and substantial being…*— G. W. F. HEGEL (1977), Preface, § 7

Nothing in biology makes sense except in the light of evolution—Theodosius DOBZHANSKY (1973), pp125–129.

[DARWINISM is like a] universal acid: it eats through just about every traditional concept…— Daniel DENNETT (1995), p63.

Abstract

The term 'hard problem' was introduced by David CHALMERS to describe the crucial hidden issue in neuropsychology: the relation of qualia to the neurophysiology of brains. The paper starts by arguing that the problem is real and inescapable and will not be eliminated by future progress in brain science. After a discussion of the evolutionary origin of 'primary' and 'secondary' consciousness the paper continues by examining what it is about brains which allows this mysterious property to obtrude. This leads to an analysis of complexity, systems properties and, finally, emergence. Two different senses of 'emergence' are defined and it is argued that the second sense, that of a property already present but hidden, is the sense required. The paper ends by considering whether the counter-intuitive phenomena of quantum physics provide an analogy for the essentially panpsychist views which this paper implies. It is argued that the seeming intractibility of the hard problem, like the so-called 'spookiness' of quantum phemomena, is due to our evolutionarily-induced epistemology.

Key words

Evolution, hard problem, complexity, emergence, qualia, consciousness, panpsychism, brains.

1. What Is the 'Hard Problem'?

The terms 'easy' and 'hard' were first introduced by David CHALMERS (1995, 1996) to distinguish between functional properties (nerve impulses, synapses and their organisation etc.) which can (in principle) be completely understood in terms of neuroscience and cognitive science and the 'problem of experience', the 'subjective aspect', the 'something that it is like to be' (NAGEL 1974) which (he says) escape such understanding.

Let me say a little more about this. For the mind/body problem is bedevilled by misunderstandings, by lack of definition, by interlocutors talking right past each other. It is *not* the way the brain calculates, processes information, organises visual input, determines motor output, etc. which is in question: all these (though very far from being understood) are open to scientific investigation and are slowly (and, hopefully, surely) being unravelled.

Quite recently, for instance, *Nature* carried a report which seems to confirm the NMDA theory of Hebb synapses, and hence provides a fundamental explanation of how some forms of associative learning take place. This work involved creating mice which over-express a specific subunit, the NR2B subunit, of the NMDA receptor in their hippocampi. Subsequent behavioural testing showed that they displayed significantly improved associative learning (TANG et al. 1999). The technical sophistication and what might be called 'the citation depth' of these and similar experiments is immense. Yet they do not touch the 'hard problem'. No: it is not problems such as these, complex and difficult as they are, which constitute the 'hard problem' but why *any* of this molecular machinery should be accompanied by subjectivity, by sensa, by feeling, by, in a word, qualia.

2. Eliminativism

To those of us who believe that the presence of mind constitutes the major and crucial problem of modern times it is baffling to find that for a large number of thoughtful and interested people the problem just does not exist. This group of people may, indeed, form a majority of those interested in mind–brain issues. They just do not see that there is an issue to be discussed. True (they say) we may not have a complete solution at hand right now but with the future progress of brain science one will surely come.

Mind, subjectivity, they say will come to be seen as just as misleading a term as the nineteenth century *élan vital* is now seen to have been. The effort to understand the relation of mind to brain is just as futile, just as misconceived, as eighteenth and nineteenth century efforts to discover 'vital forces' in living matter, or 'entelechies' in embryology. The twentieth century has provided perfectly good (though exceedingly complex) accounts in terms of biochemistry and molecular biology. Similarly (argue these eliminativist thinkers) twenty-first century brain science will provide a complete account of subjectivity in terms of a more sophisticated and complex neuroscience. It will add nothing to say that activity in such and such a part of the brain is accompanied by 'raw feels' of a certain type. This Utopian brain science will say it all, without residue.

This type of view is often found amongst scientists. They have been schooled to disdain the philosophical approach, and with good reason. The problems of this world have seldom been solved by arm chair contemplation. The scientist's cry is always for action; for a hypothesis to test in the lab. Furthermore the problem of mind has a history stretching back to at least the pre-SOCRATIC thinkers of classical antiquity. Two thousand and more years of philosophical argument, of writing and meditation, have ended—nowhere. Scientists find it baffling that philosophers of mind cannot forget PLATO and ARISTOTLE, ST. AUGUSTINE and AQUINAS, DESCARTES and HUME, KANT and HEGEL. There must, they say, be something profoundly amiss with an approach which cannot shake off and pass beyond its founders. What physicist nowadays takes DEMOCRITUS or EPICURUS seriously? What biologist spends time reading ARISTOTLE and GALEN?

Furthermore, the problem of mind has from its beginnings been deeply interwoven with theology and the theological notion of 'soul'. This, too is a notion which scientific thought has outgrown. The warfare between science and theology of course still continues: but on a much reduced scale. The major battles were won and lost in the nineteenth century. Only skirmishes remain. The gaps in the scientific enterprise, into which theologians used to squeeze their subject matter, have become ever fewer and less important. Scientists suspect that the concept of mind merely points to another of those gaps. If the future resembles the past, and it is always wise to assume that it will, this gap will eventually close.

And finally: may it not be the case (says the eliminativist) that humans have a powerful psychological defence mechanism against accepting the scientific vision of humanity, of, precisely, themselves. Perhaps the strongest of our inbuilt drives is a drive toward 'meaning'. NIETZSCHE might have called it a 'will–to–meaning'. Yet it is the central implication of Charles DARWIN's work that we can no longer expect to find 'meaning' in the world. The world revealed by science is thus bleak, inhospitable to human dreams. May it not be, therefore, that the eliminativist position is defended against, at a deep level, in the unconscious movements of the psyche, so that it is profoundly difficult to accept? Perhaps, at this deep level, we want to postpone, we want to put off solving the problem. We need to retain an area of mystery, a space for freedom, an element of hope.

3. Contra-Eleminativism

And yet, and yet. Having said all this, having felt the full weight of the eliminativist attack, is this really how it is? Is the problem of mind *really* like the problem of life in the early nineteenth century? Indeed it would make an interesting study to determine exactly what the nineteenth century meant by the term 'life'. Is mind really a problem which can be solved piecemeal, little by little, as the problems of biochemistry and molecular biology were and are? Is mind really a problem which will yield to the scientific techniques of measurement and the descriptive apparatus of mass, length and time? LEIBNIZ has a famous passage in the *Monadology* where he imagines the sensorium scaled up to the size of a Mill so that we may walk therein and yet, he says, we shall never catch sight of a perception, only parts which act one upon the other (LEIBNIZ 1714, §17).

The neurosurgeon hardly expects to come across the vivid sensory images which make up our everyday lives during his operations. A congenitally deaf neuroscientist, sometime in a utopian future, might know all there is to know about the auditory system, yet never have an inkling of what the rest of us have when a tuning fork vibrates at 256Hz. This is the 'hard problem'. Nowadays we would not expect to find the pipes and valves of DESCARTES' imaginary human brain, nor the cogs and wheels of LEIBNIZ' mill, nevertheless we should still only find fluxes of neurochemicals and cascades of charged ions across membranes—and nothing else, nothing resembling a 'raw feel'.

For are not our 'raw feels', our *qualia*, ineliminable? When we visit the dentist for a drilling or an extraction do we not insist on an anaesthetic? And surely we are not merely asking that certain groans, cries, screams are not emitted, so that the calm working of the surgery and the peace of those waiting is undisturbed? Rather, surely, that we should be spared a deeply unpleasant twenty minutes. Surely when we put our hand in the boiling water and scream something more has occurred than when a thermostat set at 100 °C activates a warning buzzer? Similarly with aches and twinges, red after images, tinnitus and a hundred and one other occurrences. We could go on to grief and joy, despair and exultation, the beauty of a sunset over the summer sea, or the grief at losing a beloved face. These surely (to adapt C. S. SHERRINGTON[1]) are the tissues of our lives. The realities within which we all, scientist or humanist, live. Argument over aesthetics, and emotions, may be interminable, *sic et non*, to and fro. At this level we are free to disagree. But surely we cannot dispute the immediacy of the tooth-ache or an interminable tinnitus. This is all we need to counter the eliminativist. Granted the 'raw feel' the rest follows.

So it would seem we could turn the tables on the eliminativist. We could say that far from our unconscious mind-sets biasing us against the eliminativist position it is the vast and ever growing 'episteme' of modern science which biases us against recognising the presence of mind.

Friedrich JACOBI (1811), at beginning of the nineteenth century, surmised that the proliferating organon of scientific thought quite misses the central issue: it diverts attention from this 'overwhelming question' "Our sciences", writes JACOBI, "are games which the human intelligence, to pass away the time, devises for itself. Devising these games, the intelligence only organises its unknowingness without coming a hair's breadth closer to a knowledge of the truth. In a certain sense the mind gets further away from the truth, because when so occupied it no longer feels the pressure of unknowingness… because the game becomes ever more complex, more delightful, huger, more enchanting…".

Indeed we might take a leaf from the work of Martin HEIDEGGER (1962) and say that the all-encompassing thought-world of late-twentieth century science helps 'Dasein' to flee from the 'uncanniness' in which it finds itself. HEIDEGGER says that the 'fleeing' is into the world of the everyday, into gossip and the quotidian, or, to use his own formulation the 'they-self'. At a slightly more sophisticated level the fleeing is equally into the world of science, where everything is seen from the 'outside', the world is a world of 'surfaces', of received knowledge, and we live our lives at second hand.

4. Primary and Secondary Consciousness

Another confusion which often bedevils consciousness studies is the conflation of consciousness with 'self-consciousness'. It is a major part of my thesis that we should keep the two separate.

The recognition that 'consciousness' can be divided into two 'levels' or two 'phases' is not new. We find it already in the fifth century AD when ST. AUGUSTINE writes: "when the mind knows itself… it is itself both knower and known" (*Trinitate 9,4*).

ST. AUGUSTINE's great seventeenth century successor, René DESCARTES, develops a similar account. In *L'Homme* he sets out a theory of conscious automata ('earthen machines') which are in all respects similar to humans except that they entirely lack self-knowledge. Only when the rational soul is introduced into (as he imagined) the pineal gland do these automata know that they perceive or know that they behave. When the rational soul is not present they are as long-distance truck drivers or (possibly) patients with blindsight.

Coming to the nineteenth century we find James FERRIER making the same point[2]. Whereas infrahuman animals show marvellous competencies, building nests and hives, migrating over vast distances, engaging in intricate reproductive rituals, none (says FERRIER) "*know* that they exist, no notion of themselves accompanies their existence and their various changes, neither do they take account to themselves of the reason which is operating within them. It is reserved for man to live this *double* life. To exist and be conscious of existence; to be rational and know that he is so… Man" concludes FERRIER "is an existent who knows that he exists. This is *the* human phenomenon" (FERRIER 1838, pp199–201).

Finally, in the twentieth century, we find one of its most influential philosophers saying much the same. At the beginning of *Sein und Zeit* Martin HEIDEGGER points out that it is only for humans that 'being' is a problem. "Dasein", he writes, "is an entity which does not just occur amongst other entities. Rather it is ontically distinguished by the fact that, in its very Being, Being is an issue for it" (1962, p12).

No other living form is puzzled. In a strong sense, they just are. It has been remarked that when an animal or a six month baby is in pain it is simply 'all pain'; when it is content it is simply 'all contentment'. Following HEIDEGGER many twentieth century thinkers in the existentialist tradition make similar observations. Jean-Paul SARTRE, for instance, having DESCARTES in mind, writes that "the consciousness which says 'I doubt' is precisely not the consciousness which doubts" (1957b, p45). And, in his major work, *Being and Nothingness* he provides many dramatic examples of the division of 'reflective' and 'pre-reflective' consciousness. The voyeur gazing through a keyhole, absorbed in the scene behind the door, hearing footsteps on the landing behind him suddenly sees himself as he presumes the other sees him; or, equally famously, the waiter waiting at table sees himself quite differently than when he is outside the restaurant, in ordinary clothes, in the street, or at home 'en famille'. SARTRE concludes that "I have my foundation outside myself. I am for myself only as I am a pure reference to the other" (1957a, p260).

5. Origin of the Twofold Structure

Whilst these analyses, from ST. AUGUSTINE in the fifth century to Ned BLOCK's P and A consciousness in the late twentieth (BLOCK 1995, 1996), are clearly not all identical, they nevertheless do show a continuing recognition that human consciousness has a two part structure.

There is no doubt that short term memory is of great significance. EDELMAN writes of the 'remembered present' (1989) and from the phenomenological tradition HEIDEGGER's student Eugen FINK contributes a useful metaphor: "Man's mind", he writes, "is like Midas of the legend: everything he touched turned to gold; even food and drink turned into hard metal for him. Likewise, everything we think turns into the hard solid form of 'being', of something that 'is' (quoted DEMSKE 1970, p198). Primary consciousness, on the other hand, is, according to SARTRE, "an impersonal spontaneity always ahead of its self-objectivication".

A future neuroscience will, no doubt, elucidate the physiology of short term memory. I mentioned the NMDA receptor and HEBB synapses in section 1. We shall then (to use MAYR's terminology) have the proximate causation of the two-part structure (MAYR 1982, p67). The mechanisms of human short-term memory are, however, unlikely to differ markedly from those at work in other mammalian brains, or indeed, if we go to the molecular level, of much lowlier organisms, sea slugs and fruitflies. Yet there is, as we have already suggested (and we shall note some evidence below), no reason to believe that consciousness in the majority of infra-human animals has a two-fold structure. It is more in keeping with the ethological evidence to conclude that whereas they may be fully equipped with a SARTREAN 'impersonal spontaneity' or a BLOCKIAN 'phenomenological consciousness' they lack self-awareness, they do not (to revert to FERRIER) know that they know.

To gain a naturalistic insight into why humans (to continue FERRIER's terminology) live a 'double life' we have to examine the second of Ernst MAYR's causations: 'ultimate' causation.

Biologists argue that the 'ultimate causation' of the bipartite structure of the mind is to be found in the evolutionary history of mankind. It originates in the biological niche into which our ancestors were forced at the beginning of the palaeolithic. Lacking the athleticism and weaponry of their competitors on the East African savannahs early hominids were forced into co-operative endeavours. Palaeoanthropologists assure us that human primitives generally live in bands of less than 100 interrelated individuals (DUNBAR 1993). This forced the development of a number of attributes. Prominent among these attributes were language, inter-individual recognition, and self-consciousness.

The origins of language have spawned a vast and (to my mind) inconclusive literature which I have neither the competence or space to review [3]. I would, however, like to devote a few sentences to the other two attributes on my list: inter-individual recognition and self-consciousness. Both, it seems to me, are at the core of the human ethogram.

The importance of inter-individual recognition is clear. The theory of selfish genes (ALEXANDER 1987; DAWKINS 1976; HAMILTON 1964) has shown us that co-operativity can only evolve amongst genetically related organisms. This puts an evolutionary premium on ability to recognise relatives from strangers. The large areas of the human infero-temporal cortex devoted to face recognition attests to this evolutionary imperative, as does the social catastrophe attendant upon the prosopagnosia following from pathology of this part of the brain (see GESCHWIND 1979; SACKS 1986).

Zoologists know of many social species. E. O. WILSON recognises "four pinnacles of sociality: the colonial invertebrates, the social insects, the non-human mammals, and the hominidae" (WILSON 1975, p379). The degree of integration falls away, he remarks, as we "ascend the zoological scale" from the Siphonophora to the Primates. Amongst the insects genetic mechanisms such as haplodiploidy make it an evolutionary stable strategy (ESS) to develop sterile castes and intense sociality (see SMITH/SZATHMARY 1995). With the interesting exception of naked mole rats, such mechanisms have not developed amongst the mammals.

Instead social interaction depends on the brain's ability to recognise other members of the group. In social mammals, therefore, the genetic material comes increasingly to depend on the excellence of the nervous system. If the sensory system cannot recognise other members of the social group its genome, as we noted above, is unlikely to be perpetuated. Genes and brains thus, as BONNER (1980) points out, become increasingly enmeshed in a feed-forward, feed-back, symbiosis.

What has all this to do with the origins of self-awareness and the bipartite structure of the human mind? It is not difficult to see the answer to this question. Nicholas HUMPHREY (1979, 1983) has dubbed the social mammals 'Nature's psychologists'. He argues that in order to forecast the actions of its fellows a social mammal necessarily develops a power of introspection. Observing a conspecific showing a number of behavioural signs it recognises that these are signs it, itself, would make in certain circumstances: it empathises, accordingly, fear or excitement, triumph or disaster and reacts accordingly. Hence, the beginnings of self-awareness.

This 'rational reconstruction' has received support from ethological studies of Primate groups. Experiments, involving the video display of problem situations, have provided evidence that chimpanzees use introspected 'states of mind' to predict behaviour (PREMACK/WOODRUFF 1978).

Other experiments with chimpanzees show that those brought up in their customary social togetherness recognise their own images in a mirror whilst those brought up in isolation do not (GALLUP 1970, 1979). So far only members of the Pongidae, in particular the great apes, have responded in this way (POVINELLI/PREUSS 1995); other primates have not responded to mirror experiments in such a way as to indicate self-awareness.

If mirror experiments have failed to give positive results in non-pongids, numerous other ethological experiments have nevertheless hinted that recognition of self and others may be widespread amongst social primates. Rank order is of great significance in primate troops and it is consequently of considerable importance to be able to recognise individuals higher and lower in the hierarchy[4]. It has been shown, for instance, that in a colony of hamadryas baboons a low ranking individual may take advantage of the presence of a dominant male to threaten a medium ranking individual by positioning himself in such a way that any return threat appears to be directed towards the dominant animal (KUMMER 1968). This, at the very least, shows consciousness (awareness) of position in the hierarchy and (more arguably) consciousness of self and others

Returning to our rational reconstruction, it can be argued that in the early hominids these primate primordia of self-consciousness are strongly reinforced by the practice of attributing praise and blame amongst members of the band.

The practice of holding individuals responsible for their actions seems to be universal in the ethnographic present and is thus, by implication, evolutionarily ancient. Again the centrality of inter-individual recognition is obvious. It could be said that if haplodiploidy forms the root of eusociality amongst the hymenoptera then deontology, the attribution of praise and blame, plays an analogous role amongst the hominids.

This, of course, has the further implication that individuals must feel themselves to be 'free'. It is impossible to work the system if those being praised or blamed do not have the feeling that they were, and are, free to have done and to do otherwise. We can see in this brief analysis that many of the outstanding features of human self-understanding can plausibly be seen as 'DARWINIAN algorithms' developed in response to the challenges of the palaeolithic life-style.

We might, indeed, go further. We might argue that the self-awareness engendered by the palaeohuman econiche led to substance dualism as the palaeolithic theory of mind.

The notion of disembodied 'souls' is found in all primitive cultures (HOCKING 1973). The existence of burial goods dating back to at least 60,000 BC suggests that the peoples of those remote epochs believed that something survived bodily death. Substance-dualism can perhaps be seen as a defence (in the psychoanalytic sense) against the sudden (evolutionarily sudden) recognition of mortality in self and others. This recognition comes, in other words, as an inevitable and unlooked for consequence of the conciousness of self on which the hugely successful hominid life-style is based. In other words, it is a tragic concomitant of one of the most deeply incorporated of DARWINIAN algorithms. It is, to use GOULD's term, a 'spandrel', a tragic spandrel.

6. Primary Consciousness

The previous paragraphs have outlined what seem to be good evolutionary reasons for the development of self-consciousness in *Homo sapiens*. But a 'reflective' or 'secondary' consciousness necessarily pre-supposes an underlying or pre-existing 'primary' or 'pre-reflective' consciousness. This, to repeat, is the ongoing 'spontaneity' of which SARTRE writes. It is the phenomena we live through when red cones are stimulated in the retina and cells are activated in V4, or when a certain group of hair cells are rhythmically distorted in the cochlea and activity occurs in the pathway leading up to and including the auditory cortex. It is our 'state of being' when the skin is indented or, and almost paradigmatically, when the finger or thumb is damaged by an unlooked for impact.

Primary consciousness is surely not restricted to mature human beings. There is, as Charles DARWIN remarks, no saltus between man and the other animals. There is, moreover, no saltus between newborn infant and mature adult. The whole of modern biology—molecular, neurophysiological, behavioural—confirms the profound continuity of the living world. The molecular mechanisms at work in the touch cells of the small nematode *Caenorhabditis elegans* are, for instance, related to the mechanisms present in our own central nervous systems and may, indeed, throw light on human neurodegenerations such as HUNTINGTON's Disease and ALS (DRISCOLL/CHALFIE 1991; HUANG/CHALFIE 1994; etc.). The cellular neurophysiology responsible for, and probably the phenomenology of, binocular rivalry is hardly different between macaques and humans although it is necessary to go back some fifty miillion years to find a common ancestor (LOGOTHETIS/SCHALL 1989; LOGOTHETIS 1998). Pre-reflective consciousness must spread deeply into animal creation. SHERRINGTON has a powerful image. He writes of watching through the lens of a microscope the action of a flea biting: "The act, whether reflex or not, seemed charged with the most violent emotion. Its Lilliput scale aside, the scene compared with the prowling lion in 'Salambo'. It was a glimpse suggesting a vast ocean of 'affect' pervading the insect world" (1947, p22).

And so we return once again to the 'hard problem'. In one of his early notebooks DARWIN writes "how does consciousness commence?" (1980, p24). How does it fit into the evolutionary scenario? How is it related to the brain? And we are forced to go further once an evolutionary schematic is accepted. Why should primary consciousness be a feature only of mammalian brains? Why should it be restricted to vertebrate brains? Why should it not also be an aspect of cephalopod brains, or even those of the large crustacea and annelid worms, not to mention SHERRINGTON's flea[5]. Furthermore, it commences afresh with each new life There has been considerable controversy over the time when it first appears in the human foetus. In a recent review BURGESS/TAWAI (1996) conclude from neuroanatomical and EEG evidence that 'raw feels' begin some thirty weeks after conception.

7. The 'Complexity' of Brains

But what's so special about brains? Like all the other things in the world they are, precisely, things: material entities. Why should this most surprising of all facts, the fact of subjectivity, be associated with, be a property of, just *these* physical entities? If, as we have done, we turn our backs on anthropocentricism, and even mammalocentrism and chordatocentrism, we are forced to consider more general properties. And, if we do this, there seems to be only one characteristic in which brains differ from other material entities: complexity—anatomical and physiological complexity. That brains constitute some of the, if not *the*, most complex entities in the physical universe is proverbial and intuitively compelling.

But what do we mean by complexity? How can it be measured? These are not easy questions (see SMITH 1994; SMITH/SZATHMARY 1995)[6]. Nevertheless we have an intuitive notion of what we mean by 'complex'. The great proteins are, for instance, complex molecules. NASA's space shuttle is a complex piece of engineering. Many computer programs are complex pieces of software. *Paradise Lost* is a complex poem. Complexity and intricacy are, in common usage, close, even synonymous.

Numbers have something to do with it. Most brains have very large numbers of cells and synaptic connections. The human brain is usually taken to have some 10^{11} neurons and 10^{14} synapses. Different morphologies are also significant. Neurons have innumerable different shapes and forms. At the molecular level there is also huge 'complexity'. Neuronal membranes are mosaics of different fluidity, studded with dozens of different receptors, different G-proteins, different effectors. It has been computed that if all the membranes in the human brain were unravelled and spread out they would cover an area upwards of a square mile (250 ha). Complexity stretches down through the orders of magnitude, from the human scale to the nano scale. Anyone who has spent time examining sections of cerebral cortex in the electron microscope will have been impressed by the overwhelming intricacy of grey matter.

This anatomical complexity is fully reflected in physiological complexity. Each neuron has been said to have its own 'personality'. Each has its own personal complement of channel proteins and transmitter substances. Indeed channel proteins are nowadays known to form great families and superfamilies whose members all have subtly different biophysical characteristics (see SMITH 1996 etc.).

The flows of electrotonic current through and around intricate dendritic arborisations, over ever-changing dendritic spines, to activate voltage-sensitive gates located at different places in the soma, but especially at the initial segment, is too complex for present-day techniques to catch.

Voltage-sensitive gates are themselves many and various, each existing in a number of different biophysical states, leading to different levels of resting potential and forms of action potential. Several dozen different neurotransmitters and modulators have been identified, ranging from small inorganics such as NO and CO to large peptides such as substance Y. Although most act directly on subsynaptic membranes via ligand-activated receptors, either in ionotropic or metabotropic mode, some diffuse through the intercellular space to affect comparatively distant cells. The intricacy of grey matter dynamics is only beginning to be guessed.

Yet there is something more. The brain is not only numerically complex and heterogeneous, it is also highly integral. The units are not independent but profoundly interdependent. The notion of a wiring diagram is significant. Many such have, of course, been published. An impressive instance has been provided by FELLEMAN/VAN ESSEN (1991) who show that the primate cortex contains upwards of two dozen intricately interconnected visual areas.

But, more than this, the columnar organisation of the cortex allows the development of what has been termed 'small world dynamics' (WATTS 1998, 1999). The well known iconography of SZENTAGOTHAI emphasises the interconnexity within a column (SZENTAGOTHAI 1978). But in addition to this, emerging from each column are U or recurrent fibres which re-enter the cortex at some more or less distant point. Activity in a column thus influences all the other columns in the cortex. 'Small-world' dynamics ensure that activity in one column quickly affects activity everywhere else in the cortex.

Furthermore, it is not only at the cellular level that there is interactional complexity. Similarly complex wiring diagrams have been published to describe G-protein/second messenger systems in neuronal membranes. Neuronal cytosols are hives of dynamic molecular interactions. The brain is interactionally complex and heterogeneous all the way down (see SMITH 1994, 1996). Unlike other tissues, for instance the liver, every part of the cortex is tied into and influences every other part. The 'wiring' ensures that activity in one part of the system effects spatially distant parts. The whole is bound together into a unitary activity

8. Systems and System Properties

It is a common observation that systems have properties not present in their components. A watch, to use PALEY's model, is a system. Time-keeping is not a property of its individual cogs, springs and escapements. Time-keeping is a system property, it is a property shown when all the parts are assembled in a proper way. A chess-playing computer displays system, or network, properties. The knight's move is not a feature of any of its component transistors, resistors etc.

The brain also has system properties. Catching a cricket ball, ice-dancing, or down-hill skiing are outcomes of which no single neuron knows. They are outcomes of the billions of neurons, the trillions of synapses, acting together, bound together in a system. Or, to take a simpler example, the receptive field of a retinal ganglion cell is the outcome of several thousand ignorant bipolar, horizontal, amacrine, and receptor cells interacting systematically together.

But subjectivity is more than this. Ice-dancing, returning the tennis ball, driving to work, are often almost automatic. Utopian machines should be able to mimic these sophisticated movements. After all, they are already able to challenge the world champion at chess. They are, to use SEARLE's analysis, syntactic, not semantic, machines. Subjectivity is something else again.

9. Emergence and Emergent Properties

'Emergence': a term to conjure with. Can we make sense of the proposition that "consciousness 'emerges' when matter becomes sufficiently complex"? An attractive idea and one which is appealed to again and again. But what does it mean? It is often said that new properties *emerge* when parts are put together to form a system. Mario BUNGE (1979), for instance, defines an emergent property as follows: "Let 'x' be a concrete complex thing and P a property possessed by 'x'. Then P is an emergent or collective property of 'x' iff no component of 'x' possesses P." It is easy to think of examples where this definition applies. Chess playing computers, motor car engines, retinal receptive fields. But, this sense of emergence leads merely to system properties[7]. Let us call this a weak sense of emergence, or E_1. I am arguing that we need something more.

Emergence, in the sense I want to use, is that of a property which is already there, but hidden. This is, indeed, the sense which the *Random House Dictionary* provides: "to come forth into view or notice, as from concealment or obscurity—*a ghost emerging from the grave; a ship emerging from the fog…*" (for further discussion see SMITH 1983). When conditions are correct the formerly hidden property 'emerges' into view. Another example has a contemporary resonance: emergent diseases. The Western World is nowadays concerned that diseases which have been for a long period endemic on a small scale in sub-Saharan Africa may, because of the efficiencies of modern transport, '*emerge*' as full scale pandemics in Europe or North America. Once again the sense is of something already present showing itself when the conditions are correct. For our purposes, let us define this strong sense of emergence, E_2, as follows: "Let 'x' be a concrete complex thing and P a property possessed by 'x'; then P is an *emergent* property of 'x' iff it is present but undetected in some or all the components of 'x'".

Do we have any examples in the physical world? The most frequently discussed examples of E_2 emergence include gravitational force, subnuclear forces and quantum effects.

The gravitational force in the microworld of subnuclear particles is undetectable by contemporary physical methods, it being some 36 orders of magnitude less than the predominating coulombic forces. Yet, in the macroworld which we humans inhabitat it emerges as perhaps the most signficant force of all.

POPPER (1974) points out that short range subnuclear forces are not detectable in the 'free-living' nucleons of intergalactic space. Yet when nucleons are crushed together by huge stellar pressures so that they come to within 10^{-15} m of each other these forces are 'unveiled'. They are responsible for holding together all the more complex atoms of the Universe.

Finally, at temperatures close to 0 °K, quantum effects, normally swamped by statistics, emerge into the human world as superconductivity and superfluidity etc.

All these examples have one thing in common: in appropriate conditions unsuspected properties of the isolated constituents of a complex entity are unveiled, or emerge, as properties of the complex thing itself.

But: note, importantly, these are all *analogies*. I'm not suggesting that qualia could somehow be measured, or physically detected, as G, or the strong nuclear force, could, given some Utopian measuring device, be detected. Qualia cannot be '*observed*', at whatever level.

What is the 'appropriate' condition which allows the E_2 emergence of subjectivity? Is it possible to argue that this appropriate condition is the huge interconnected complexity of brains and, in particular, *grey matters*. This brings us close to LEIBNIZ' monadology and to the process philosophy of WHITEHEAD and HARTSHORNE which DE QUINCY calls pan-experientalism or, to use the older term, panpsychism (see also SPRIGGE 1994; CHALMERS 1996). This is essentially, as the title of DE QUINCY's (1995) paper puts it, 'consciousness all the way down'. It is to say that there is a 'whisper' of consciousness even at the level of atoms. It is to say that only when the monads are aligned in a common hugely complex coherence does this 'strange' property emerge, otherwise it is somehow cancelled out, like quantum effects by the statistics of large numbers.

This is where brains differs so markedly from the complex mechanisms of human technology (see SMITH 1983). The brain shows to an unparalleled degree a property which DENBIGH (1975) defines as 'integrality'. Change in one part leads to alteration in all other parts. Integrality reaches down to the molecular level, and through the molecule to the atom and all the way back up to the square metre sheet of the cortex. In this way the brain is very unlike its silicon-based counterpart.

This huge enbound integrality is unique to brains. Remove a circuit, alter a capacitor, in a silicon based computer and the whole system may break down; it is certainly not the case that small compensatory alterations in the structure of circuits throughout the rest of the system occur. There is, as it is said, no 'graceful degradation'. A computer is engineered from the outside, brains from the inside. The material from which a computer is built is amorphous, it is indifferent to the part it plays in the circuits, it is in no way organised at the molecular level, there is no analogue of the immunological terms 'self' and 'not-self'. I've always liked the quotation from GALEN's *On the Natural Faculties* (1952): "…Praxiteles and Phidias and all the other sculptors used merely to decorate their material on the outside, insofar as they were able to touch it; but its inner parts they left unembellished, unwrought, unaffected by art or forethought, since they were unable to penetrate therein and reach and handle all portions of the material. It is not so, however, with Nature." And, I would add, it is especially not so with brains.

10. Is Consciousness Associated with Large Numbers of Neurons Acting Together?

Is it possible to argue, then, that primary consciousness is a 'brute fact' associated with the integrated activity of large volumes of grey matter? This does not imply that subjectivity has any 'function', any selective advantage, in the evolutionary process. Movement away from harmful and toward survival-enhancing situations could equally well be programmed into silicon or other artefacts without a whisper of concsciousness. Note, also, that the hypothesis does not restrict the coherent activity to cortical grey matter alone: other grey matters, especially central grey matters, may also be involved. Nor, finally, does it imply that what appear to us as simple sensations, the red after-image, the tone of middle C, are necessarily correlated with small areas of cortex, while complicated activities such as ice dancing or riding a bicycle, are correlated with large areas. This may well be a fallacy. Quite possibly it is contrariwise. For my proposal is that consciousness is associated with large volumes of grey matter tied together in coherent activity.

Clearly it is going to be difficult to confirm or disconfirm this hypothesis. As noted in section 7 we do not yet have an adequate measure of the complexity of brains. Nor do we have a good measure of the quantity of grey matter active when primary consciousness occurs. However these measurements are not *in principle* inaccessible. The hypothesis is not *in principle* unfalsifiable. Indeed, with the contemporary explosive development of neuroscience, especially non-invasive imaging, these parameters may well be determined in the forseeable future; at present I can only point to a few suggestive findings in the literature.

The most fashionable evidence at present for the association of consciousness with the integrated activity of large volumes of grey matter is the detection of time-locked electrical and magnetic 40Hz oscillations. These have been shown to occur in wide areas of neocortex, palaeocortex and thalamus in response to sensory stimulation of awake and attending subjects and also in REM sleep (LLINAS/RIBARY 1993; CRICK 1994; DESMEDT/TOMBERG 1994; PARÉ/LLINAS 1995; LLINAS/PARÉ 1996; KOCH 1996; SINGER 1996; etc.). LLINAS conceives of the thalamus as a hub from which looping interconnections to all parts of the cortex engender an 'oneiric' consciousness continuously modified by input from the senses (LLINAS/RIBARY/CONTRERAS 1998).

Turning to that most well known of sensory systems, the primate visual system, there is also much evidence that, although single 'cognitive' cells can be detected (e.g., face-recognition cells, PERRETT et al. 1992; etc.), large areas of parieto-occipital and tempero-occipital cortices are nevertheless involved in visual consciousness (FELLEMAN/VAN ESSEN 1991; ROLAND/GULYAS 1994; MILNER 1995; etc.). The long-continued and incisive experiments of LOGOTHETIS and colleagues on binocular rivalry (perhaps some of the most convincing experiments yet on the correlation of consciousness and brain activity) show consciousness-correlated activity in the inferotemporal cortex and superior temporal sulcus (but not in areas lower in the visual pathway: V1, V2, V4 and V5). The temporal cortex has massive connections to other parts of the brain. (STERNBERG/LOGOTHETIS 1997; LOGOTHETIS 1998).

In addition there is much evidence that deeper and evolutionarily more ancient brain regions are crucially involved. DEVINSKY/MORRELL/VOGT (1995), for instance, review evidence which supports a convincing case for the significance of cingulate gyrus, the amygdala and the periaqueductal grey in 'emotion' and 'emotional responses'.

The various types of non-invasive neuroimagining (EEG, MEG, rCBF, fMRI, PET) as well as the optical techniques of BLASDELL/SALAMA (1986) and T'SO et al. (1990) all indicate that consciousness is associated with activity in macroscopic volumes of cerebral material.

The well known neuro-chronometry of Benjamin LIBET generates the concept of 'neuronal adequacy'. LIBET argues that a subject does not become aware of sensory stimuli until a minimum (and quite large) number of neurons have been recruited (LIBET 1985 etc.). This conclusion has been supported by work reported by Ravi MENON at the millennial meeting of the *Society for Experimental Biology*. Using fMRI techniques he shows that the transition from unconscious to consciousness of patterns of black and white stripes is related to the number of cortical neurons activated.

Note, finally: it is not simply the areas which light up which are associated with consciousness—rather it is the pattern of activity and inactivity across the whole brain which is the physical correlative. LLINAS/PARÉ (1996) emphasise that silent neurons are as important as active neurons in forming the cerebral pattern associated with consciousness. William JAMES, in some ways the presiding genius of neurophilosophy, said much the same: "consciousness 'corresponds' to the entire activity of the brain, whatever that may be at the moment" (JAMES 1890).

All of these investigations, and many others which space does not allow, point to the same conclusion: consciousness is associated with the activity, and probably time-locked activity, of large volumes of grey matter.

Perhaps the best analogy is to a musical instrument such as a flute. It is not the precise stop which is being covered which generates the sound but the whole instument.

11. Evolutionary Epistemology

Emergence of subjectivity, in the sense proposed in this paper, requires that a 'whisper' of subjectivity is present in the elements of which the brain is composed. It is at this point that a connection might, perhaps, be made with the 'paradoxes' of quantum physics which STAPP (1993), PENROSE (1994) and others have so ably discussed.

It is likely that our deep puzzlement at quantum mechanics is due to our evolutionarily inbuilt concepts of space, time and substance. It has often been remarked that systems evolved to solve the metre-sized problems of survival on a small rocky planet are unlikely to be at home in the atto-metre subnuclear or multi-tera-metre extragalactic worlds.

In the quantum world happenings appear to influence each other as if distance did not count, the notion of a discrete particle at a definite place disappears. And, most importantly, the deep demarcation (as FERRIER puts it) between 'observer' and 'observed' (or, to use our terminology, 'experiencer' and 'experienced') on at least some accounts, also disappears. On these accounts perception is necessary to collapse the wave-function and transform 'propensities' into actualities[8]. Intuitions which have developed over half a billion years of animal evolution on planet Earth no longer serve.

The twentieth-century revolution in physics has, however, hardly affected the biologist, the molecular biologist, the biochemist. The neurobiologist does not concern himself with events at the quantum level. He accepts that the chemical bond, especially the all-important covalent bond, is quantum based, but he accepts this as given, and his attention is directed at far higher orders of magnitude. At these orders of magnitude the 'common-sensical' classical assessment seems fully adequate. Matter is still seen as Isaac NEWTON taught us to see it, as consisting, ultimately, of "solid, massy, hard, impenetrable particles". We are all familiar with the ball and stick models of the great biological molecules. These are, of course, huge and crude abstractions. When we contemplate the 'hard problem' we may wonder whether part of the problem and perhaps the major part, is that these crudities and our classical intuitions mislead us. The tentative suggestion being made in this paper is that the integrated complexity of cerebral greys allows features of matter, counter to these intuitions, to obtrude.

HALDANE famously remarked that the world is not only strange but probably stranger than we can imagine. Many have thought that the presence of mind demands an expansion of our understanding of physics. E. A. BURTT, long ago, noted that classical, Newtonian, physics simply has no place for mankind (BURTT 1932). T. L. S. SPRIGGE (1994) supporting his panpsychic thesis writes that "if this essentially panpsychic view of the inner being of matter is rejected, then I think the only alternative is some form of dualism". Does the vast integral complexity of the brain allow subjectivity, a counter-intuitive property, hidden in the quantum world to 'emerge'? Does the seeming entanglement of observer and observed which has puzzled the most eminent of quantum physicists support this panexperientalist argument? Are quantum physicists and neurophilosophers of mind perhaps catching sight of the same issue from different directions?

12. Are the Cartesian Categories Darwinian Algorithms?

Consciousness in the view adumbrated in this paper is a 'brute fact' in both senses of that phrase. We are back once again in the vicinity of Martin HEIDEGGER and the hidden problem of Being. But we have, I believe, made some progress. I would like to argue that we can begin to see that the 'hardness' of the hard problem, like the paradoxes of quantum theory, are consequences of our evolutionarily-induced epistemology.

It is not so much that quantum theory underlies exocytosis at pre-synaptic terminals and thus 'explains' our introspective freedom (ECCLES 1994) or that it accounts for microtubular information processing (PENROSE 1994; HAMEROFF 1994) but that it undermines the 'CARTESIAN categories' which have been deeply 'incorporated' (to use NIETZSCHE's terminology) into our understandings by our evolutionary history. The DARWINIAN acid (to revert to our opening quotation from DENNETT) begins to attack our confidence in that seventeenth century settlement in which 'feeling', 'sensation', *res cogitans*, is said to be one thing and matter, *res extensa*, quite another.

I am saying, in other words, that the folk-categories of objectivity and subjectivity will become less clear-cut when the full implications of the twentieth-century revolution in physics have been absorbed[9]. We shall have to come terms with a post EINSTEIN/BOHR/HEISENBERG concept of matter. We shall have to accept that experiencer and experienced are in some deep and as yet mysterious way inextricably entangled. When a recogntion of this entanglement is combined with a recognition that brains are composed of this mysterious substratum and show 'integrality' all the way down we may at last be a little closer to answering ROCKWELL's question: "what is it, amongst all the things that are, that makes brains uniquely associated with consciousness?" (ROCKWELL 1994).

13. Conclusion

In one of his better-known sayings Sherlock Holmes pointed out that "When you have eliminated the impossible, whatever remains, *however improbable*, must be the truth". It is in this spirit that this paper has been offered. It consists, of course, only of a series of tentative pointers. Pointers to directions which, it seems to me, are worth pursuing. It is motivated by monism. My objective is to point the way to a unitary theory of mind in nature. My hope is that as the sciences of matter, especially neuroscience and fundamental physics, evolve into the twenty-first century others will pursue some of the hares I have started.

My proposal is materialistic provided that the term 'materialist' is not tied rigidly to its seventeenth century connotation, but allowed to evolve with modern physics into the twentieth and twenty first centuries. It is also, it seems to me, consistent with the flow of contemporary research in evolutionary biology and, in particular, neurobiology

It restates the classical panpsychist position in taking consciousness, subjectivity, to be inherent in the world from the beginning. It suggests that our antagonism to this assessment is due to our deeply engrained palaeolithic epistemology[10]. It suggests that the insights of twentieth century physics into the nature of matter have placed that epistemology in doubt. It suggests that the coherent activity of volumes of hugely complex matter (cerebral greys) allows the emergence (E_2 emergence) of this inherent subjectivity.

It does not, of course, suggest that we shall ever be able to 'read off' the consciousness of another from an understanding of his/her material brain state. Privacy remains inviolable. We can no more experience another's subjectivity than we can *be* another's brain. An objective phenomenology remains a contradiction in terms.

But my tentative proposal does, it seems to me, provide a way of healing the deep fracture in our world-view: the palaeolithic fracture which DESCARTES clarified and codified in the seventeenth century, which puzzled DARWIN in the nineteenth, and was the source of so much argument in the twentieth.

It seems to me, to revert to my opening quotation, to provide a route towards recovering that "*lost sense of solid and substantial being*" which HEGEL mourned[11] and, to use the famous words of John DONNE in the seventeenth century, to a reassertion that, far from being "lost in a world we never made", we are in fact "all parts of the Maine".

Author's address

C. U. M. Smith, Vision Sciences, Aston University, Birmingham, B4 7ET, UK
Email: smithcum@aston.ac.uk

Notes

1 "…our psychical lives are tissues of purposes" (SHERRINGTON 1947, Introduction).

2 It is interesting to note that Charles DARWIN carefully read FERRIER's papers when they appeared in *Blackwood's Magazine* in 1838. They provided a significant challenge to the evolution theory he was endeavouring to construct during his immediate post-Beagle years (see SMITH 1978)

3 Indeed, so acrimonious has the debate been that in 1866, section 2 of the statutes of the Parisian Société de Linguistique issued a famous embargo against any further discussion: "La Société de Linguistique n'admet aucune communication concernant… l'origine du langage…" (JESPERSEN 1922, p96). Of course, as with most embargoes, this was more honoured in the breach than in the observance. Perhaps the outstanding and outstandingly accessible modern account is that of PINKER (1994).

4 It is pertinent to note that the overwhelming importance of individual recognition is mirrored in the antomy and physiology of primate brains. Physiologists have shown that the temporal cortices of macaques are populated with very precise face-recognition cells (PERRETT et al. 1992)

5 The molecular biological revolution of the latter part of our century is showing that the physiological mechanisms at work in (for example) memory are likely to be much the same from *Drosophila*, through *Aplysia* to *Mus* and *Homo* (see SMITH 1996)

6 In a review article John MADDOX observes that "Everybody wants to measure complexity but nobody quite knows how it should be done" (MADDOX 1990) whilst as long ago as 1966 VON NEUMANN remarked that "…none of this can get out of the realm of vague statement until one has defined the concept of complication correctly… there is nothing new about this… the simplest mechanical and thermodynamical systems had to be discussed a long time before the correct concepts of energy and entropy could be extracted from them" (VON NEUMANN 1966).

7 It is in this weak sense that 'life' emerges from the hugely complex interactions of molecules. 'Life' is a true system property; biologists no longer look for 'vital' forces; 'subjectivity' is, however, altogether different.

8 Obviously this is no place to discuss the various interpretations of quantum mechanics. It was VON NEUMANN in 1932 who first pointed out the seeming necessity of a conscious observer to 'collapse' the otherwise continuously evolving wave-function (VON NEUMANN 1932). This apparent entanglement of observer and observed has puzzled physicists inclined to a 'realist' metaphysics ever since. These physicists included Albert EINSTEIN. The puzzlements include SCHROEDINGER's stressed-out cat and WIGNER's schizophrenic friend. A daring speculation, in the context of the present paper, might seek to tie together the 'remembered present' which we discussed in the sections on primary and secondary consciousness with the transformation of 'propensity' or 'potentiality' into 'actuality' which BOHM (1951), HEISENBERG (1958) and POPPER (1967) see as the essence of a recording or measurement event. In other words, is it possible to see the 'objective tendencies' or 'propensities' of the quantum physicists as in some way equivalent to SARTRE's 'impersonal spontaneity' and the 'hard', 'solid' 'being' of 'secondary consciousness', the 'remembered present', as the outcome of a 'recording' or 'observation' event?

9 One is indeed reminded of the famous passage in *The Gay Science* where NIETZSCHE (1974) writes of the nineteenth century's insensitivity to the death of God: "…this tremendous event is still on its way… it has not reached the ears of man. Lightning and thunder require time, the light of stars requires time, deeds require time even after they are done, before they can be seen and heard. This deed is still more distant from them than the most distant stars—and yet they have done it themselves".

10 A comparison (and perhaps a profound one) can be made between the distaste expressed by many for this type of 'reductionism' and the reductionism implied by evolution theory. The young DARWIN, perceptive as ever, writes in the *C Transmutation Notebook* (1980, p196) that "Man in his arrogance thinks himself a great work worthy the interposition of a deity. More humble and I believe truer to consider him created from animals". He has many similar passages in his early writings and in general he rebuts the diminishment implied by reductionism by arguing that our estimate of animals should be raised: "Animals whom we have made our slaves we do not like to consider our equals… animals with affections, imitation, fear of death, pain, sorrow for the dead—respect" (1980, p231).

11 "Spirit… demands from philosophy not so much *knowledge* of what it *is*, as recovery through its agency of that lost sense of solid and substantial being… Formerly… (men)… had a heaven adorned with a vast wealth of thoughts and imagery. The meaning of all that is, hung on the thread of light by which it was linked to that heaven. Instead of dwelling in this world's presence, men looked beyond it, following this thread to an otherworldly presence, so to speak. The eye of the Spirit had to be forcibly turned and held fast to the things of this world; and it has taken a long time before the lucidity which only heavenly things used to have could penetrate the dullness and confusion in which the sense of worldly things was enveloped… Now we seem to need just the opposite: sense is so firmly rooted in earthly things that it requires just as much force to raise it…" (HEGEL 1977, Preface, paragraphs 7–8).

References

Alexander, R. D. (1987) The Biology of Moral Systems. Aldine: New York.

Block, N. (1995) On a confusion about a function of consciousness. Behavioural and Brain Sciences 18: 227–247.

Block, N. (1996) Nature of Consciousness. Wiley: Chichester.

Blasdell, G. G./Salama, G. (1986) Voltage sensitive dyes reveal a modular organisation in monkey cerebral cortex. Nature 321: 549–555.

Bohm, D. (1951) Quantum Theory. Prentice Hall: New York.

Bonner, J. T. (1980) The Evolution of Culture in Animals. Princeton University Press: Princeton.

Bunge, M. (1979) Some topical problems in biophilosophy. J. Social Biol. Struct. 2: 155–172.

Burgess, J. A./Tawia, S. A. (1996) When did you first begin to feel it–locating the beginning of human consciousness. Bioethics 10: 1–26.

Burtt, E. A. (1932) The Metaphysical Foundations of Modern Physical Science. Routledge and Kegan Paul: London.

Chalmers, D. J. (1995) Facing up to the problem of consciousness. Journal of Consciousness Studies 2: 200–219.

Chalmers, D. J. (1996) The Conscious Mind. Oxford University Press: New York.

Crick, F. H. C. (1994) The Astonishing Hypothesis. Simon and Schuster: London.

Darwin, C. R. (1980) M Notebook. In: Barrett, P. H./Gruber, H. E. (eds) Metaphysics, Materialism and the Evolution of Mind: Early Writings of Charles Darwin. (Originally published in 1838). University of Chicago Press: Chicago, pp. 6–37.

Darwin, C. R. (1980) Transmutation Notebooks. In: Barrett, P. H./Gruber, H. E. (eds) Metaphysics, Materialism and the Evolution of Mind: Early Writings of Charles Darwin. (Originally published in 1838–39). University of Chicago Press: Chicago, pp. 179–200.

Dawkins, R. (1976) The Selfish Gene. Oxford University Press: Oxford.

De Quincy, C. (1995) Consciousness all the way down. Journal of Consciousness Studies 1: 217–229.

Demske, J. M. (1970) Being, Man and Death. University Press of Kentucky: Lexington KY.

Denbigh, K. G. (1975) Non-Conserved Function for Organized System. In: Kubat, L./Zeman, J. (eds) Entropy and Information in Science and Philosophy. Elsevier: Amsterdam, pp. 83–92.

Dennett, D. (1995) Darwin's Dangerous Idea. Simon & Schuster: New York.

Desmedt, J. E./Tomberg, C. (1994) Transient phase-locking of 40Hz electrical oscilations in pre-frontal and parietal human cerebral cortex reflects the process of conscious somatic perceptions. Neuroscience Letters 168: 126–129.

Devinsky, O./Morrell, M. J./Vogt, B. A. (1995) Contributions of anterior cingulate gyrus to behaviour. Brain 118: 279–306.

Dobzhansky, T. (1973) Nothing in biology makes sense except in the light of evolution. American Biology Teacher 35: 125–129.

Driscoll, M./Chalfie, M. (1991) The mec-4 gene is a member of a family of Caenorhabditis elegans genes that can mutate to induce neuronal degeneration. Nature 349: 588–593.

Dunbar, R. I. M. (1993) Coevolution of neocortex size, group size and language in humans. Brain and Behavioural Sciences 16: 681–735.

Eccles, J. C. (1994) How the Self Controls Its Brain. Springer-Verlag: Berlin.

Edelman, G. (1989) The Remembered Present: A Biological Theory of Consciousness. Basic Books: NewYork.

Felleman, D. J./van Essen, D. (1991) Distributed hierarchical processing in the Primate cerebral cortex. Cerebral Cortex 1: 1–47.

Ferrier, J. (1838) An introduction to the philosophy of consciousness. Blackwood's Magazine 43: 187–201.

Galen (1952) On the Natural Faculties, Book 2,1,3. The Loeb Classical Library. Heinemann: London. Latin original: De Naturalibus Facultatibus. Translated by J. A. Brock

Gallup, C. G. (1970) Chimpanzee self-recognition. Science 167: 86–87.

Gallup, C. G. (1979) Self-awareness in Primates. American Scientist 67: 417–421.

Geschwind, N. (1979) Specialisations of the human brain. Scientific American 241(3): 158–168.

Hameroff, S. (1994) Quantum coherence in microtubules: a neural basis for emergent consciousness. Journal of Consciousness Studies 1: 91–118.

Hamilton, W. D. (1964) The genetical theory of social behaviour. Journal of Theoretical Biology 7: 1–52.

Hegel, G. W. F. (1977) Phenomenology of the Spirit. Oxford University Press: Oxford. German original: (1807) Phänomenologie des Geistes. Translated by A. V. Miller.

Heidegger, M. (1962) Being and Time. Blackwell: Oxford. German original: (1927) Sein und Zeit. Translated by J. Macquarrie and E. Robinson.

Heisenberg, W. (1958) Physics and Philosophy. Harper and Row: New York.

Huang, M./Chalfie, M. (1994) Gene interactions affecting mechanosensory transduction in Caenorhabditis elegans. Nature 367: 467–470.

Humphrey, N. (1979) Nature's Psychologists. In: Josephson, B./Ramachandran, B. S. (eds) Consciousness and the Physical World. Pergamon: New York.

Humphrey, N. (1983) Consciousness Regained: chapters in the development of mind. Oxford University Press: Oxford.

Hocking, C. F. (1973) Man's Place in Nature. McGraw Hill: New York.

Jacobi, F. (1811) Brief an Fichte. (Quoted in: Caton, H. (1973) The Origins of Subjectivity. Yale University Press: New Haven, p. 95).

James, W. (1890) Principles of Psychology. New edition 1950. Dover: New York.

Jespersen, O. (1922) Language: Its Nature, Development and Origin. Allen and Unwin: London.

Koch, C. (1996) A neuronal correlate of consciousness. Current Biology 6: 492–495.

Kummer, H. (1968) Social Organisation of Hamadryas Baboons. Karger: Basel.

Libet, B. (1985) Unconscious cerebral initiative and the role of conscious will in voluntary action. Behaviour and Brain Sciences 8: 529–539.

Llinas, R./Ribary, U. (1993) Coherent 40Hz oscillation characterises dream state in humans. Proceedings of the National Academy of Sciences 90: 2078–2081.

Llinas, R./Paré, D. (1996) The Brain as a Closed System Modulated by the Senses. In: Llinas, R./Churchland, P. S. (eds) The Mind–Brain Continuum. MIT Press: Cambridge MA and London.

Linas, R./Ribary, U./Contreras, D. (1998) The neuronal basis for consciousness. Philosophical Transactions of the Royal Society B 353: 1841–1849.

Logothetis, N. K./Schall, J. D. (1989) Neuronal correlates of subjective visual perception. Science 245: 761–763.

Logothetis, N. K. (1998) Single units and conscious vision. Philosophical Transactions of the Royal Society B 353: 1801–1818.

Maddox, J. (1990) Complicated measures of complexity. Nature 344: 705.

Mayr, E. (1982) The Growth of Biological Thought. Harvard University Press: Cambridge MA.

Milner, A. D. (1995) Cerebral correlates of visual awareness. Neuropsychologica 33: 1117–1130.

Nagel, T. (1974) What is it like to be a bat? Philosophical Review 4: 435–50.

Nietzsche, F. W. (1974) The Gay Science. Vintage Books: New York. German original (1882) Die fröhliche Wissenschaft. Translated by W. Kaufmann.

Paré, D./Llinas, R. (1995) Conscious and pre-conscious processes as seen from the standpoint of sleep-waking cycle neurophysiology. Neuropsychologica 33: 1155–68.

Penrose, R. (1994) Shadows of the Mind. Oxford University Press: Oxford.

Perrett, D. I./Hietanan, J. K./Oram, M. W./Benson, B. J. (1992) Organisation and functions of cells responsive to faces in the temporal cortex. Philosophical Transactions of the Royal Society B 335: 23–30.

Pinker, S. (1994) The Language Instinct. Penguin: Harmondsworth.

Popper, K. (1967) Quantum Physics without "The Observer". In: Bunge, M. (ed) Quantum Theory and Reality. Springer: Berlin.

Popper, K. (1974) Scientific Reduction and the Essential Incompleteness of All Science. In: Ayala, F. J./Dobzhansky, T. (eds) Studies in the Philosophy of Biology. Macmillan: London, pp.259–282.

Povinelli, D./Preuss, T. (1995) Theory of Mind: evolutionary history of a cognitive speciality. Trends in Neurosciences 18: 418–24.

Premack, D./Woodruff, G. (1978) Does the chimpanzee have a theory of mind? Brain and Behavioural Sciences 4: 515–26.

Rockwell, W. T. (1994) On what the mind is identical with. Philosophical Psychology 7: 307–323.

Roland, P. E./Gulyas, B. (1994) Visual imagery and visual representation. Trends in Neurosciences 17: 281–289.

Sacks, O. (1986) The Man who Mistook his Wife for a Hat. Picador: London.

Sartre, J.-P. (1957a) Being and Nothingness. Methuen: London. French original: (1943) L'Être el le Néant. Translated by H. E. Barnes.

Sartre, J.-P. (1957b) The Transcendence of the Ego. Noonday: New York. French original: (1936) La Trascendance de l'Ego. Translated by F. Williams and R. Kirkpatrick.

Sherrington, C. S. (1947) Integrative Action of the Nervous System, 2nd edition. Cambridge University Press: Cambridge.

Singer, W. (1996) Neuronal Synchronisation: A Solution to the Binding Problem. In: Llinas, R./Churchland, P. S. (eds) The Mind–Brain Continuum. MIT Press: Cambridge MA and London.

Smith, C. U. M. (1978) Charles Darwin, the origin of consciousness and panpsychism. Journal of the History of Biology 11: 245–267.

Smith, C. U. M. (1983) Anatomical concepts and the problem of mind. Journal of Social and Biological Structures 6: 381–392.

Smith, C. U. M. (1987) "Clever beasts who invented knowing": Nietzsche's evolutionary biology of knowledge. Biology and Philosophy 2: 65–91.

Smith, C. U. M. (1994) The Complexity of Brains: A Biologist's View. In: Stonier, R. J./Yu, X. H. (eds) Complex Systems: Mechanisms of adaptation. IOS Press: Amsterdam, pp.93–100.

Smith, C. U. M. (1996) Elements of Molecular Neurobiology (2nd ed.).Wiley: Chichester.

Smith, C. U. M. (1998) Descartes' pineal neuropsychology. Brain and Cognition 36: 57–72.

Smith, J. M./Szathmary, E. (1995) The Major Transitions in Evolution. W. H. Freeman: Oxford.

Sprigge, T. L. S. (1994) Consciousness. Synthese 98: 73–93.

St. Augustine (1983) The Trinity. Catholic University of America Press: Washington DC. Latin original: De Trinitate. Translated by S. McKenna.

Stapp, H. P. (1993) Mind, Matter and Quantum Mechanics. Springer-Verlag: Berlin.

Sternberg, D. L./Logothetis, N. K. (1997) The role of temporal cortical areas and perceptual organisation. Proceedings of the National Academy of Sciences 94: 3408–3413.

Szentagothai, J. (1978) The neuron network of the cerebral cortex: A functional interpretation. Proceedings of the Royal Society B 201: 219–248.

T'so, D. Y./Frostig, R. D./Lieke, E. E./Grinvald, A. (1990) Functional organisation of Primate visual cortex revealed by high resolution optical imaging. Science 249: 417–420.

von Neumann, J. (1932) Mathematical Foundations of Quantum Mechanics. Princeton University Press: Princeton.

von Neumann, J. (1966) Theory of Self-reproducing Automata (Edited and completed by W. A. Burks). University of Illinois Press: Urbana IL.

Watts, D. J./Strogatz, S. H. (1998) Collective dynamics of "small-world" networks. Nature 393: 440–442.

Watts, D. J. (1999) Small Worlds: The Dynamics of Networks between Order and Randomness. Princeton University Press: Princeton.

Whitehead, A. N. (1978) Process and Reality: An Essay in Cosmology. Cambridge University Press: Cambridge.

Wilson, E. O. (1975) Sociobiology: The New Synthesis. Harvard University Press: Cambridge MA.

William F. Flack, Jr.

Emotional Self-Perception, Social Adaptation, and Interpersonal Disruption

A Functional Perspective on Severe Psychopathology

Abstract

The aim of this article is to explain the role of the relationship between emotional expression and emotional experience in an interpersonal perspective on severe psychopathology. The link between emotional expression and experience is understood as a process of self-perception, and it is assumed to malfunction in persons with psychotic disorders. This malfunction is one source of disruption during social encounters, which results in the identification of certain persons as abnormal. Findings from studies of emotional judgments, and of the effects of expressive behavior on experience, in both normals and psychotics are summarized.

Key words

Emotion, self-perception, psychopathology, social interaction.

After publishing his magnum opus on the theory of evolution in 1859, Charles DARWIN sought to buttress his arguments about the continuity of human beings and other animals by using emotional expression as an example. As is so often the case in works of scientific genius, DARWIN's treatise, The Expression of the Emotions in Man and Animals (1998)[1], is replete with ideas that would later prove to be of considerable heuristic value in a variety of fields of investigation. In this chapter, I will outline a functional perspective on severe psychopathology (e.g., FLACK et al. 1998) in which the central concepts are predicated on evolutionary theory. Levels of analysis included in this thinking are both psychological and microsociological. Central to this perspective is the adaptive interplay of behavioral and subjective phenomena in the context of social encounters. Within this context, concepts of particular importance are the meaning of emotion, the self-perception of emotion, and interpersonal adaptation and disruption, concepts each of which can be traced, at least in nascent form, to DARWIN's work.

The first section of this chapter is devoted to the relationship between emotional expression and emotional experience. William JAMES' (1884) theory, an early version of which is presaged in DARWIN (1998), is the starting point for an account of this relationship. The cognitive component of this relationship is highlighted in the second section of the chapter. This component is the process of self-perception (BEM 1967), which is used by LAIRD (e.g., 1984) to explain the relationship between emotional behavior and emotional feeling. In the third and fourth sections of the chapter, an interpersonal perspective on normality and psychopathology, first developed by MILLER (1963), is summarized. As this perspective has been elaborated (e.g., FLACK et al. 1998; FLACK/MILLER 1992; MILLER 1983; MILLER/JAQUES 1988), the relationship between emotional expression and self-perception has assumed an increasingly important role. The chapter ends with a summary of illustrative data that have been collected in descriptive and experimental studies of nonverbal communication and of emotional expression and experience in normal and psychiatric groups.

The Meaning of Emotion: Emotional Expression and Emotional Experience

"The free expression by outward signs of an emotion intensifies it. On the other hand, the repression, as far as this is possible, of all outward signs softens our emotions... Even the simulation of an emotion tends to arouse it in our minds." (DARWIN 1998, pp359–360).

These observations, regarding the relationship between the expression and the experience of emotion in human beings, are found at the very end of DARWIN's book on emotion. Like the concept of evolution, this idea, that expressive behavior causes changes in emotional feelings, was not new; indeed, DARWIN cites MAUDSLEY, WUNDT, BRAID, and GRATIOLET as making similar observations. DEWEY (1894, 1895) conjectures that the roots of this idea date back to ARISTOTLE, although he cites HEGEL specifically.[2] Whoever the original progenitor, the idea that changes in expression lead to changes in feeling is promoted to the status of theory by William JAMES (1884, 1890), in his famous conception of emotion. JAMES attempts to give a functional explanation of subjective emotional experience by locating the cause of emotional feeling squarely within the physical body (MYERS 1986). He asserts:

"Our natural way of thinking about these standard emotions is that the mental perception of some fact excites the mental affection called emotion, and that this latter state of mind gives rise to the bodily expression. My thesis on the contrary is that the bodily changes follow directly the *perception* of the exciting fact, and that our feeling of the same changes as they occur IS the emotion... Without the bodily states following on the perception, the latter would be purely cognitive in form, pale, colorless, destitute of emotional warmth." (JAMES 1890, pp449–450, emphases in the original).

JAMES' theory[3], in a nutshell, runs something like this. First, there is a perceptible stimulus that leads to excitation. The perception of this stimulus is the initial phase of an emotional episode. The second phase consists of a bodily response which, at least in the case of what JAMES calls the 'coarser' emotions (e.g., anger, fear, disgust, surprise), is instinctive. This part of the theory is a direct descendant of DARWIN's notions about the adaptive qualities of emotional expression. The third phase is actually contemporaneous with the second; that is, emotional feeling is the subjective experience of what one's body is doing when behaving emotionally. Without some bodily arousal, there is, for JAMES, no opportunity for emotional experience. Thus, JAMES takes DARWIN's handful of observations about a potential relationship between emotional expression and feeling, and makes the much stronger argument that expressive behavior is a requisite of emotional experience.

Although JAMES and his theoretical followers have been criticized for emphasizing the physical characteristics of emotion at the expense of cognitive ones, this criticism is unjustified. In the first place, one must be cognizant that the stimulus that sets off an emotional episode is something about which one should become excited. Therefore, judgment enters into the process at the very outset; emotional responses are set off by the "making and further processing of sensory differences" (WIMMER/CIOMPI 1996). In addition, although implicit in JAMES' account, one must also be able to identify which emotion is being enacted. "Pure arousal", if such a thing were possible, would be maladaptive in the same way that "free-floating anxiety" or any other chronic mood state is; one becomes unable to act adaptively during these states, in part, because there is no information value to the emotional experience. Even in the case of a "fight–or–flight" response, one must be able to choose which of the alternatives is most likely to result in adaptive success. Furthermore, the functional value of the subjective aspect of emotion is that it provides an ongoing 'readout' of information about what one is doing, so that a course of further action may be chosen and carried out (DEWEY 1894, 1895; LAIRD/APOSTOLERIS 1996). In short, on JAMES' account, cognitive distinctions about the environment and about internal feeling states are required for a meaningful experience of emotion. DEWEY (1894, 1895) goes further, stating that all such aspects of an emotional episode are experienced as an organic whole, and not in the sort of step–by–step fashion seemingly implied by JAMES' account. As scientists who often find ourselves pulling apart complex phenomena as one means of understanding them, we would do well to bear DEWEY's point in mind.

Self-Perception of Emotion

The JAMESIAN theory of emotion has enjoyed something of a renaissance among social scientists during the past thirty years (CORNELIUS 1996). The resurgence of interest in this theory can be traced directly to the influence of psychologist Silvan TOMKINS (e.g., 1963). TOMKINS agrees with JAMES that expressive behavior is the proximal cause of emotional experience. More specifically, TOMKINS focuses on the role of the facial musculature, hypothesizing that our ability to distinguish among numerous emotional states is made possible by the precision and specificity of expressive movements of the face. TOMKINS' influence can be gauged, in part, by the extent to which subsequent theoretical and empirical work on this topic continues to be referred to under the rubric of the "facial feedback" hypothesis (MCINTOSH 1996).

LAIRD (1974) was among the first to subject the JAMESIAN/TOMKINSIAN theory to experimental verification. The crux of his method, which has been used in various permutations by many other investigators, is to manipulate an emotional behavior, and then to measure the resulting emotional experience. If JAMES and TOMKINS are correct, then the production of specific sets of muscle movements (i.e., facial expressions) should elicit changes in emotional feelings. In fact, this is exactly what happens. The results of over fifty published studies are consistent in demonstrating that manipulations of emotional facial expressions cause reliable changes in subjective emotional experience (see reviews of the literature by, for example, ADELMAN/ZAJONC 1988; CAPELLA 1993; CORNELIUS 1996; IZARD 1990; LAIRD 1984; MCINTOSH 1996).[4] When subjects in these experiments are instructed to adopt a facial expression of anger, for example, they report feeling angry, whereas the adoption of a sad expression leads to increased feelings of sadness. More recently, manipulations of bodily postures (DUCLOS et al. 1989; FLACK/LAIRD/CAVALLARO 1999b; STEPPER/STRACK 1993) and vocal expressions (SIEGMAN/BOYLE 1993) have been shown to have similar effects on emotional feelings. Thus, in each modality of emotionally expressive behavior that has been tested, evidence has been found for the effect proposed by JAMES.

Critics of this work have pointed to a number of problems. Specifically, they have complained that the results of such studies are attributable to experimental demand characteristics, and that they are also both nonspecific and small in magnitude (MATSUMOTO 1987; TOURANGEAU/ELLSWORTH 1979; WINTON 1986). Each of these criticisms has been addressed in recent studies. In order to address the role of demand characteristics, a number of investigators (DUCLOS et al. 1989; STRACK/MARTIN/STEPPER 1988) have employed cleverly disguised procedures for producing expressions, and have obtained results similar to those found in earlier studies using undisguised methods. More recently, my students and I (FLACK et al., in preparation) have shown that subjects' awareness of the purpose of expression manipulations makes no difference in the effects of facial expressions on feelings. The first studies in this area demonstrated that manipulated expressions could produce changes in feelings, but these early results addressed only global distinctions between pleasant and unpleasant feelings. The results of subsequent research have demonstrated that specific expressions cause changes in equally specific feelings (DUCLOS et al. 1989; FLACK/LAIRD/CAVALLARO 1999a; FLACK et al. 1999); facial expressions of anger, for example, lead to increases in feelings of anger, but not in feelings of sadness, fear, or happiness. The criticism regarding magnitude sizes of expression effects has merit; that is, manipulations of facial expressions tend to cause relatively small changes in the ratings subjects make of their feelings. This can be explained by JAMESIAN theory, however, in that only single expressive modalities (i.e., facial expressions) have been manipulated in most of these studies. Presumably, the manipulation of multiple, matching, and simultaneous expressive modalities (e.g., facial expressions and bodily postures) would cause greater magnitudes of change in feelings than individual expressions alone. In fact, this is precisely what was found in another of our recent investigations (FLACK/LAIRD/CAVALLARO 1999b).

Thus, it is clear that, at the very least, a weak version of JAMES' theory of emotion (i.e., emotional expressions *can* cause emotional feelings) is supported by current research. But how to explain the link between expression and experience? LAIRD (e.g., 1984) has proposed that the connection is a cognitive one, namely, the process of self-perception. At the center of self-perception theory (BEM 1967) is the notion that we come to understand ourselves in the same way that we come to understand others—that is, by observing our own behavior. In the case of *self*-perception, this process can be automatic and unconscious, as it is in most cases of emotional response. The self-perception of emotional conduct provides the cognitive link between emotional behavior and emotional feeling; we are able to discern our own, discrete emotional states because we are able to perceive our own emotional behavior. The subjective experience, or feeling, of emotion is made possible by the fact that we are able to discern accurately what our bodies are doing during emotional episodes. Our ability to do so makes possible the subsequent planning of an adaptive course of action, which helps to explain the functional significance of emotional feeling (DEWEY 1894, 1895; LAIRD/APOSTOLERIS 1996). Additional aspects of the adaptive value of emotional experience and expression are evident when considered in the context of social interactions, the subject of the next section.

Interpersonal Adaptation

"The movements of expression in the face and body, whatever their origin may have been, are in themselves of much importance for our welfare. They serve as the first means of communication... We readily perceive sympathy in others by their expression... and mutual good feeling is thus strengthened." (DARWIN 1998, p359).

Here again, DARWIN is pointing to an aspect of emotional behavior, namely, its interpersonal, communicative value, which would later be the subject of much productive theorizing and empirical research. Emotional expression and, more broadly, nonverbal behavior is a vital component in the process of social interaction. At a still more general level of analysis, the adaptive significance of social interactions is highlighted in the thinking of anthropologist David ABERLE and colleagues (e.g., ABERLE et al. 1950). These workers point out that survival ultimately requires cooperative behavior within groups. Members of the human group must engage in certain kinds of behaviors with their fellows that will enable the group to continue over time. ABERLE et al. (1950) refer to these activities as the functional prerequisites of a society. Functional social prerequisites include activities such as procreation, socialization of the young, distribution of goods and services, and the regulation of deviance.

Success at fulfilling the functional prerequisites requires the successful carrying out of face–to–face interactions between individual members of a social group. Such interactions are themselves comprised of complex structures, which have been studied by, among others, psychologists DUNCAN, FISKE, and their colleagues (e.g., DUNCAN/FISKE 1985). These investigators have developed finely detailed descriptions of interactive structures, one of which is the conversational turn–taking system. In this system, verbal and nonverbal behaviors function as signals that are exchanged by interacting partners. More specifically, nonverbal behaviors, such as changes in vocal pitch, hand gestures, and gaze direction are timed precisely in order to indicate the beginning, continuance, and ending of speaking turns. Partners to an interaction know when it is appropriate to take over or relinquish the conversational floor by paying attention to such signals. Of course, such 'attention' is usually automatic, so that we only notice the signals and their timing when they are either absent or are used in unusual ways.

Similarly, interaction partners come to understand each other's perspectives, including especially their emotional states, by perceiving and then mimicking each other's expressive behaviors. MEAD (1934, 1982) argues that this is part of the process by means of which human beings develop a sense of self. It also becomes, over the course of ontogenesis, the means by which we are able to empathize with others. Empathy is made possible not only by being cognizant of, or decoding, the emotional expressions of others, but also by mirroring, or adopting in attenuated form, those expressions as well. The purpose, as MEAD points out, is to enable the sharing of subjective experience, since similar expressions among social partners should, on the JAMESIAN theory, elicit similar internal states (see also CAPELLA 1993, and HATFIELD/CACCIOPPO/RAPSON 1994 for similar views). This mutuality of expression and experience lends a sense of shared understanding to both partners in social situations.

The sharing of mutual perspectives and interpersonal understanding are required for the development of human relationships in any domain of activity. Our abilities to use our own and others' nonverbal behaviors in taking turns at speaking and listening enables us to communicate in ways that lead to the successful outcomes of our interactions. Our capacities to understand our own and others' emotions through the mechanisms of mimicry and peripheral feedback adds a dimension of warmth and cohesion that makes for the kind of "mutual good feeling" referred to by DARWIN. Each of these factors is crucial in understanding the qualities of adaptive interpersonal behavior. When these qualities break down, when our ability to carry out the most basic of social interactions becomes disrupted, we confront phenomena judged as psychopathological.

Interpersonal Disruption: A Functional View of Severe Psychopathology

MILLER (1963; MILLER 1983; MILLER/JAQUES 1988) has proposed, and later MILLER and I (FLACK/MILLER 1992) have elaborated upon, a perspective on the meaning of psychopathology that is predicated on the notion of interpersonal adaptation. This line of thinking has been offered as an alternative to the atheoretical manner in which psychopathology is usually defined[5] at present (e.g., American Psychiatric Association 1994). We begin at the level of the social group, because this is the level at which psychopathology has meaning; that is, psychopathology has functional significance, ultimately, not because of abnormal biochemistry or brain physiology, but because of its interpersonal consequences. From our perspective, normal behavior (including action, thinking, and feeling) is defined as conduct that enables members of any social group to fulfill the functional prerequisites of their society. More specifically, psychosocial normality is conceived as interpersonal activity that makes possible the successful carrying out of social encounters. Abnormality, by contrast, is defined as conduct that interferes with the working of such relations. That is, psychosocial abnormality, or psychopathology, is seen, from this perspective, as behavior that causes the breakdown of social interactions. Such a conception helps to explain why certain forms of conduct come to be viewed not only as unusual, but also as worthy of notice and intervention; activity that makes impossible the most basic forms of interaction required for survival is a concern in any society. That is why all societies have developed means of dealing with social behavior that is severely disruptive.

Sometimes the sources of interpersonal disruption are obvious. One example is the deterioration of grammar in verbal conversation, referred to as 'word salad', observed in some people diagnosed as chronic schizophrenics. Other examples are the withdrawal and wholesale shutting down of social behavior seen in some catatonic states. However, in our clinical experience, the most common sources of interpersonal disruption are far more subtle, although nonetheless destructive to social welfare. These sources are, we believe, located primarily in the nonverbal channel of communication, a category of activity that includes both conversational signals and emotional expressions. In short, it is not so much what people say, but rather how they say it, that causes them to get into trouble, and hence to be identified as different or strange, abnormal or deviant, and ultimately, psychopathological.

The misuse of nonverbal signals occurs on a continuum, so that some instances are relatively minor. Others, however, are so severe that they lead to a complete breakdown in communication, and thus require special identification. Nonverbal behavior that is extremely deviant is often characteristic of people identified as psychotic.[6] Interactions with psychotic individuals are difficult because they misuse nonverbal behaviors that are required for the smooth functioning of social encounters. Thus, although the beginning of a conversation usually requires mutual verbal greetings and gestures (handshakes, social smiles), the psychotic may start speaking as though such formalities are not required. During the course of an interaction in which the partners are standing, the psychotic may violate culture-specific norms with respect to interpersonal distance, standing too close (symbiosis), too far away (withdrawal), or even alternating distances (ambivalence). The rate of speech may be too slow, or its volume too loud. Gestures may be too frequent or nonexistent. Certain facial signals may be missing, or unavailable to the social partner when the psychotic looks down or away too much.

Psychoanalysts (e.g., ROSENFELD 1987) have observed that they often feel uncomfortable, perplexed, and even angry when conversing with psychotic individuals, although the contents of their verbal communications do not give ample reason for such negative reactions. One reason for those untoward responses is that psychotics fail to keep up their end of the implicit interactive bargain. That is, they may use the nonverbal signals that regulate conversation in unusual ways, at abnormally low frequencies or in unexpected sequences of occurrence. Either form of misuse throws off the normal partner to a conversation, who is left not knowing when to start or stop speaking. The synchronous flow of verbal conversation and nonverbal behavior that usually develops between normal partners is replaced by a fitful, choppy quality in interactions with a psychotic. Some clinicians refer to their own negative reactions to such patients as 'countertransference', denoting the unconscious origins of those responses. The reason that normal partners cannot identify such sources of disruption is that they are usually not attended to. That is, the conversational regulators that DUNCAN and others have identified function automatically, outside of our awareness, and much like the series of actions that are used in driving a car or riding a bicycle. Thus, when the regulators are used in abnormal ways, the encounter becomes disrupted, but the partners are unable to find the cause and address it. The result, often enough, is feelings of frustration and perplexity on both sides, and ultimately the utter collapse of communication.

In much the same way, problems with emotional expression and experience in psychosis can also disrupt social encounters. Mutuality of emotional expression is usually assumed by interacting partners to indicate mutuality of emotional experience. Individuals usually assume that if a partner is smiling, then they are also feeling happy, and the latter has the right to make the same assumption about the former. However, this may not be the case in interactions with psychotics. KRAEPELIN (1917) and Eugen BLEULER (1950) were among the first to observe that psychotics' emotional expressions are often inconsistent with self-reports of their internal experiences, a phenomenon referred to as "inappropriate affect". Thus, a psychotic may smile while experiencing murderous impulses toward a conversational partner. Without the usual signal denoting such impulses, namely an expression of anger or rage, the normal partner has no way of knowing what is, in fact, going on in the subjective experience of the psychotic.

Psychotics may also fail to mirror the emotional expressions of their normal conversational partners. Rather than adopt a social smile at the outset of an interaction, they may instead remain expressionless, a phenomenon called "affective flattening". A lack of expression throws the normal individual into another set of difficulties; they do not know how the psychotic is feeling because they do not know what a lack of expression denotes. Attempts on the part of a normal partner to mirror a flat expression are likely to result in a general dampening of feeling (LAIRD/BRESLER 1992)

Unfortunately, the dampening of expression in psychosis can sometimes be attributed to the action of antipsychotic medications. According to CIOMPI (1988, 1997), such medications appear to act primarily by shutting down the emotional response system. One reason that these medications 'work' is that they reduce the emotional arousal often associated with cognitive symptoms, thereby reducing the symptoms themselves but also causing a diminution of emotional responsiveness. In effect, a treatment that is designed to help patients to communicate in a normal manner may, in fact, make it more difficult for them to communicate normal emotions (see GRANT 1972 for a similar view).

We believe that such disruptions in communication constitute a continuum of severity in psychopathology, and that such a conception is meaningful across cultures and historical periods. However, the specific forms of disruptive behavior are likely to differ across societies. At a universal level of analysis, all societies require the sorts of successful encounters that enable fulfillment of the prerequisite social functions. However, the normative rules for interpersonal interaction will vary in different societies, so that violations of those rules, and hence the forms taken by psychosis in any given group, are culture-specific.

Nonverbal Behavior and Emotional Self-Perception in Severe Psychopathology: Illustrative Findings from Studies of Schizophrenia and Depression

My colleagues and I have conducted a number of studies of nonverbal social behavior and emotion in people diagnosed with psychotic and nonpsychotic disorders in order to test aspects of the thinking contained in previous sections of this chapter. Some of the results of these studies are summarized in this final section of the chapter, in order to provide evidence for the propositions outlined previously.

In one series of studies (FLACK/MILLER 1992; FLACK/CAVALLARO/MILLER 1999) devoted to the examination of nonverbal behavior, we videotaped adults diagnosed with schizophrenia while they were interacting with normal partners. Interactions with patients were compared with similar interactions between two normal adult partners. The conversations were unstructured, consisting mostly of the sort of talk[7] that adults engage in when getting acquainted, and lasted five minutes. The videotapes were observed, and computer-aided codings were made of specific vocal behaviors (speaking turns, brief verbal responses) and nonverbal behaviors (head nods, gestures, changes in gaze direction). These data were then analyzed to look for differences between the two sets of interactions. The results of statistical analyses revealed that the schizophrenic group used their nonverbal behaviors at significantly lower frequencies than did normal partners in either set of interactions. Further analyses indicated that the schizophrenic group also used nonverbal behaviors in different sequences, relative to the beginnings and endings of speaking turns, than did the normal subjects. Both sets of findings are consistent with our position on severe psychopathology; that is, adults with schizophrenia tend to use their nonverbal behaviors in ways that are likely to disrupt their interactions with others. Reductions in the numbers of nonverbal behaviors shown by schizophrenics mean that the sorts of signals upon which social interaction depends are frequently missing in their encounters with others. In addition, the misplacement of those nonverbal behaviors that do occur, relative to the onset and offset of turns at speaking, is likely to cause perplexity on the part of normal partners, who do not know whether to start or stop talking. These initial studies were conducted with inpatients; a more recent investigation of outpatients with schizophrenia or depression is currently in the phase of videotape analysis.

We have also conducted experiments on emotional expression, experience, and understanding in schizophrenia and depression. In the first of these investigations (FLACK/LAIRD/CAVALLARO 1999a), we examined the effects of manipulating facial expressions and bodily postures on subsequent emotional feelings in two separate studies, one using normal adults, and the other outpatient male military veterans diagnosed with schizophrenia, depression, or nonpsychiatric medical conditions. The findings from the normal sample were consistent with those of previous studies of normal subjects; facial expressions of anger, sadness, fear, happiness, and disgust caused increases in feelings matching the expressions; and bodily postures of anger, sadness, and fear resulted in similar patterns.[8] Results from the clinical groups were strikingly different. While the schizophrenic group was appropriately responsive to their facial expressions of sadness, fear, and happiness, they reacted abnormally to all bodily postures except that for anger. In fact, the depressive and nonpsychiatric groups also showed unusual patterns of response. Depressives tended to respond normally only to their facial expressions and bodily postures of sadness, whereas the normal patient group responded normally only in their expressions and postures of anger. Findings for the schizophrenic group indicate that their emotional feelings are likely to match their facial expressions in at least some categories of emotion, but that this is less likely to be the case with respect to their bodily postures. Instances of social expression that fall into the latter group are likely to be a significant source of disruption during interpersonal encounters, for reasons already outlined. The results in the depressive group indicate that they, too, are likely to experience a disconnection between their expressive behavior and their emotional feelings, except in the case of sadness. This finding may imply that the emotional problem in depression is not the experience of sadness per se, but rather that depressives do not experience any other emotions, even when helped to express them. COYNE and colleagues (e.g., JOINER/COYNE 1999) has written extensively on the ways in which depressives' behavior causes negative reactions in their social partners; a failure in the normal range of expressive behavior to lead to normal feelings may be another source of such reactions.

The second of our clinical investigations on emotion (FLACK et al. 1997) was focused on the adoption ('encoding') and understanding ('decoding') of facial expression in schizophrenia and depression. There are numerous studies of schizophrenics' emotional decoding in the literature (see BELLACK et al. 1993, for a review), all of which have found psychotics to be lacking in this ability relative to normal and psychiatric control groups. A much smaller number of studies have been devoted to emotional encoding in this disorder, but their results have been similarly discouraging. In stark contrast to such previous findings, our results revealed that schizophrenics and depressives are able to encode and to decode emotional facial expressions of anger, sadness, fear, happiness, disgust, and surprise. What might account for the difference between our findings and those of other investigators? The most likely candidate is the method we used to have subjects and observers make ratings of the expressions. In the case of the decoding task, subjects were asked to say how much of each of six different emotions was being expressed. Unlike previous research, in which subjects were forced to choose one emotion descriptor from among a list, our method is much closer to the type of distinctions among multiple emotions that people make in their encounters with others. In the encoding task, subjects were asked to adopt deliberately the six different expressions, which were videotaped and then rated by a group of three observers. The observers' ratings were made using the same procedure and scale used by the subjects during the decoding procedure, with the same result; schizophrenic, depressive, and medical control groups were all successful in adopting the six facial expressions.

Thus, it appears that while a disconnection between emotional expression and emotional experience is a likely source of interpersonal disruption in people who have been identified as either schizophrenic or depressive, the capacities to understand others' expressions and to adopt deliberately a set of expressions that are universally understood is not. One of the major limitations of the research to date is that we have examined the expression–experience relationship in solitary situations. In order to understand more fully how disruptions of this relationship are associated with disruptions of social interactions, we are currently focusing on instances of emotional expression and experience as they occur in the course of interpersonal encounters.

Author's address

William F. Flack, Jr., Department of Psychology, Bucknell University, Lewisburg, Pennsylvania 17837.
Email: wflack@bucknell.edu

Notes

1 Investigators of emotion owe a considerable debt of gratitude to Paul EKMAN for his recent scholarly restoration and elaboration of DARWIN's book on emotion.

2 One example from HEGEL (1971): "In all the corporealizations of the mental or spiritual [feelings] just considered, only that externalization of the emotions occurs which is necessary for them to be felt or which can serve to indicate the inner sensations" (p. 85; brackets added by the present author).

3 Published at about the same time as, and bearing important similarities to, as well as differences with, the thinking of Carl Lange (JAMES/LANGE 1922).

4 Sole exception is a study by TOURANGEAU/ELLSWORTH (1979).

5 This thinking provides a theoretical perspective on the meaning of severe psychopathology, and does not limit prematurely the range of potential theories of explanation.

6 Although this designation usually refers to extreme abnormalities of thinking, it is used here to denote severe abnormalities of acting, thinking, and feeling.

7 Examination of the audio portion of the videotapes revealed almost no evidence of abnormalities in the contents of schizophrenic patients' speech.

8 The posture for happiness was an exception to this rule, although the results of other studies are consistent in demonstrating an effect of happy postures in normal subjects.

References

Aberle, D. F./Cohen, A. K./Davis, A. K./Levy, M. J./Sutton, F. X. (1950) The functional prerequisites of a society. Ethics 60: 100–111.

Adelman, P. K./Zajonc, R. B. (1988) Facial efference and the experience of emotion. Annual Review of Psychology 40: 249–280.

American Psychiatric Association (1994) Diagnostic and statistical manual of mental disorders, 4th edition. American Psychiatric Association: Washington DC.

Bellack, A. S./Blanchard, J. J./Mueser, K. T./Sayers, S. L. (1993) Affect perception in schizophrenia. Schizophrenia Bulletin 9: 93–104.

Bem, D. J. (1967) Self-perception: An alternative interpretation of cognitive dissonance phenomena. Psychological Review 74: 183–200.

Bleuler, E. (1950) Dementia praecox or the group of schizophrenias. International Universities Press: New York. German original: (1911) Dementia Praecox oder Gruppe der Schizophrenien. In: Aschaffenburg, G. (ed) Handbuch der Psychiatrie. Franz Deuticke: Leipzig. Translated by J. Zinkin.

Capella, J. N. (1993) The facial feedback hypothesis in human interaction: Review and speculation. Journal of Language and Social Psychology 12: 13–29.

Ciompi, L. (1988) The psyche and schizophrenia: The bond between affect and logic. Harvard University Press: Cambridge MA.

Ciompi, L. (1997) The concept of affect logic: An integrative psycho-socio-biological approach to understanding and treatment of schizophrenia. Psychiatry 60: 158–170.

Cornelius, R. R. (1996) The science of emotion: Research and tradition in the psychology of emotions. Prentice-Hall: Upper Saddle River NJ.

Darwin, C. (1998) The expression of the emotions in man and animals, 3rd edition. With an introduction, afterword and commentaries by P. Ekman. Oxford University Press: New York. Originally published 1872.

Dewey, J. (1894) The theory of emotion I: Emotional attitudes. Psychological Review 1: 553–569.

Dewey, J. (1895) The theory of emotion II: The significance of emotions. Psychological Review 2: 13–32.

Duclos, S. E./Laird, J. D./Schneider, E./Sexter, M./Stern, L./Van Lighten, O. (1989) Emotion-specific effects of facial expressions and postures on emotional experience. Journal of Personality and Social Psychology 57: 100–108.

Duncan, S. D., Jr./Fiske, D. W. (1985) Interaction structure and strategy. Cambridge University Press: Cambridge, England.

Flack, W. F., Jr./Canty, T./Edwards, R./Hellman, B. (in preparation) The effect of facial expression on emotional feeling: Does experimental demand account for the relationship? Manuscript in preparation.

Flack, W. F., Jr./Cavallaro, L. A./Laird, J. D./Miller, D. R. (1997) Accurate encoding and decoding of emotional facial expressions in schizophrenia. Psychiatry 60: 197–210.

Flack, W. F., Jr./Cavallaro, L. A./Miller, D. R. (1999) Sources of interpersonal disruption in schizophrenics' conversational turn-taking behavior. Unpublished manuscript.

Flack, W. F., Jr./Laird, J. D./Cavallaro, L. A. (1999a) Emotional expression and feeling in schizophrenia: Effects of specific expressive behaviors on emotional experiences. Journal of Clinical Psychology 55: 1–20.

Flack, W. F., Jr./Laird, J. D./Cavallaro, L. A. (1999b) Separate and combined effects of facial expressions and bodily postures on emotional feelings. European Journal of Social Psychology 29: 203–217.

Flack, W. F., Jr./Laird, J. D./Cavallaro, L. A./Miller, D. R. (1998) Emotional expression and experience: A psychosocial perspective on schizophrenia. In: Flack, W. F., Jr./Laird, J. D. (eds) Emotions in psychopathology: Theory and research. Oxford University Press: New York, pp. 315–322.

Flack, W. F., Jr./Miller, D. R. (1992) Normality, abnormality, and the meaning of schizophrenia. In: Cang, S./Cason, K. (eds) Festschrift for Elliott Jaques (Limited edition). Cason Hall & Co: Arlington VA, pp. 175–189.

Grant, E. C. (1972) Non-verbal communication in the mentally ill. In: Hinde, R. A. (ed) Non-verbal communication. Cambridge: Cambridge University Press, pp. 349–358.

Hatfield, E./Caccioppo, J. T./Rapson, R. L. (1994) Emotional contagion. Cambridge University Press: Paris.

Hegel, G. F.W. (1971) Hegel's philosophy of mind. Clarendon Press: Oxford. German original: (1830) Philosophie des Geistes, Teil 3 of Encyklopädie der philosophischen Wissenschaften. Translated by W. Wallace and A. Miller.

Izard, C. E. (1990) Facial expressions and the regulation of emotion. Journal of Personality and Social Psychology 58: 487–498.

James, W. (1884) What is an emotion? Mind 9: 18–205.

James, W. (1890) The principles of psychology. Dover: New York.

James, W./Lange, C. G. (1922) The emotions. Williams & Wilkins: Baltimore. Originally published 1885.

Joiner, T./Coyne, J. C. (eds) (1999) The interactional nature of depression. American Psychological Association: Washington DC.

Kraepelin, E. (1917) Clinical psychiatry: A textbook for students and physicians. Macmillan: New York. German original: (1899) Psychiatrie. Ein Lehrbuch für Studirende und Ärzte. 6., vollständig umgearbeitete Auflage. 2 Bände. Barth: Leipzig. Translated by A. R. Diefendorf.

Laird, J. D. (1974) Self-attribution of emotion: The effects of expressive behavior on the quality of emotional experience. Journal of Personality and Social Psychology 33: 475–486.

Laird, J. D. (1984) The real role of facial response in the experience of emotion: A reply to Tourangeau and Ellsworth, and others. Journal of Personality and Social Psychology 47: 909–917.

Laird, J. D./Apostoleris, N. H. (1996) Emotional self-control and self-perception: Feelings are the solution, not the problem. In: Harre, R./Parrott, G. W. (eds) The emotions: Social, cultural and physical dimensions. Sage: London, pp. 285–301.

Laird, J. D./Bresler, C. (1992) The process of emotional feeling: A self-perception theory. In: Clark, M. (ed) Emotion: Review of personality and social psychology, vol. 13. Sage: Newbury Park CA, pp. 223–234.

Matsumoto, D. (1987) The role of facial response in the experience of emotion: More methodological problems and a meta-analysis. Journal of Personality and Social Psychology 52: 769–774.

McIntosh, D. N. (1996) Facial feedback hypotheses: Evidence, implications, and directions. Motivation and Emotion 20: 121–147.

Mead, G. H. (1934) Mind, self, and society: From the standpoint of a social behaviorist. University of Chicago Press: Chicago.

Mead, G. H. (1982) The individual and the social self: Unpublished work of George Herbert Mead. University of Chicago Press: Chicago.

Miller, D. R. (1963) The study of social relationships: Situation, identity, and social interaction. In: Koch, S. (ed) Psychology: A study of a science. Vol. 5: The process areas, the person, and some applied fields: Their place in psychology and in science. McGraw-Hill: New York, pp. 639–737.

Miller, D. R. (1983) Self, symptom, and social control. In: Sarbin, T. R./Scheibe, K. E. (eds) Studies in social identity. Praeger: New York, pp. 319–338.

Miller, D. R./Jaques, E. (1987) Identifying madness: An interaction frame of reference. In: O'Gorman, H. J. (ed) Surveying social life: Essays in honor of Herbert H. Hyman. Wesleyan University Press: Middletown CT, pp. 265–286.

Myers, G. E. (1986) William James: His life and thought. Yale University Press: New Haven.

Rosenfeld, H. (1987) Impasse and interpretation. Tavistock: London.

Siegman, A. W./Boyle, S. (1993) Voices of fear and anxiety and sadness and depression: The effects of speech rate and loudness on fear and anxiety and sadness and depression. Journal of Abnormal Psychology 102: 430–437.

Stepper, S./Strack, F. (1993) Proprioceptive determinants of emotional feelings. Journal of Personality and Social Psychology 64: 211–220.

Strack, F./Martin, L. L./Stepper, S. (1988) Inhibiting and facilitating conditions of the human smile: A non-obtrusive test of the facial feedback hypothesis. Journal of Personality and Social Psychology 54: 768–777.

Tomkins, S. S. (1963) Affect, imagery, consciousness. Vol. 2: The negative affects. Springer: New York.

Tourangeau, R./Ellsworth, P. C. (1979) The role of facial response in the experience of emotion. Journal of Personality and Social Psychology 37: 1519–1531.

Wimmer, M./Ciompi, L. (1996) Evolutionary aspects of affective-cognitive interactions in the light of Ciompi's concept of "affect-logic". Evolution and Cognition 2: 37–58.

Winton, W. M. (1986) The role of facial response in self-reports of emotion: A critique of Laird. Journal of Personality and Social Psychology 50: 808–812.

W. Todd DeKay and Todd K. Shackelford

Toward an Evolutionary Approach to Social Cognition

Abstract

We discuss the benefits of combining evolutionary and traditional cognitive approaches in the study of social cognition. A central goal of applying a traditional cognitive approach is to develop a computational theory (through task analysis) of the domain under study. A task analysis includes a specification of the computational problems that the system must solve for it to operate successfully. Evolutionary psychological theory provides a useful guide in developing a computational theory because it allows a top-down analysis of the constraints that a complex, information processing system must meet to exist in its current form, and because it provides principled criteria for evaluating the success of the system. Research on grandparental investment is used to illustrate the advantages of using evolutionary principles within a task analysis. We conclude that evolutionary approaches to social phenomena can be made more rigorous by adding a computational layer of analysis, and functionally agnostic approaches within psychology can better avoid blind alleys of investigation by adding a rigorous evolutionary approach.

Key words

Theory of evolution, evolutionary psychology, psychology, social cognition, grandparenting, altruism.

Humans are intensely social, and thus it is reasonable to assume that a large proportion of the human cognitive architecture is dedicated to social tasks (see FISKE 1991; HIRSCHFELD/GELMAN 1994; LESLIE 1987; TOOBY/COSMIDES 1992). The investigation of social cognitive mechanisms is not easy, however, because they are sensitive to contextual variation, and because multiple mechanisms are simultaneously engaged by the same social task (TOOBY/COSMIDES 1992). As a result, many of the most powerful conceptual tools developed by cognitive scientists have not been immediately useful to social psychologists (but see HASTIE 1988, and SMITH/ZARATE 1992, for applications to person perception and simple social judgement). Recent advances in evolutionary psychology can help social psychologists develop a rigorous top-down approach to social cognition by adding a layer of analysis to the 'top' of the traditional approach employed in cognitive science.

In this article, we first re-introduce to psychologists the multilevel approach often used in cognitive science, and we argue that this approach can strengthen social psychological research and theory. Second, we argue that the effective use of a multilevel approach in social domains will require the addition of an evolutionary level of analysis. Finally, we argue that current work in evolutionary psychology, especially work on social phenomena, can be made more rigorous if evolutionary theory is used in conjunction with the multilevel approach of cognitive science.

In his work on visual perception, David MARR (1977, 1982; MARR/POGGIO 1977) proposed that any information processing system can be understood at three levels of analysis (see Table 1). The first level, the computational theory, specifies the problem to be solved, the information available for solving the problem, and the sub-tasks that must be completed to solve the problem, given the informational constraints imposed on the system. Developing a computational theory entails decomposing a larger problem (e.g., visual perception) into the sub-problems (e.g., edge detection, computation of surface orientation) that must be solved, in a real-world environment, to accomplish the larger task.

The second level, the algorithm, specifies the rules used by the system for solving the problem. The algorithm is specified in abstract terms, independent of the physical system on which it is implemented. The algorithm specifies the nature of

I. Computational theory (Task Analysis)
What is the goal of the computation, why is the computation appropriate, and what is the logic of the strategy by which the computation can be carried out?
II. Algorithm (Cognition)
How can the computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation of input to output?
III. Implementation (Hardware)
How can the algorithm and representation be realized physically?

Table 1: MARR's Three Levels of Analysis. Adapted from MARR (1982, p25).

the input into the system, in addition to the operations performed by the system that solve the subproblems specified by the computational theory (e.g., for edge detection, computing the second derivative of light intensity, and using zero-crossings to indicate edges; MARR 1982).

The third level, the implementation (or hardware), concerns the physical instantiation of the algorithms specified at the second level. Detailing algorithmic implementation amounts to a description of the mechanical processes by which the system carries out the operations specified abstractly at the second level (e.g., for edge detection, 2 geniculate X-cells, one on-center and one off-center, connected by an 'and' gate; MARR 1982).

For cognitive scientists, and most psychologists, a primary goal is to understand the abstract rules according to which the mind works. This corresponds to MARR's second level of analysis and captures what psychologists mean by cognition—the ways in which information is used by the mind. The algorithmic level of analysis captures a unique class of generalizations not reflected in purely physical or purely behavioral descriptions. It is the algorithmic level of analysis that is typically of primary interest to psychologists (CHOMSKY 1957; FODOR 1968; PYLYSHYN 1984).

According to MARR (1977, 1982), the most profitable strategy for identifying the algorithms used by the mind is to adopt a top-down strategy by first developing a computational theory for the domain in question. Careful consideration of the specific sub-tasks required in a domain, together with a consideration of the information available for accomplishing these tasks, limits the potentially viable algorithms. Only algorithms capable of successfully completing the series of required sub-tasks, using information available to the system in real-world situations, are worth consideration. Employing a computational theory prevents the researcher from considering the huge array of algorithms that are impossible given the real-world constraints on the system.

The set of viable algorithms is relatively unconstrained, however, by the physical hardware on which the algorithms are implemented. Given a complete description of the physical elements of a system, in addition to a complete description of how those physical elements are connected, the researcher remains uninformed as to the abstract rules that are instantiated in that physical system. In MARR's words:

"Although algorithms and mechanisms are empirically more accessible, it is the top level, the level of computational theory, which is critically important from an information-processing point of view. The reason for this is that the nature of the computations that underlie perception depends more upon the computational problems that have to be solved than upon the particular hardware in which their solutions are implemented. To phrase the matter another way, an algorithm is likely to be understood more readily by understanding the nature of the problem being solved than by examining the mechanisms (and the hardware) in which it is embodied" (MARR 1982, p27).

A simple thought experiment illustrates MARR's point. Imagine a giant machine of some unknown but definite function. Further imagine that this machine is constructed entirely of Tinker Toys. Each piece is made of wood, and is one of a small set of types (e.g., cylindrical with holes, or long and stick-like). What algorithms are instantiated in this machine? That is, what are the rules that guide its operation? The answers to these questions are relatively unconstrained by a physical examination of the machine or by a description of its specific movements or 'behaviors'. The set of possible rules guiding the operation of this machine is substantially reduced, however, if we first formulate a hypothesis about the problem the machine solves: It plays, and never loses at, tic–tac–toe (this machine does exist; DEWDNEY 1989). This piece of information now constrains the possible algorithms instantiated in the machine to such a degree that only a small number are viable. For example, algorithms such as, "Cooperate on the first move. Then defect if your partner defects, and cooperate if your partner cooperates," and "Flip pancakes after cooking for two minutes," are no longer worth consideration. These algorithms are technically possible given the physical medium involved but are unrelated to the machine's function and so represent 'blind alleys' of investigation. However, this machine may include a system for representing the possible configurations of X's and O's in a 3 by 3 environment, and may perhaps embody a set of strategic algorithms such as, "If two of your opponent's symbols are in adjacent squares, place your next symbol in the next successive square," and "If moving first, always place your symbol in the center square". These algorithms embody solutions to specific sub-tasks necessary to accomplish the larger goal—never losing at tic–tac–toe.[1]

MARR's (1982) computational approach revolutionized theories of visual processing, in particular, and cognitive science, in general. It provided a framework for understanding perceptual processes through a careful consideration of the demands of the problem. MARR's ideas, however, have not penetrated social psychological circles to the same degree. One possible reason is that MARR (1982) and others concerned primarily with perceptual processes have had it comparatively easy in applying a computational approach. The existence of a perceptual process is evident and need not be inferred. Furthermore, once a perceptual process is identified, a task analysis can proceed 'from the ground up'. MARR (1982) relied heavily on observation (what features are, and are not, detectable by people) and intuition (for example, to see a table, we must perceive its edges and its surfaces) to develop his computational theory of vision. Operations were identified as 'appropriate' based on intuitions about the relevant processes and relations between them. For example, MARR explains that the reason a cash register uses addition, rather than multiplication, is that "the rules that we *intuitively feel to be appropriate* for combining the individual prices in fact define the mathematical operation of addition" (1982, p22, emphasis added). And so the addition of prices would be part of any computational theory of cash registers. Thus, observation and intuition give the perception researcher an initial foothold from which to identify sub-processes, and the analysis can then proceed 'from the top down', with the computational theory providing constraints on possible algorithms (see Figure 1).

Developing computational theories in social domains is not as straightforward. Humans pursue friendships, form social hierarchies, and become romantically involved with others, for example. These social phenomena are accompanied by a baffling array of psychological processes, but even the existence of these processes (unlike most perceptual processes) is not always immediately evident. Furthermore, it is not always clear when a social agent is operating successfully, or even what an appropriate metric for 'success' is (Is success equivalent to subjective happiness? To the maximization of pleasure? To the minimization of pain?). Nor is it intuitively obvious what sub-problems must be solved by a social agent in order to be successful (What must we do to ascend successfully a social hierarchy?).

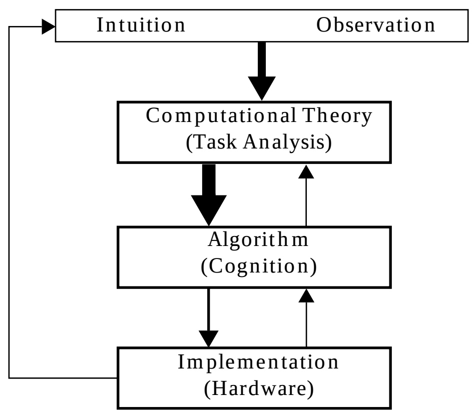

Figure 1: MARR's (1982) computational approach. Arrows represent the direction of inferences from one level to another. The thickness of the arrow indicates the relative importance of the associated inference process to MARR's approach.

Intuition and observation provide little help. Our intuitions often are blind to the complexity of tasks that we perform well with little or no conscious effort (COSMIDES/TOOBY 1994). The apparent (phenomenological) ease with which we use language to convey and consume ideas, for example, masks the cognitive complexity that underlies even the simplest conversation about the weather (PINKER 1994; SPERBER/WILSON 1995).

To be effective in social domains, this intuition-dependent approach must be replaced with an approach that specifies the tasks and sub-tasks present in an information processing domain and that specifies (1) what it means for a social psychological process to be successful and (2) what specific problems an organism must solve to be successful in a particular social domain. Evolutionary psychological theory can provide social psychologists with a powerful set of tools for these purposes.

The Use of Evolutionary Theory within a Cognitive Approach

Recent advances in evolutionary psychology enable the application of a computational approach to social psychological domains. Computational modeling provides a framework for decomposing complex psychological phenomena into the specific information processing tasks required for the phenomena to occur in real-world circumstances. Evolutionary psychology can strengthen the computational approach and allow its application to social phenomena by supplanting the role of intuition in the process of generating computational theories, and by providing a theoretically grounded metric for deciding when psychological processes are operating successfully. Combining evolutionary psychological theory with a key component of the computational approach, a task analysis, can greatly reduce the chance of making conceptual errors when building theories of social cognitive processes.

Task analysis

A computational approach provides a heuristic for building viable theories of cognitive processes. The computational approach to the mind, in addition, is an epistemological statement: Mental processes are computational (i.e., they involve representations, operations, manipulation of symbols); the mind 'computes' (HOBBES 1974; PYLYSHYN 1984). It is useful, however, to distinguish the epistemology of computational theory from the procedural heuristic of task analysis in thinking about social cognitive processes.

Adopting a task analysis involves breaking a psychological process into the sub-tasks that must be solved for that process to occur. Adopting a task analysis does not require the belief that the mind computes anything in the formal sense (although this approach often will lead to that conclusion). Nor does it require formal computational (mathematical) models of social cognition, or a working computer simulation. In adopting a task analysis, one simply is disposed toward decomposing a proposed psychological process into its constituent parts by specifying the tasks that must be solved to reach the end state.

Using evolutionary theory

This approach involves one simple assumption: that any proposed set of cognitive processes must be evolvable. Because the cognitive architecture that underlies social cognition is a product of natural selection, it must be structured in such a way that is consistent with this causal history. Detailed models of the process of natural selection, including a specification of relevant selection pressures and ancestral environments, can provide a principled set of tools to begin specifying the cognitive tasks that are solved in a particular domain.

The first step in applying an evolutionarily informed computational approach is to identify the selective constraints on potential cognitive adaptations in a particular domain. Game theory models can be used profitably at this step. In the domain of cooperation, for example, the theories of kin selection (HAMILTON 1964) and reciprocal altruism (AXELROD 1984; COSMIDES/TOOBY 1992; TRIVERS 1971) provide models of some of the constraints any set of cognitive processes must have satisfied recurrently if they presently guide cooperative or altruistic behaviors (e.g., 'HAMILTON's rule'; HAMILTON 1963). These models suggest a series of sub-tasks that must be performed to meet these constraints. Using these models, we can generate hypotheses about the algorithms that guided cooperative behavior under ancestral conditions. For example, in cooperation and altruism, the uncertain nature of kinship (due to the possibility of cuckoldry) imposes a class of constraints that may have designed mechanisms for evaluating the uncertainty of relatedness (DALY/WILSON 1982; DEKAY 1995, 2000; EULER/WEITZEL 1996; see below).

Studies of hunter–gatherer societies (HILL/HURTADO 1996), paleontological research (TRINKAUS/ZIMMERMAN 1982), and reverse engineering (reconstructing the past from examining currently existing adaptations; DENNETT 1995; PINKER 1997) aid in identifying the environmental features available for exploitation, and in identifying the conditions under which the proposed algorithms must have been operative. The end result is a set of proposed cognitive processes, which can (and could) solve the problem in question, given the constraints operative in the relevant ancestral environments (see Figure 2).

Grandparental Investment: An Example of an Evolutionary Approach to Social Cognition

DEKAY (1995, 2000) has applied an evolutionary–cognitive approach to cooperation by developing an analysis of the relevant evolvability constraints, performing a task analysis, and proposing previously undiscovered psychological processes

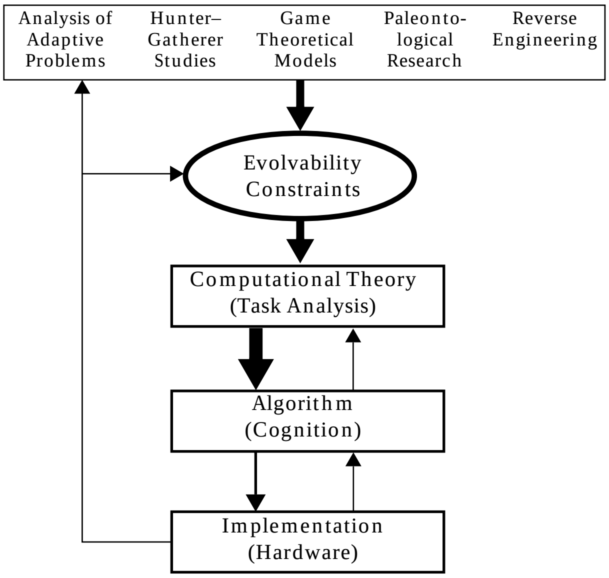

Figure 2: An evolutionary approach to social cognition. Arrows represent the direction of inferences from one level to another. The thickness of the arrow indicates the relative importance of the associated inference process to the evolutionary approach.

involved in kinship interactions. This work is presented as an example of the approach we advocate.

Evolvability constraints on cooperation and altruism

The first step in DEKAY's analysis was to recognize a universal evolvability constraint—that any set of processes, on average and over time, must have led to greater reproductive benefits than costs. This general constraint applies to the evolution of any mechanism, including those involved in cooperation and altruism, and subsumes both the processes of natural selection and sexual selection.

Step 2 in DEKAY's analysis was to express the initial constraint in Inclusive Fitness terms (HAMILTON 1963, 1964). That is, the reproductive costs and benefits specified in step 1 include not only the costs and benefits to the individuals directly involved in the cooperative interaction, but also the sum of the costs and benefits to all individuals affected by the cooperative interaction, times the degree of relatedness (*r*) between each individual affected and each directly participating individual. Degree of relatedness is defined as the probability that two individuals share some genetic unit (e.g., gene) due to common descent, or by direct generational genetic transmission. An individual's degree of relatedness with siblings, parents and children is 0.50; with grandparents it is 0.25; and with cousins it is 0.125.

Step 3 focused on the kinship aspect of these constraints. True relatedness is a function of kinship category membership (e.g., sibling, parent, half-sibling, grandparent), but also is a function of relational certainty (*R*), or the probability that two individuals are related, independent of their putative kinship status. Kinship is usually uncertain due to the possibility of cuckoldry severing the line of descent from one individual to another. This is the case for any two individuals related, at some point in their ancestry, through a common male. For example, two grandmothers can be differently uncertain about their relatedness to their grandchild. One grandmother, with a grandchild by her son, is certain that her son is her genetic kin, but is less than 100% certain that her son's child is actually his. The other grandmother, with a grandchild by her daughter, is 100% certain that her daughter is actually hers, and that her daughter's child is actually her daughter's. This is an important aspect of kinship and must be part of the evolvability constraints on kin-directed cooperation and altruism.

Step 4 in the analysis elaborated the constraint of *R* to recognize that *R* is a function of the product of the probability of cuckoldry in each generation separating two putatively related individuals. The result is a detailed set of constraints that any set of mechanisms involved in cooperation and altruism must have met, on average and over time, to have been favored by selection and to exist in their current form (see Table 2).

Task analysis

The evolvability constraints identified by DEKAY allow a detailed task analysis of mechanisms involved in cooperation and altruism (see Table 3). For example, these mechanisms must include processes designed to identify and recognize putative kin, and processes designed to identify and deal appropriately with cheaters—individuals who accept the benefits of an interaction without incurring the costs (COSMIDES 1989; COSMIDES/TOOBY 1992). Also included in the sub-tasks involved in cooperation and altruism are some that previously have been overlooked. In addition to mechanisms designed to identify and recognize putative kin, cooperation and altruism directed towards kin must also involve processes designed to (1) evaluate relational certainty (*R*), (2) weigh putative relatedness by *R*, and (3) weigh cost/benefit evaluations appropriately.

$$B_1 > C_1$$ *Step 1*: For mechanisms involved in reciprocal and unidirectional 'altruism' the average reproductive benefits (B_1) must outweigh the average reproductive costs (C_1).
$$\sum_{i=1}^{x} r_{(1,i)}B_i > \sum_{i=1}^{x} r_{(1,i)}C_i$$ *Step 2*: The reproductive benefits and costs are equal to the sums of the reproductive benefits and costs for all individuals (i) affected by the interaction, times the degree of relatedness (r) between the individual directly participating in the interaction and each individual affected (inclusive fitness).
$$r_{\text{true}(1,i)} = r_{\text{put}(1,i)}R_{(1,i)}$$ *Step 3*: True relatedness (r_{true})is equal to putative relatedness (r_{put}) times 'relational certainty' (R).
$$R_{(1,i)} = \prod_{j=0}^{y} [1 - p(\text{cuck})_j]$$ *Step 4*: Relational certainty is the product of one minus the probabilities of cuckoldry [$p(\text{cuck})$] in each generation (j) separating two individuals.
$$\sum_{i=1}^{x}\left[r_{\text{put}(1,i)}\left(\prod_{j=0}^{y} 1 - p(\text{cuck})_j\right)B_i\right] > \sum_{i=1}^{x}\left[r_{\text{put}(1,i)}\left(\prod_{j=0}^{y} 1 - p(\text{cuck})_j\right)C_i\right]$$ *Result*: Detailed constraints on the evolution of cooperation and altruism. Any set of mechanisms must have, on average and over time, satisfied these constraints.

Table 2: An Analysis of the Evolvability Constraints on Cooperation and Altruism. Adapted from DEKAY (2000).

Cognitive processes involved in cooperation and altruism must include mechanisms that:	
1	Recognize different individuals
2	Identify kin and distinguish between individuals based on degree of relatedness (r)
3	Evaluate the 'relational certainty' of kin (DEKAY 2000)
4	Weigh degree of relatedness by relational certainty (DEKAY 1999)
5	Estimate the costs and benefits of an interaction to oneself and to others
6	Weigh the costs and benefits by degree of relatedness and relational certainty
7	Store information about past interactions
8	Detect and punish, or exclude, cheaters (COSMIDES/TOOBY 1992)

Table 3: Results of a Task-Analysis of Social Exchange. Adapted from COSMIDES/TOOBY (1992) and DEKAY (2000).

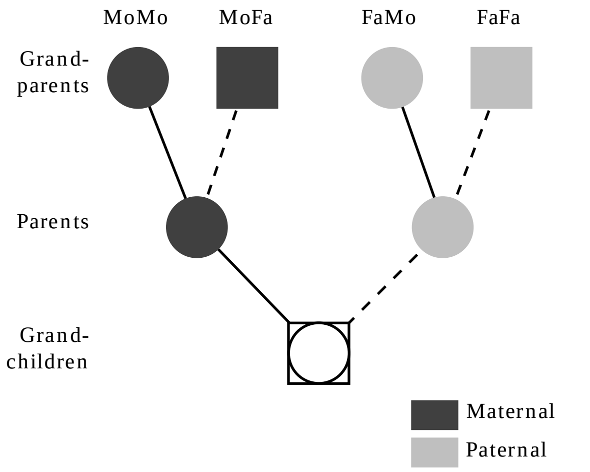

Figure 3: Relational certainty between grandparents, parents, and children/grandchildren. Dashed lines indicate uncertain connections due to the possibility of cuckoldry.

Processes involved in kin-directed cooperation and altruism

If people have psychological processes designed to assess relational certainty, and if they use these assessments in their helping decisions, we should see predictable patterns in grandparental investment. All else equal, mother's mother (MoMo) should be the most investing because she has no uncertainty about her relationship to her daughter's child. Father's father (FaFa) should be the least investing because he is 'doubly uncertain'—uncertain about his relatedness to his own son, and uncertain about his son's relatedness to his grandchildren. Mother's father (MoFa) and father's mother (FaMo) should be intermediate in investment because each have one place in the line of

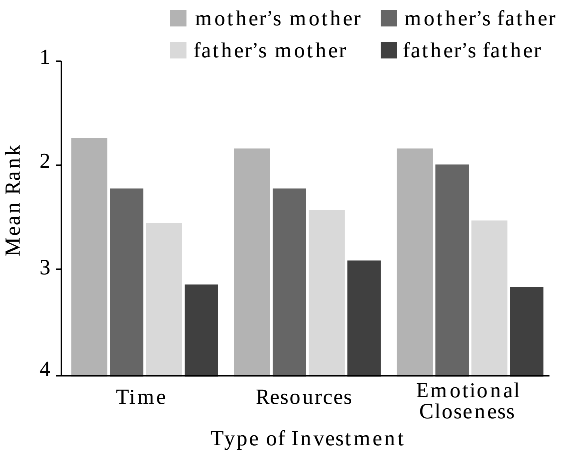

Figure 4: Mean rank of grandparents by grandchildren on the dimensions of 'emotional closeness', 'time spent', and 'resources (gifts, money, etc.)'. Lower ranks mean more investment (adapted from DEKAY 1999).

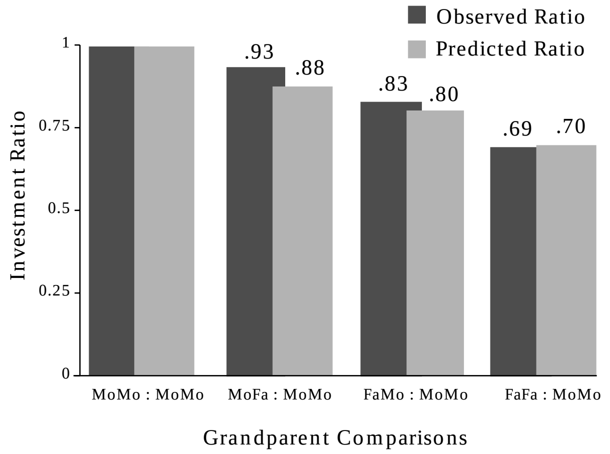

Figure 5: Predicted and observed ratios between Mother's Mother and other grandparents. Predictions are based on self-reported infidelity rates for female grandparent and parent cohorts (from LAUMANN et al. 1994).

descent between them and their grandchildren where cuckoldry could sever relatedness (see Figure 3).

If people use cues to infidelity in their assessments of relational uncertainty, and if they are reasonably accurate in these assessments, we can make additional predictions about grandparental investment patterns. If women's sexual infidelity is greater in the grandparental generation compared to the parental generation, FaMo should invest more than MoFa. This is because the uncertainty for MoFa lies in his own (grandparental) generation—he is uncertain about his relatedness to his own daughter. FaMo is less uncertain because her uncertainty lies in her son's generation. If women's sexual infidelity is greater in the parental generation compared to the grandparental generation, we should expect the opposite investment pattern (MoFa invests more than FaMo). If women's sexual infidelity is equal across generations, we should expect equal investment from MoFa and FaMo (see Figure 3).

In a recent study of the sexual attitudes and behaviors of a nationally representative sample (LAUMANN et al. 1994), 19.9% of women between 43 and 53 years old admitted to an extramarital affair. These women are the approximate age of the mothers of typical college students. In contrast, 12.4% of women 53 to 63 years old admitted to an extramarital affair. These women are the approximate age of the grandparents of a typical college sample. Therefore, among college-age subjects, we should expect the following general pattern of investment: MoMo > MoFa > Fa Mo > FaFa. Actual rankings by undergraduates of their grandparents' investment of time, resources, and emotional closeness confirm these predictions (DEKAY 2000; see also EULER/WEITZEL 1996, for supportive data from Germany; see Figure 4).

If people are accurate in their assessments of the probability of infidelity, we can also make point predictions about the relative investments of various grandparents. Using the figures from LAUMANN et al. (1994), we can calculate the expected ratio of investment, based on relational certainty, using mother's mother as the 'standard', because she is the only grandparent with no uncertainty. Relational certainty (*R*) is equal to one minus the product of the probability of cuckoldry in each generation separating two putative kin members. For MoFa, the single chance for cuckoldry exists in the grandparental generation. Since 12.4% of grandparent-aged women reported extramarital affairs, the approximate relational certainty (*R*) for MoFa is 0.88, and the expected ratio of investment between MoMo and MoFa is also 0.88. For FaMo, the single chance for cuckoldry exists in the parental generation. Since 19.9% of parent-aged women reported extramarital affairs, the approximate R for FaMo is 0.80, and the expected ratio between the investments of MoMo and FaMo is 0.80. For FaFa, there are two chances of cuckoldry, in both the parental and grandparental generations, and so *R* for FaFa is the product of those two probabilities, or approximately 0.70, and the expected ratio of investment between MoMo and FaFa is also 0.70. Actual reports by undergraduates of grandparental investment (rated on a 7-point scale) closely match these predicted ratios (DEKAY 2000; see Figure 5).

Summary

DEKAY's work on grandparental investment illustrates the advantages of using a top-down, task-analytic approach guided by evolutionary theory. The resulting task analysis revealed previously ignored processes that are necessary for cooperation and altruism between related individuals. This example represents only a first step in uncovering social cognitive processes using an evolutionary approach. Further decomposition of this domain into more specific sub-tasks will reveal additional layers of cognitive complexity. Cost/benefit evaluations, for example, are central to cooperation between strangers and between relatives (COSMIDES/TOOBY 1992; DEKAY 1995), but the specific processes underlying these evaluations are unknown. These examples also are not without controversy (see, for example, SPERBER/CARA/GIROTTO 1995). An evolutionary approach is one strategy for generating viable hypotheses about cognitive domains, not a guarantee that the resulting hypotheses will prevail over alternatives. A computational theory limits, but still underdetermines, the algorithm level of analysis and empirical research must decide between alternative accounts of the same phenomena.

The process of combining evolutionary and cognitive approaches promises to strengthen both evolutionary psychological work and social cognitive work. Each approach has limitations that can be avoided through the use of task analysis within an evolutionary framework. We discuss some of these limitations below.

Evolution without 'Cognition'

A common criticism of evolutionary approaches to social phenomena is that they are post hoc, or are simply stories that are untestable. Although this criticism is often misinformed, evolutionary psychologists can avoid certain pitfalls by using evolutionary theory within the top-down system described above. Often, these pitfalls occur because general evolutionary principles are applied directly to behavior, and bypass the psychological, or cognitive, level of analysis.

Process versus product

A common problem with evolutionary approaches to social phenomena is that they conflate evolutionary processes with evolutionary outcomes, or adaptations. This occurs when there is insufficient attention paid to the task demands of a domain. For example, 'HAMILTON's rule' states that natural selection favors designs for helping genetically related others if, on average and over time, the benefits to the helpee, times the degree of relatedness between the helper and helpee, are greater than the costs to the helper (HAMILTON 1963, 1964).

HAMILTON's rule represents a description of one aspect of the evolutionary process (or the conditions under which selection might operate), but is not itself a model of psychological mechanisms. There is no reason, based upon HAMILTON's rule alone, to expect that people have processes that compute the inequality 'on-line' during social interactions. Nor is there reason to expect that people's behavior in any specific instance will conform to the rule. To do so would be to jump erroneously from a description of the process of natural selection to features of organisms. Rather, HAMILTON's rule is best conceptualized as an evolvability constraint on the evolution of mechanisms involved in cooperation between related individuals. It sets conditions that processes within the domain must satisfy to be evolvable. There are unlimited ways to satisfy the condition, and computing HAMILTON's rule in each situation and using the outcome in behavioral decisions is only one of these possibilities. HAMILTON's rule, when combined with information about the past ecology and social structure of a species, can provide the foundation for a task analysis of the processes involved in kin-directed altruism. For example, DEKAY's work on grandparental investment patterns used this approach (see above).

Sociobiology and fitness maximization

Sociobiologists view social behaviors as adaptively patterned: Organisms currently behave in such a way as to maximize their inclusive fitness, or their relative genetic contribution to future generations (ALEXANDER 1979). According to this perspective, behavioral patterns vary across ecological and social contexts because particular behaviors have different fitness consequences in different contexts, and organisms are able to assess these fitness consequences and adjust their behavior in order to maximize fitness outcomes. Behavior is viewed as highly flexible, shifting in response to shifting fitness consequences.

Sociobiology has been criticized on several grounds (BUSS 1991, 1995; SYMONS 1989; TOOBY/COSMIDES 1990). A key problem of sociobiology is that it (often implicitly) proposes an impossible psychology. For an organism to adjust its behavior to maximize its inclusive fitness, it must have cognitive processes designed to evaluate its current fitness trajectory, decide if that trajectory could be improved, associate behaviors with increases or decreases in fitness, and predict the fitness outcomes associated with changes in behavior. Fitness consequences, however, are temporally distal to the organism. Fitness is not observable by the organism because it is determined in hindsight, based on the reproductive outcomes of an individual's children's children's... *ad infinitum*. It also is not clear how the organism might decide what to do, assuming it could determine that its current fitness trajectory was not maximally positive. The problem space is not sufficiently narrow to restrict potential solutions much beyond random trial–and–error, not a generally effective method for living (and dying!) things.

To solve the general problem specified by sociobiology—that of evaluating a current fitness trajectory and adjusting behavior to maximize that trajectory—a cognizing agent must solve sub-tasks that cannot be solved within real-world constraints. A model that proposes a set of cognitive processes that functions to evaluate on-line fitness and adjust behavior accordingly is not a viable model.

Standard Social Science

Standard social science (SSS) is a broad label for a heterogeneous endeavor. A full characterization of SSS is beyond the scope of this article, and is provided elsewhere (TOOBY/COSMIDES 1992; WILSON 1998). Common to variants of SSS is the assumption that the mind is initially content-free—a blank slate upon which experience writes, or a general-purpose computer, with 'culture', 'socialization', or 'learning' writing the programs. According to SSS models, social cognition and consequent behaviors are a function of general learning processes that have accumulated patterned cognitive contents (such as schemas and scripts, association strengths, or social roles), and have conditioned patterned behavioral responses, over the life of an individual. This view permeates current social science theory and research, although it often remains implicit (SYMONS 1989; TOOBY/COSMIDES 1992; WILSON 1998).

Organisms alter their behavior partly as a result of experience. The assumption that the mind is largely content-free, however, encourages the belief that a true causal process is being invoked by references to general, content-free concepts like 'culture' or 'learning'. An examination of the cognitive requirements for even the 'simplest' forms of learning (i.e., classical and operant conditioning) reveals a great deal of complexity (GALLISTEL 1990). Simple, general accounts of social phenomena fail partly because they underspecify the complexity of social cognition. The predominance of behaviorism, for example, crumbled under the weight of evidence demonstrating that its general principles, such as equipotentiality, could not account for the complexity of observations (GARCIA/ERVIN/KOELLING 1966; GARCIA/KOELLING 1966).[2]

Advantages of an Evolutionary Cognitive Approach

The rigorous application of an evolutionary approach to social cognition provides benefits that address some of the conceptual weaknesses present in other approaches. We elaborate these benefits below.

An evolutionary cognitive approach forces explicit examination of the nature of the mind and its component psychological mechanisms. Much social psychological and sociobiological theorizing proceeds without careful consideration of the assumptions made about the nature of the mind. Domain-general constructs such as 'fitness maximization', 'culture', and 'learning' often pass without a rigorous examination of the model of the mind they require (one that is equipotential and unstructured). An evolutionary cognitive approach, by contrast, forces an explicit description of the nature of the mind and demands that this description be empirically and theoretically scrutinized before research proceeds. Sociobiological and SSS notions of the mind are limited because they assume a relatively content-free, and hence impossible, architecture.

An evolutionary cognitive approach forces explicit examination of the adaptive tasks and subtasks required in a psychological domain. By considering the computational tasks associated with a psychological phenomenon, and by investigating only those solutions to the specified tasks that are possible in a real-world environment, one is prevented from disguising complexity with underspecified, global concepts such as 'socialization', 'culture', and 'learning'. In addition, new areas of inquiry emerge as the problem space is defined more clearly. GALLISTEL (1990) has applied such an approach to non-human learning and has documented numerous and complex information processing sub-tasks that are required for 'simple' learning processes (e.g., representations of the time, rate of event occurrence).

Authors' address

W. Todd DeKay, Department of Psychology, Franklin and Marshall College, Lancaster, PA 17604-3003, USA. Tel: ++1 (717) 291-4202; Fax: ++1 (717) 291-4387.
Email: t_dekay@acad.fandm.edu
Todd K. Shackelford, Department of Psychology, Florida Atlantic University, Davie, FL 33314, USA. Email: tshackel@fau.edu

An evolutionary cognitive approach demands explicit consideration of the cognitive algorithms that underlie psychological phenomenon. In employing this approach, one is forced to describe the rules by which a psychological process operates. This results in a sustained focus on the most appropriate level of analysis for understanding cognition, and for describing cognitive adaptations.

An evolutionary cognitive approach prevents proposing simplistic notions of causality. Decomposing a psychological phenomenon into its component processes prevents one from being lulled into proposing or believing that the cognitive tasks performed by the mind are simple, as can happen when our intuitions overlook complex tasks that we experience as effortless. An evolutionary cognitive approach forces one to examine the cognitive complexity underlying global concepts like 'culture' and 'learning'. By failing to appreciate the complexity implied by these underspecified concepts, sociobiology and SSS models have failed to produce models of cognition that respect the complexity of the mind. Consequently, sociobiology and SSS models have stopped short of investigating the complex cognitive details of social phenomena.

An evolutionary cognitive approach prevents proposing impossible solutions to psychological problems. Because it forces a decomposition of larger information processing problems into smaller ones, an evolutionary cognitive approach can expose proposed solutions that are not viable given the information available in real-world environments. The evolutionary component of this approach demands a focus on ancestral environments, both for specifying the constraints that must be met for a task to be solved, and for evaluating whether the proposed algorithm could have accomplished its presumed task. Sociobiological analyses, for example, fail partly because these analyses propose an impossible psychology, as fitness consequences are too distal to be used by cognitive processes.

Conclusion

Using an evolutionary cognitive approach is difficult because it requires interdisciplinary knowledge. It requires an understanding of psychology, anthropology, evolutionary biology, and cognitive science. It is increasingly clear, however, that psychological research can profit from such an interdisciplinary, integrated approach. This approach can help researchers avoid the pitfalls inherent in more traditional approaches and can help psychology to become integrated with the other life sciences.

Acknowledgement

The authors thank April BLESKE, David BUSS, Martie HASELTON, and Erika HOFF-GINSBERG for helpful comments on an earlier version of this article.

Notes

1 For a task as relatively simple as playing tic–tac–toe, with a limited number of possible states (board configurations), a simple 'look-up' table can suffice. Still, this look-up table simultaneously embodies a representation of possible board configurations as well as strategic algorithms, though they are simple and are equal in number to the possible board configurations themselves (e.g., if board configuration is X, then do Y). As the problem increases in complexity, however, the efficiency of a look-up table algorithm plummets, and algorithms are likely to profit from shortcuts or privileged hypotheses cutting through the exhaustive solution space.

2 Domain-general, content-free theories of social cognition also are problematic because, like sociobiology, they propose an impossible psychology. Any information-processing device must solve successfully the 'frame problem' (PYLYSHYN 1987; TOOBY/COSMIDES 1992). As a problem space increases (e.g., by the addition of dimensions to consider), the number of possible solutions to consider increases exponentially. Any information-processing device (including the human mind) must be able to limit the possibilities in order to operate successfully in real time. This means that the size of a problem space must be limited by imposing 'frames' (privileged hypotheses, domain-specific mechanisms). The domain-general procedures proposed by SSS fail to solve the frame problem and, therefore, constitute an impossible conception of human cognitive processes.

References

Alexander, R. D. (1979) Darwinism and human affairs. University of Washington Press: Seattle.

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

Buss, D. M. (1991) Evolutionary personality psychology. Annual Review of Psychology 42: 459–491.

Buss, D. M. (1995) Evolutionary psychology: A new paradigm for psychological science. Psychological Science 6: 1–30.

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

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

Cosmides, L./Tooby, J. (1992) Cognitive adaptations for social exchange. In: Barkow, J./Cosmides, L./Tooby, J. (eds) The adapted mind. Oxford University Press: New York, pp. 163–228.

Cosmides, L./Tooby, J. (1994) Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. Cognition 50: 41–77.

Daly, M./Wilson, M. (1982) Whom are newborn babies said to resemble? Ethology and Sociobiology 3: 69–78.

DeKay, W. T. (1995) Grandparental investment and the uncertainty of kinship. Paper presented at the annual meeting of the Human Behavior and Evolution Society, Santa Barbara, CA.

DeKay, W. T. (2000) The evolutionary psychology of cooperation: Grandparental investment as a test case. Manuscript under editorial review.

Dennett, D. C. (1995) Darwin's dangerous idea. Simon & Schuster: New York.

Dewdney, A. K. (1989) Computer recreations: A Tinkertoy computer that plays tic-tac-toe. Scientific American 261: 120–123.

Euler, H. A./Weitzel, B. (1996) Discriminative grandparental solicitude as reproductive strategy. Human Nature 7: 39–59.

Fiske, A. (1991) Structures of social life: The four elementary forms of social behavior. Free Press: New York.

Fodor, J. (1968) Psychological explanation. Random House: New York.

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

Garcia, J./Ervin, F. R./Koelling, R. A. (1966) Learning with prolonged delay of reinforcement. Psychonomic Science 5: 121–122.

Garcia, J./Koelling, R. A. (1966) The relation of cue to consequence in avoidance learning. Psychonomic Science 4: 123–124.

Hamilton, W. D. (1963) The evolution of altruistic behavior. American Naturalist 97: 354–356.

Hamilton, W. D. (1964) The genetical evolution of social behavior. Journal of Theoretical Biology 7: 1–52.

Hastie, R. (1988) A computer simulation model of person memory. Journal of Experimental Social Psychology 24: 423–447.

Hill, K./Hurtado, M. (1996) Ache life history. Aldine de Gruyter: Hawthorne NY.

Hirschfeld, L./Gelman, S. (eds) (1994) Mapping the mind: Domain specificity in cognition and culture. Cambridge University Press: New York.

Hobbes, T. (1974) Leviathan. New York: Penguin. Originally published in 1651.

Laumann, E./Gagnon, J./Michael, R./Michaels, S. (1994) The social organization of sexuality: Sexual practices in the Unites States. The University of Chicago Press: Chicago.

Leslie, A. (1987) Pretense and representation: The origins of "theory of mind". Psychological Review 94: 412–426.

Marr, D. (1977) Artificial intelligence—a personal view. Artificial Intelligence 9: 37–48.

Marr, D. (1982) Vision: A computational investigation into the human representation and processing of visual information. Freeman: San Francisco.

Marr, D./Poggio, T. (1977) From understanding computation to understanding neural circuitry. Neurosciences Research Progress Bulletin 15: 470–488.

Pinker, S. (1994) The language instinct. Morrow: New York NY.

Pinker, S. (1997) How the mind works. Norton: New York.

Pylyshyn, Z. (1984) Computation and cognition. MIT Press: Cambridge MA.

Pylyshyn, Z. (ed) (1987) The robot's dilemma. Ablex: Norwood NJ.

Smith, E. R./Zarate, M. J. (1992) Exemplar-based model of social judgement. Psychological Review 99(1): 3–21.

Sperber, D./Cara, F./Girotto, V. (1995) Relevance theory explains the selection task. Cognition 57: 31–95.

Sperber, D./Wilson, D. (1995) Relevance. 2nd edition. Blackwell: Cambridge MA.

Symons, D. (1989) If we're all Darwinians, what's the fuss about? In: Crawford, C./Smith, M./Krebs, M. (eds) Sociobiology and psychology. Erlbaum: Hillsdale NJ, pp. 121–146.

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

Tooby, J./Cosmides, L. (1992) Psychological foundations of culture. In: Barkow, J./Cosmides, L./Tooby, J. (eds) The adapted mind. Oxford University Press: New York, pp. 19–136.

Trinkaus, E./Zimmerman, M. R. (1982) Trauma among the Shanidar Neanderthals. American Journal of Physical Anthropology 57: 61–76.

Trivers, R. L. (1971) The evolution of reciprocal altruism. Quarterly Review of Biology 46: 35–57.

Wilson, E. O. (1998) Consilience: The unity of knowledge. Knopf: New York.

Zusammenfassungen der Artikel in deutscher Sprache

Jaak Panksepp/Jules B. Panksepp

Die sieben Sünden der 'Evolutionären Psychologie'

In diesem Artikel wird eine fundamentale Kritik an einigen der basalen Voraussetzungen der 'Evolutionären Psychologie' (EP) geübt. Die Kritik gründet auf Resultaten ethologischer, ontogenetischer und neurobiologischer Untersuchungen. Der EP wird dabei vorgeworfen, Erkenntnisse dieser Disziplinen nicht zu berücksichtigen und eine 'Überinterpretation' der menschlichen Natur vorzunehmen, die völlig von der Idee der 'inclusive fitness' beherrscht ist. Vor allem die neurologischen Hintergründe von affektiven und motivationalen Prozessen von Menschen und Tieren stehen dabei im Hintergrund und führen zu einer ungerechtfertigten Ausdehnung des Modulkonzeptes und der damit verbundenen erkenntnistheoretischen Implikationen. Kritisch betrachtet wird auch das Konzept von genetisch determinierten Modulen im Bereich des Neocortex, welche aus der Perspektive der EP ganz bestimmte Funktionen (special purpose functions) erfüllen, deren evolutionäre Hintergründe aus den Überlebensbedingungen des Pleistozäns ersichtlich gemacht werden können.

Dagegen legen die Ergebnisse neurobiologischer Forschungen am menschlichen Gehirn bzw. am Säugergehirn ganz allgemein nur die Existenz von subkortikalen Modulen nahe. Im Bereich 'höherer' Hirnregionen erscheint aus neurobiologischer Perspektive nur die Annahme der Existenz eines Sprachmodules als angebracht.

Der Neocortex wir im Gegensatz zur EP eher als 'Generalist' interpretiert, d.h. ein auf keine spezifischen Funktionen hin evoluiertes Organ, welches ontogenetischen Einflüssen und Anforderungen gegenüber in hohem Ausmaß offen ist.

Die sieben Sünden der EP werden in folgenden Bereichen geortet:

1. Überbetonung der pleistozänen Umgebungsbedingungen, welche sich formend auf menschliches Sozialverhalten ausgewirkt haben sollen
2. Extreme Zentrierung der EP auf die menschliche Spezies
3. Adaptationismus
4. Massive Modularität
5. Vermischung von Emotion und Kognition
6. Das Fehlen glaubwürdiger neuronaler Perspektiven
7. Anti-organische Tendenzen bzw. der Mythos vom computationalen Repräsentationalismus

Gordon G. Brittan

Die Unvermeidlichkeit des Anthropomorphismus

Anthropomorphismus wird als jene, von Ethologen strikte zu vermeidende Haltung bezeichnet, die darin besteht, tierisches Verhalten mit menschlichen Attributen und Eigenschaften in Beziehung zu setzten bzw. im Rahmen menschlicher Empfindungs- und Erfahrungsdimensionen zu interpretieren.

In diesem Artikel wird, nach einer kritischen Analyse des Anthropomorphismusbegriffes die Behauptung aufgestellt, daß es Ethologen in einem gewissen Ausmaß doch möglich ist sich in Tiere hineinzuversetzen. Weiters wird die These vertreten, daß intentionale Erklärungen tierischen Verhaltens möglich sind. Intentionale und funktionale Erklärungen ergänzen einander, wobei beide auf unterschiedlichen Analyseebenen Anwendung finden.

Jeanette Emt/Sverre Sjölander

Utopien. Eine unvermeidliche Konsequenz menschlichen Denkens?

Ausgangspunkt ist die Feststellung, daß utopische Vorstellungsformen im weitesten Sinne ein zentrales Charakteristikum des Menschseins überhaupt darstellen. Der Artikel versucht die Hintergründe und Ursachen dieser gedanklichen Dimensionen aufzuzeigen.

Eine der Erzeugung von Utopien zugrundeliegende kognitive Fähigkeit ist die Bildung von Vorstellungen welche von den aktuellen zeitlichen und räumlichen Rahmenbedingungen abgehoben werden können. Diese Fähigkeit entstand ursprünglich im Zusammenhang mit komplexen sozialen Interaktionen, Jagd und Sammlertätigkeiten und führte in weiterer Folge dazu, daß bessere interne bzw. externe Zustände im Vorstellungsraum repräsentiert sein konnten. Im subjektinternen Bereich werden unterschiedliche anzustrebende Zustände erreicht durch eine Vielfalt von Praktiken der Meditation, Selbstkontrolle, Drogeneinnahme usw. Dagegen steht der Versuch verbesserte externe Bedingungen durch politische Aktionen, Wissenschaft, Kunst und dgl. mehr zu erreichen.

Dabei bedingt die kognitive Ausstattung des Menschen auch die Gefahr der kritiklosen Übernahme und blinden Gefolgschaft von Ideologien und Ideologen. Dahingehend erscheint es notwendig sich dieser Anfälligkeiten und der—historisch aufzeigbaren—vielfach destruktiven Komponenten, welche mit der Realisierung utopischer Szenarien einhergehen ständig bewußt zu sein

Das Wissen um diese Anfälligkeiten, sowie die historisch vielfach belegten destruktiven Komponenten, welche mit der Realisierung von Utopien einhergehen, sollten neben den positiven Wirkungen utopischer Dimensionen immer bewußt bleiben.

C.U.M. Smith

Evolutionsbiologie und das 'hard problem'

Ausgangspunkt ist das sog. 'hard problem'—ein Problembereich welcher sich mit den Beziehungen zwischen den sog. Qualia, das sind spezifische mentale Zustände (z.B. die subjektive Qualität einer Farbwahrnehmung) und den zugrundeliegenden neurophysiologischen Prozessen befaßt.

Hinsichtlich des Bewußtseins erweist sich die Unterscheidung zwischen primärem und sekundärem Bewußtsein als hilfreich—eine Unterscheidung, die sich auch in zahlreichen philosophischen Überlegungen (AUGUSTINUS, HEIDEGGER, SARTRE etc.) findet. Darin wird das unmittelbare (praereflexive) Bewußtsein, von dem reflexiv—distanzierten Bewußtsein unterschieden, wobei Letzteres das praereflexive Bewußtsein zum Gegenstand haben kann.

Die Untersuchung der Hintergründe dieser Bewußtseinsstruktur führt zu einer Analyse von Komplexität, Systemeigenschaften und Emergenz.

Als Resultat dieser Untersuchungen wird eine materialistische Position formuliert welche in Anlehnung an die zeitgenössische Physik, die Evolutionsforschung und die Neurobiologie entstanden ist. Enthalten ist dabei auch eine Neuformulierung der 'klassischen' panpsychistischen Sichtweise, welche Bewußtsein und Subjektivität als inhärente Bestandteile dieser Welt interpretiert.

William F. Flack, Jr.

Emotionale Selbst-Wahrnehmung, soziale Anpassung und interpersonelle Störungen: eine funktionelle Perspektive hinsichtlich schwerer psychopathologischer Störungen

Ausgehend von DARWINs Werk 'Der Ausdruck der Gemütsbewegung bei Mensch und Tier' ('The Expression of the Emotions in Man and Animals', 1872) werden die Beziehungen zwischen emotionalen Ausdrucksphänomenen, emotionaler Erfahrung und sozialen Anpassungsvorgängen untersucht. Der Bezug zwischen emotionalem Ausdruck und emotionaler Erfahrung ist in der Selbst—Wahrnehmung gegeben, welche bei psychotischen Störungen unterschiedliche Fehlfunktionen aufweisen kann.

Der erste Abschnitt bezieht sich—unter Bezugnahme auf W. JAMES—vor allem auf die Beziehungen zwischen emotionalem Ausdruck und emotionaler Erfahrung.

Dieser Bereich wird im zweiten Abschnitt mit Vorgängen der Selbstwahrnehmung in Beziehung gesetzt, wobei diese das kognitive Bindeglied zwischen Emotionsausdruck und subjektiver Erfahrung (dem Gefühl) darstellt.

Im Bereich zwischenmenschlicher Beziehungen sind die emotionalen Komponenten menschlichen Verhaltens von besonderer Bedeutung. Reziproke Abstimmung von Verhaltensmustern im Verlauf von Gesprächen, sowie Empathie und Verstehen sind notwendige Voraussetzungen für das Entstehen zwischenmenschlicher Beziehungen. Sind derartige Fähigkeiten gestört, wird soziale Interaktion erschwert bzw. unmöglich. Die entsprechenden Verhaltensmuster werden vielfach als abnorm bzw. pathologisch klassifiziert.

Im letzten Abschnitt werden Daten referiert, die aus den Untersuchungen von als psychotisch diagnostizierten Patienten (Schizophrenie, Depression) entstanden sind. Dabei ist es vor allem die Trennung von emotionalem Ausdruck und emotionaler Erfahrung (Gefühl), welche für Schizophrene und Depressive bezeichnend ist.

W. Todd DeKay/Todd K. Shackelford

Soziale Kognition aus evolutionärer Perspektive

Das Phänomen sozialer Kognition als Gegenstand der Sozialpsychologie wird in diesem Artikel mit D. MARRS Arbeiten zur visuellen Wahrnehmung sowie mit soziobiologisch—evolutionären Überlegungen in Beziehung gesetzt. Ausgangspunkt ist dabei die bei Marr vorgenommene Unterscheidung zwischen drei Analyseebenen bei Wahrnehmungsvorgängen: Aufgaben- bzw. Problemanalyse, kognitive Prozesse, physische Strukturen. Nach MARR liefert bei derartigen Analyseformen die Aufgabenanalyse (oberste Ebene) wichtige Rahmenbedingungen, die zum Verständnis der zugehörigen kognitiven Prozesse unabdingbar sind. Sozialpsychologische Theorien weisen dahingehend oft beträchtliche Defizite auf, indem sie sich nur mit der kognitiven Ebene auseinandersetzen. Um derartige Aufgaben- bzw. Problemanalysen im sozialpsychologischen Bereich möglich zu machen werden evolutionäre Überlegungen herangezogen.

Die kognitiven Fertigkeiten, die im sozialen Bereich zum Ausdruck kommen werden als Produkte evolutionärer Prozesse aufgefaßt, d.h. sie gingen aus einem Selektionsprozeß hervor, der sich unter jeweils spezifischen Umgebungsbedingungen vollzog. Anwendung finden dabei vor allem soziobiologische Modelle, wie sie in der Analyse von Verwandtschaftsbeziehungen gebraucht werden. Diese evolutionäre Dimension soll dabei ergänzend zu den bisherigen Ansätzen im Bereich der Sozialpsychologie Verwendung finden.