

To appear in *Metascience*.

Archived at [HomePage for the Rutgers University Research Group on Evolution and Higher Cognition](#).

GRIFFITHS, ELIMINATION & PSYCHOPATHOLOGY

Dominic Murphy
Department of Philosophy
Rutgers University
New Brunswick, NJ 08901
DMu260967@aol.com

and

Stephen Stich
Department of Philosophy and Center for Cognitive Science
Rutgers University
New Brunswick, NJ 08901
stich@ruccs.rutgers.edu

Contents

1. Introduction
2. First Argument - The Heterogeneity of Emotions
3. The Eliminativist Conclusion of The First Argument
4. Against Particular Emotions
5. The Second Argument - Cladistics and Ecology
6. The Third (Dennettian) Argument
7. Extension to psychopathology
8. Specific pathologies
9. Conclusion

Paul Griffiths, *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press, 1997. Pp. xi + 286. US\$27.50.

1. Introduction

Paul Griffiths' excellent book, *What Emotions Really Are*, features valuable discussions of many important issues in evolutionary psychology and the philosophy of science. At the heart of the book, however, we find a trio of eliminativist arguments directed at our understanding of emotion. The first argument is aimed at the superordinate concept of emotion: Griffiths claims that emotion research shows that three very different sorts of states fall into the extension of our folk concept of emotion, and that as a consequence we cannot develop a general theory of emotion. The other two arguments are variations on a

single theme. They apply to concepts of individual emotions, not the concept of emotion as a whole. Griffiths maintains that in many instances folk psychology sees a distinct and unitary kind where in fact there is a heterogeneous category. In these cases, taxa which properly belong to two different sciences are contingently united. Whereas the first argument is claimed to follow from the current state of play in the literature, the latter two arguments reflect explanatory bets: Griffiths expects psychology to make an evolutionary turn which will undermine our existing taxonomies.

Like Griffiths, we foresee an evolutionary future for psychology, and we are interested in exploring the extent to which his arguments can be generalized. Although Griffiths does not discuss it, his book suggests the intriguing possibility that very similar reasoning to his might also apply to a wide range of other concepts which are invoked in both folk and scientific psychology. In this paper we'll begin by sketching our interpretation of Griffiths' three eliminativist arguments. Then we'll explore the application of his approach in a domain that is of enormous importance both theoretically and practically: psychopathology.

2. First Argument - The Heterogeneity of Emotions

Griffiths distinguishes three different kinds in the extension of our folk concept of emotion: affect programs, higher cognitive states, and social constructions. He suggests that these are *radically* different kinds for which no unified theory is likely, and that this trifurcation is a good reason for skepticism about the theoretical utility of our concept of emotion. Much of the research into affect programs adopts what Griffiths calls the "psychoevolutionary approach," which, he argues, was inaugurated by Darwin in *The Expression of the Emotions in Man and Animals* (1872). In chapter three Griffiths provides a fascinating survey of the modern psychoevolutionary movement, concentrating on the work of Paul Ekman (Ekman 1971, 1980, 1992; Ekman & Friesen 1971). What emerge from Griffiths' portrait of this research is a unified explanation of a number of paradigmatic "basic" emotions: surprise, fear, anger, disgust, sadness, joy and perhaps contempt. These are explained as complex, co-ordinated and automatic responses to events in the environment, involving physical changes at a number of levels, from the muscles of the face down to the autonomic nervous system.

The computational psychology of these mechanisms is modular, in a very Fodoresque sense (Fodor 1983). The affect programs are mandatory responses triggered by information derived from a very limited range of perceptual input - input which is not shared with more central processes, such as those underlying intentional action. Griffiths is careful to keep in mind the distinction between the input and output of affect programs. The output consists of a specific suite of physical changes. Data strongly suggest that these responses are pancultural; "people in all cultures respond in a similar way to things that frighten them." (p.55). Although the input which elicits these responses varies across cultures it is always the product of biased learning mechanisms, so some associations are learned much more readily than others.

Their mandatory, fast and passive character makes the affect programs candidates for the reference of some vernacular emotion concepts. However, folk psychology also recognizes other types of emotion which lack this strongly modular nature. Griffiths maintains that these other states cannot be explained by the same means as the affect programs. The fact that the psychoevolutionary results fail to generalize to other emotions is the cornerstone of Griffiths' overall argument for conceptual elimination and an important illustration of the mismatch between folk concepts and scientific categories.

In chapter five Griffiths discusses the "higher cognitive emotions" such as guilt, shame, sexual jealousy and loyalty. Some evolutionary psychologists try to explain the higher cognitive emotions along the lines of affect programs (Tooby & Cosmides 1990; Pinker 1997; Frank 1988). However, Griffiths departs from

these theorists in denying that the mechanisms implicated in higher cognitive emotions should be viewed either as instances or as further effects of affect program mechanisms. His argument turns on the fact that affect programs have a number of salient features which the cognitive emotions lack, and vice versa. Affect programs are reflex-like responses to a limited class of events in the environment while higher cognitive emotions are sensitive to a much wider range of information, including one's own thoughts. The higher cognitive emotions also fail to display stereotypical physiological effects, persist longer, and are well-integrated with cognitive activity such as planning. They are sensitive to a much wider range of information than the strongly encapsulated affect programs. Thus the higher cognitive emotions are not triggered as one would predict by assimilating them to affect programs. "If Othello's sexual jealousy had been an affect program or a downstream cognitive effect of such a program, he would have had to catch Desdemona in bed with Cassio, or at least have seen the handkerchief, before his jealousy was initiated (p.117)". Griffiths aligns himself with the work of Robert Frank who sees higher cognitive emotions as short-term irrational responses designed to keep one rational in the long term. For example, loyalty often leads to long-term co-operation rather than short-term defection in social interactions which have the structure of an iterated prisoner's dilemma (Frank 1988). A little cautious in chapter five, by the end of the book Griffiths clearly views the higher cognitive emotions as resting on psychological mechanisms that are quite different from those that subserve the affect programs. Cognitive emotions are not just further effects of affect programs.

Socially constructed states make up the last prong of Griffiths' trident. A paradigm example would be a state like *amok*, a violent frenzy found in southeast Asian cultures. Griffiths endorses what he calls the "social role" theory of socially constructed emotions. "A social role is a characteristic pattern of behavior found in a particular society" (p.139). So for example, one fills a social role when one is elected to Parliament. MPs enter a network of social practices in which they play a particular role. The roles they take up are relatively enduring, and overt in the sense that everyone agrees that all there is to being an MP is to be treated in a certain way. Griffiths thinks that emotions like being amok are *transitory covert* social roles. They are transitory because one fills these roles only in short-lived and stressful situations. They license behaviors which would otherwise be unacceptable; one takes advantage of the passivity usually attributed to emotions in order to evade the consequences of acting in that way. They are covert in the sense that society does not acknowledge their function or the social practices in which they are embedded. To acknowledge them in this way would lay bare their utility and in doing so destroy it.

Social role emotions are unlike other emotions not only in being culturally specific but also in their psychology. Griffiths thinks they are largely unconscious attempts to mimic some cluster of features characteristic of other emotions. This means that their etiology includes unconscious mechanisms of social cognition and not the perceptual mechanisms underlying the affect programs or the cognitive mechanisms which subserve the higher cognitive emotions. Therefore, they cannot be explained analogously to either. They are an independent kind of state which is parasitic on the universal evolved emotional repertoire.

3.The Eliminativist Conclusion of The First Argument

Griffiths applies to this three-fold picture of emotion the understanding of natural kinds recommended by Richard Boyd (1989, 1991). On Boyd's view a natural kind is a real category in nature tied together by a causal homeostatic mechanism which underlies projectibility and inductive reasoning. Griffiths' contention is that our concept of emotion is a cluster concept, with the core feature being passivity (p. 245). It does not pick out a natural kind which can be used to ground induction or projection across the range of emotions. The collection of features we think characterize emotions are explained by different causal mechanisms in different cases. "In the light of these findings, the category of emotion will have to be replaced by at least two more specific categories" (p. 242) -- the affect programs and the cognitive

emotions. Griffiths considers the case for identifying the vernacular concept with one of these categories to be no better than that for identifying it with the other. Since it cannot be identified with both, it should be identified with neither (i.e. eliminated from scientific psychology). He doesn't appear to think that the social roles are even a candidate for a successor to our emotion concept. We suspect this is because of their parochiality. Griffiths thinks that authentic psychological taxa must be species-wide (although not found in every member of the species), and social roles do not qualify. Of course, this does not mean that there could not be a psychology of the social cognition mechanisms which devise, implement and preserve the social roles.

Prior to Griffiths, most of the literature on eliminativism has focused on beliefs and desires (Churchland 1981; Stich 1983) and on qualia (Dennett 1988). In this literature, both eliminativists and their opponents adopted a common framework. They assumed that the battle would be lost or won when we determined how much of our commonsense psychological theory is true. Eliminativists claimed, and their opponents denied, that commonsense psychology offers a seriously mistaken account of the processes that give rise to behavior. Both sides agreed that if this was so then we could conclude that beliefs, desires and the other "posits" of commonsense psychology do not exist, and that terms or concepts putatively denoting them would find no place in any mature science that explained the workings of the mind or the brain. Recently, this framework itself has come under attack (Stich 1996). The critics maintain that there is no easy or obvious argument from the falsehood of a theory to the non-existence of the entities to which the theory's terms putatively refer. So, for example, from the fact that the ancients had a wildly mistaken theory about stars, it certainly does not follow that stars do not exist.

It is important to see that both the eliminativist conclusion Griffiths is urging and his arguments for that conclusion are largely orthogonal to these debates. As we read him, Griffiths' thesis is not that emotions do not exist, but rather that our *current concept* of emotion is ill suited for use in a scientific psychology, and that it will have to be replaced by several new concepts - concepts which refer to categories that are not recognized by commonsense psychology. To make the case he must impugn the scientific utility of our current conception of the emotions. But, as he recognizes, he must also offer an account of concept identity which sustains the claim that none of the concepts invoked in the new theories which he sees emerging - theories about affect programs, higher cognitive emotions and social role emotions - can be identified with our current concept of emotion. Griffiths allows that "concepts can retain their identity across radical changes in theory." (192) But he also thinks that "the intension and extension of concepts change as the result of changes in theory." (192) So what, exactly, *is* retained when a concept retains its identity across radical changes in theory? Though Griffiths has lots to say on the topic, we must confess that we can offer no clear account of how he would answer this question. On our view, the account of concept identity that Griffiths sets out (in chapter seven) is the least successful and most obscure part of the book. But we are also inclined to think that this shortcoming in no way undermines what we take to be Griffiths' most exciting and important conclusions.

What is really central to Griffiths' account is the claim that the *theory* (or *theories*) that are emerging in the scientific study of the emotions do not map comfortably onto our current commonsensical thinking about the emotions. Those scientific theories will pick out categories that commonsense does not pick out, draw boundaries that commonsense does not draw, recognize distinctions that commonsense does not recognize, and it will use terms and concepts that refer to these categories. Will *some* of our current "folk" concepts survive in this radically transformed theoretical environment? Will some of the concepts in this new science really be *identical* to folk concepts, despite the fact that the claims made using those concepts are quite different from claims we now make? Will some future theoretical *terms* turn out to have the same extensions (or intensions) as current "folk" terms, even though what we would now *judge* to fall in the extensions of these terms is quite different from what those who embrace the new theory will judge to fall within the extension of any of its terms? The answers, of course, depend in large measure on which accounts of reference and of concept identity turn out to be correct. And while many philosophers

have strong views on these matters, we're inclined to suspect that there may be nothing to be correct *about*. We are far from sure that there *is* a right theory of reference or of concept identity.⁽¹⁾ But, and this is the important point, even if our suspicions are correct, it should hardly matter for Griffiths. On our reading of his view, what is important is the claim that the emerging theoretical landscape will be very different from the one suggested by commonsense. Questions about co-reference and concept identity are at best a philosophical sideshow.

4. Against Particular Emotions

So far we have focused on Griffiths' arguments about the scientific viability of the superordinate concept of emotion. We turn now to his two rather more speculative arguments about the fate of some of some particular emotion concepts. Both of these arguments are motivated by Griffiths' view of levels of explanation in the life sciences. He picks out four levels in biology (pp. 216-218). The topmost level is that of population genetics, whose laws allow generalizations across populations of any species. We then have the ecological level at which traits are grouped in terms of task descriptions like mate selection, regulation of body temperature or communication over an acoustic channel. These tasks may be variously realized by different structures in different lineages. Following Sterelny (1990) Griffiths argues that the level of task description traditionally recognized by cognitive psychologists is an ecological level.

Ecological traits are classified by the adaptive problems they solve. One level down from the ecological we have in biology the cladistic or "natural-historical" level. Whereas the ecological level groups traits in purposive terms a cladistic taxonomy looks for homologies. These are biological traits shared by groups in virtue of descent from a common ancestor. The parallelism between biology and psychology now starts to break down. Whereas ecological kinds in biology correspond to task descriptions in psychology, cladistic kinds may be realized in psychology either by computational programs or by their implementations in the brain. Both the second and third arguments which Griffiths runs turn on his contention that our folk emotion concepts pick out clusters of features which belong to different levels. Griffiths' first argument envisaged a "horizontal" fragmentation of the folk category of emotion into three others at the same general level of explanation. His latter two arguments suppose that particular emotion concepts will be torn into strata as we recognize different levels of explanation within one vernacular kind.

5. The Second Argument - Cladistics and Ecology

Our folk concepts, Griffiths thinks, may have both ecological and cladistic features which cannot be conjoined in one theory. If psychology is to be integrated into the life sciences these different levels must be separated out. In that case, many folk-psychological concepts will not be clearly identifiable with any one successor concept, since different features of the folk concept will be preserved at different levels of explanation in the mature science. One level is ecological. Griffiths thinks such a general ecological science trading in context-invariant adaptive pressures may never develop, but even if does it will apply only at the level of task descriptions. These can be multiply realized by very different computational or physiological mechanisms. So, for example, although disgust may be a good ecological category, we cannot extrapolate from our knowledge of the mechanisms subserving disgust in birds to the mechanisms subserving disgust in rats because they have different computational and neurological ways of satisfying the disgust task description.

Cladistic classifications, however, are likelier to focus on the same underlying mechanisms. Griffiths argues that two closely related species will probably share since they inherited it from a common ancestor. Even if they use the mechanism for different adaptive purposes we will still be able to generalize across

them at the architectural level. The point may emerge more clearly if we imagine a hypothetical case of a rabbitlike population speciating as it moves into two different environments. We may suppose that the ears of the two new species remain similar physically, but are used differently. One successor species uses its large ears for body-temperature regulation in a hot, dry environment whereas the other lives in a crowded ecosystem in which its large ears are useful aids to improved hearing. Hence we may find generalizations at different levels - one cladistic, one ecological - invoking the successors of a single vernacular concept. But these successor concepts will not be co-extensive, and there is no theoretical reason to identify the vernacular concept with either of its successors.

6. The Third (Dennettian) Argument

Griffiths' third argument aims to show that computational homologies might be physically implemented in different ways in different lineages. Some of the features which Griffiths claims to be "constitutive" for a concept might be computational whereas other constitutive features, the qualitative ones, could be the results of our brain chemistry. This means that particular emotions may not be preserved as taxa even within cladistic schemes of classification, because the computational and implementational levels in psychology which the cladistic level in biology maps on to may come apart. Griffiths uses Dennett's discussion of pain as a model for the analysis which he expects affect programs (and the concept of mood) to receive. Dennett (1978) suggests that our everyday concept of pain is ultimately incoherent, since it gives rise to contradictory intuitions. On the one hand, we think that nothing could be in pain unless it feels a certain way. But we can also imagine attributing pain to creatures based on their behavior (such as aversive responses to harmful stimuli) and the computational architecture that subserves that behavior. These two components of the concept of pain are not co-extensive. As Griffiths notes (pp. 226, 256), Dennett argues for a computational program -- generating an aversion or avoidance response -- that could exist across species, together with a distinctively human component, which in the case of pain is qualitative. The latter is realized by neural structures occurring in only a few species. In the case of humans, and maybe some other species with a neomammalian brain, the two components coincide, but this coincidence is not necessary. Our concept of pain applies to the contingent conjunction of two different categories. So we don't know whether or not to apply it in the case of robots, extraterrestrials, octopi and so on, since in those cases the behavioral program does not coincide with a humanlike neurophysiology. "Do squid or computers really feel pain? The uncertainty of ordinary speakers on this question leaves it unclear which of the theoretical kinds should be identified with pain" (226). There is certainly no reason to suppose that the computational program underlying a given affect program is implemented in the same brain chemistry in all the species which share the computational program. It is unclear whether we should identify the vernacular concept with the computational or implementational level of explanation. If we cannot find a way of identifying the folk concept with a scientific successor, then we will not get a general theory of what we take our affective states to be.

7. Extension to psychopathology

So we have two "vertical" arguments which aim to pull particular emotion concepts apart at different levels, and one general "horizontal" argument applying to the superordinate category of emotion. If Griffiths' arguments are good ones -- and with the qualifications we've noted, we think they are -- then we should expect them to apply to many other psychological concepts. And in fact psychopathology appears to offer an arena in which Griffiths' arguments are particularly likely to work. The conceptual havoc which Griffiths expects from an evolutionarily-informed psychology of emotion is also likely to follow from an evolutionary psychology of mental disorder. We should expect very similar patterns of conceptual transformation to affect our understanding of a number of pathologies, and perhaps the superordinate concept of mental disorder itself.

The general picture in psychopathology exhibits a conceptual territory that is very similar to the one Griffiths recounts in the psychology of emotion. Phobias, for example, have been seen as extreme forms of adaptive responses to specific input; this view characterizes normal fears in much the same way that the affect program views basic emotions (Seligman 1970; Garcia & Koelling 1966; Marks 1987). There are likely to be socially constructed "mental disorder" roles too. Indeed, Griffiths notes the strong parallels between Ian Hacking's (1995) treatment of Multiple Personality Disorder and the social role account of emotions like amok. In addition, there are lots of mental disorders with a higher, less encapsulated, cognitive dimension. Depression, for example, can be diagnosed on the basis of patterns of pessimistic, self-critical thought with which it is reliably correlated, whereas phobias are classified in terms of their stimuli.

Given the strong similarity between psychopathology and the way Griffiths treats emotion, we might expect the eliminativist arguments in Griffiths' book to go through for psychopathology also. In the following section we will assemble some evidence that suggests that there are indeed some currently recognized disorders conceptualized in ways which are undermined by Griffiths' arguments. A more general issue is whether there can be no science of "psychopathology" for the same reasons that Griffiths denies there can be a science of "emotion". We'll take that up in the Conclusion.

8. Specific pathologies

Particular pathologies may show the same sort of pattern that Griffiths suggests for pain and mood - that is, our everyday concept may pick out conjunctions of properties at different levels of explanation within an evolutionary psychology. Consider again the example of depression. We noted above that depression can be diagnosed in humans based on cognitive factors. But there is also a behavioral profile of psychomotor retardation which is part of the contemporary psychiatric concept of depression, including symptoms like excessive (or insufficient) sleep and psychomotor retardation. This suite of behaviors is also sufficient for diagnosing depression, and much the same behavioral profile has been found in other species, ranging from rats to chimpanzees. In non-human primates the behavioral profile regularly appears in individuals who have lost contests for dominance. These animals also exhibit a dip in their levels of the neurotransmitter serotonin, as do many depressed humans.

That primates who have lost contests for dominance show the same neurotransmitter changes as depressed humans is a matter of shared neurochemistry. But the neurochemical changes (and the "blue" feelings which they may cause) may not be necessary to implement the behavioral profile or cognitive responses characteristic of depression. According to the social competition theory of depression, the behavioral and cognitive components of depression are adaptive computational responses to loss of status (Price et al. 1994). A central claim of the theory is that depression causes one to rethink one's social strategies when they are unsuccessful in enhancing status. While the behavioral and cognitive aspects of depression may be caused by a suite of neurochemical changes in us and related species, these changes and the low affect that they also cause, may be only contingently associated with the computational mechanism of depression, despite its central role in our talk about depression.

So, is depression a cognitive, qualitative or behavioral category? The moves that might be made here are analogous to those in Dennett's pain argument but with cognitive processes as well as qualia to consider. Are we, or are we not, to say that rats or primates may become depressed? They share behavior and brain chemistry with people who pay a fortune for psychotherapy, yet it is singularly implausible to attribute to them not only feeling blue but also the kinds of elaborate self-deprecatory thoughts which haunt our depressed conspecifics. We have here, as in the case of pain, an example of a human concept several components of which have wider-than-human extensions. Perhaps the cognitive patterns characteristic of human depression are computational processes which differ in degree only from those of primates - we

can just think about more things than they can. In that case, we might still have a shared adaptive response to loss of status. Or there might be distinctively human aspects to depressive cognition which have evolved in response to specific selection pressures. There is also the question of affect - the feel of depression - to consider. In a Dennettian spirit, we might conjure a robot who exhibits both the behavioral and the cognitive components of depression but who, unlike us, does not feel blue. Our depression concept, it seems, mixes up ecological, computational and neural classifications (that is, classifications in terms of adaptive niches, computational programs and neural implementation) and a future science may well have to pull them apart.

Though less familiar, Panic Disorder is an even clearer candidate for the application of a Griffiths-style eliminativist argument. Its symptomatology includes both physical and cognitive elements. Clark (1988, 1996) and McNally (1994) have developed theories which rely on cognitive misinterpretations of physical arousal. Quite ordinary bodily sensations are interpreted as signs of incipient death, madness or loss of control. But it has also been claimed that primates can show "affective distress," after lactate infusion for example, which also tends to induce panic in human sufferers (Sunderland, Friedman & Rosenblum 1989). If panic is a cognitive kind it is hard to suppose that these apes are panicking; do they think they are about to die or lose control, let alone become detached from themselves (depersonalized) or from reality (derealized)? These cognitive traits are all symptoms of panic. According to current psychiatry, however, there are also straightforward physical symptoms which are diagnostic of panic -- they include sweating, difficulty in breathing and accelerated heart rate -- and there is certainly no problem in ascribing these to primates. There are also symptoms with a qualitative or phenomenological component which pull our intuitions in both directions, for Dennettian reasons; these include nausea, dizziness, lightheadedness, and distress.

What are we to make of the multifarious bases for diagnosing panic disorder? The physical components might be an example of a shared autonomic response in several lineages. If Clark's influential theory is correct though, and panic disorder is fundamentally a matter of cognitive misinterpretation of bodily responses, then only in humans is there the right cognitive profile. The currently mixed physical/cognitive symptomatology seems to lump together distinctive human cognitive responses with extra-human components. Within one psychopathological concept we have a mixture of physiological response and higher cognitive activity. Moreover, this is not just a folk concept. Panic is a distinctive clinical concept with its own body of theory, which emerged from another one, anxiety neurosis, in 1959. Despite these scientific trappings it seems unlikely that we can generalize from the cognitive profile of human panickers across the various lineages which might share the physiology of panic. If we take the physiology as constitutive of panic we will be left missing important features of the relevant human psychology. In this case Griffiths would probably issue the same recommendation about the clinical concept of panic as Frank Lloyd Wright reportedly did when the city fathers of Pittsburgh asked him how to improve their town: "Abandon it".

9. Conclusion

We've seen that there are some mental disorders whose concepts exhibit the same mishmash of levels of explanation that Griffiths sees in our emotion concepts. Given the general revisionary thrust which Griffiths sees in a mature evolutionary psychology this is not surprising. What can be said, now, about the superordinate concepts of psychopathology -- concepts like mental disorder? If, as we suspect, the examples in the previous section are just the tip of the iceberg, then the prospect of a diversity of kinds of mental illness opens up, with cognitive, affective, biological or other bases, each with proprietary theories of their causes and mechanisms. Some of these will be distinctively human, and some will be very strongly homologous to traits in other species. By analogy with Griffiths' first argument, against the superordinate concept of emotion, we should expect the conclusion that our concept of mental disorder is

a prime candidate for elimination.

This conclusion might well be premature, however. Doubtless there are many ways for the mind to go wrong if it is an evolved computational system. What might unite these misbehaviors is a normative feature that they all share -- they are what we think of as ways of going *wrong*. It is extremely unlikely that a satisfactory non-normative conception of mental illness could be developed and it is the normative character of our conceptions of psychological disorder that will continue to unify our thinking about it, even if we end up with very many different scientific approaches to the taxa it includes. The scientific fragmentation of a concept might be counteracted by the non-scientific needs it serves. These give it a unity, robustness and longevity which it would lack on purely theoretical grounds, and enable us to make sense of a diverse range of behaviors by keeping their normative and social dimension in mind.

References

Boyd, R. 1989. What realism implies and what it does not, *Dialectica*, 43, 5-29.

Boyd, R. 1991. Realism, anti-foundationalism and the enthusiasm for natural kinds, *Philosophical Studies*, 61, 127-148.

Churchland, P. M. 1981. Eliminative materialism and the propositional attitudes, *Journal of Philosophy*, 78, 67-90.

Clark, D. 1988. A cognitive model of panic attacks. In S. Rachman and J. Maser (eds) *Panic: Psychological Perspectives*. Hillsdale NJ: Erlbaum.

Clark, D. 1996. Panic disorder and social phobia. In D. Clark & C. Fairburn (eds) *The Science and Practice of Cognitive Behavior Therapy*. Oxford: Oxford University Press.

Darwin, C. 1872. *The Expression of the Emotions in Man and Animals*. Reprinted, Chicago, University of Chicago Press 1965.

Dennett, D. 1978. Why you can't make a computer that feels pain. In *Brainstorms*. Brighton: Harvester

Dennett, D. 1988. Quining qualia. In A. Marcel and E. Bisiach (eds) *Consciousness in Contemporary Science*. New York: Oxford University Press.

Ekman P. 1971. Universals and cultural differences in facial expressions of emotion. In J. K. Cole (ed) *Nebraska Symposium on Motivation* 4. Lincoln: University of Nebraska Press.

Ekman, P. 1980. *The Face of Man*. New York: Garland.

Ekman, P. 1992. Are there basic emotions? *Psychological Review* 99 (3); 550-53.

Ekman, P. and Friesen, W. 1971. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17, 124-29.

Fodor, J. 1983. *Modularity of Mind*. Cambridge; MIT Press.

Frank, R. 1988. *Passions Within Reason*. New York; Norton.

- Garcia, J. & Koelling, R. 1966. Relation of cue to consequence in avoidance learning. *Psychonomic Science*, 4, 123-4
- Griffiths, P. 1997. *What Emotions Really Are*. Chicago; University of Chicago Press.
- Hacking, I. 1995. *Rewriting The Soul*. Princeton: Princeton University Press
- Marks, I. 1987. *Fears, Phobias and Rituals*. Oxford: Oxford University Press
- McNally, R. 1994. *Panic Disorder*. New York: Guilford.
- Pinker, S. 1997. *How the Mind Works*. New York: Norton.
- Price, J., Sloman, L., Gardner, R., Gilbert, P. & Rohde, P. 1994. The Social Competition Hypothesis of Depression. *British Journal of Psychiatry*, 164, 309-315.
- Seligman, M. 1970. On the generality of the laws of learning. *Psychological Review*, 77, 406-418.
- Sterelny, K. 1990. *The Representational Theory of Mind*. Oxford: Blackwell.
- Stich, S. 1983. *From Folk Psychology to Cognitive Science*. Cambridge, MA: Bradford Books / MIT Press.
- Stich, S. 1996. *Deconstructing the Mind*. Oxford: Oxford University Press.
- Sunderland, G., Friedman, S. & Rosenblum, L. (1989). Imipramine and alprazolam treatment of lactate-induced acute endogenous distress in nonhuman primates. *American Journal of Psychiatry*, 146, 1044-1047.
- Tooby, J. & Cosmides, L. 1990. The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology*, 11, 375-424,
- Wakefield, J. 1992. The concept of mental disorder. *American Psychologist* 47; 373-388.
- Wakefield, J. 1997. Diagnosing DSM-IV, Part 1: DSM-IV and the concept of disorder. *Behavior Research and Therapy*, 35; 633-649.

NOTES

¹ For further discussion of these suspicions, see Stich (1996), Ch. 1.