Chapter 7

Functions and Norms

> Nothing makes sense in biology, except in the light of evolution
> — Dobzhansky

> For a long time the biologist treated teleology as he would a woman he could not do without, but did not care to be seen with in public. The concept of [genetic] programme has made an honest woman of teleology.
> — Jacob, *The Logic of Living Systems*

In chapters 5 and 6 I argued that we can define the truth and meaning of a belief in terms of the successful actions that it plays a role in. This then left the problem of determining what actions should be considered successful, and why? Darwin offers one possible answer; and in this chapter I discuss why, and how, we should use it.

7.1 Functional Explanation and Darwinian Norms

According to Darwin, a product of natural selection is successful if it contributes to the reproduction of the organism that carries it[1]. This criterion of success, or *norm*, can be applied to traits, such as hearts, or particular events, such as a single beat of a heart: if a vertebrate had no heart, or if it had a coronary attack, then it would be less likely to reproduce than one that did. We can also apply this criterion of success quite naturally to *behaviour*, and to particular actions and the internal representations that underlie them. For example, a frog is more likely to reproduce if it catches a fly than if it misses. Therefore those occasions on which the neurons in its retinotopic map fire when a fly is present can be considered successes: its proto-belief about the presence of an 'eatable thing' was correct. (We shall see how we can extend this criterion of correctness to the products of social evolution, including human beliefs, in the final chapter.)

Why should we use this Darwinian norm rather than any other? The reason is that it makes possible a new way of making sense of systems that evolve over time. Remember that the original problem Darwin pondered was that the natural world seemed so *successful*. Every organism

[1] Or, more accurately, if it contributes to the ability of that organism to pass on its genes. The issue of the relationship between the replication of genes and the reproduction of organisms will be the subject of chapters 8 and 9.

seemed perfectly adapted to its ecological niche, and every part of every organism seemed optimally designed to allow the whole to survive and flourish. Every part of nature seemed to fulfill a *function* for the whole. Paley drew the obvious conclusion that this design was God-given, but Darwin dared to suggest that this order could be explained by the mechanism of natural selection. In other words Darwin *naturalised* functional descriptions of nature (in the sense defined in chapter 3). Paley knew *that* the natural world was functionally organised, but only Darwin could explain *why*.

Naturalisation normally works 'from below' by showing how the behaviour of a whole is produced by the workings of its parts. But once the mechanism of natural selection is in place then a new kind of explanation is possible, namely *functional explanation*. It is now possible to explain the workings of the parts in terms of the contribution that they make to the whole. For example, why do hearts pump blood? The usual naturalised, i.e. *mechanical*, explanation is that hearts pump because they have chambers and valves and rhythmic control systems. But the functional explanation is that hearts pump because this helps the animal survive. Functional explanation is explanation from above, rather than from below; it accounts for the existence of an entity in terms of its effects, rather than its causes. This type of explanation only works because natural selection can explain why effects of that type in the past have led to the existence of entities of that type today. Hearts exist *because* they (or rather, their ancestors) pump blood, and so on. Millikan defines this as the Direct Proper Function of an entity (1984): it is what the entity exists in order to do; it is the Purpose of the entity; what it exists *for*.

The advantage of using the Darwinian norm, rather than any other, is that it enables us make sense of the history of the system. This is because we are using the same criterion of success that evolution uses. The Darwinian norm is a force that operates within nature, rather than a criterion that we are imposing on nature from without. The Darwinian norm is a *naturalised* norm. Consider this example. It is fairly crass, but it makes the distinction clear. Suppose we are contemplating a bluebell wood in spring, and choose to analyse it using the norm of 'looking pretty'. The wood succeeds in meeting that norm, and this in turn defines functions for its various parts: the bluebells have the function of providing a splash of colour, the leaves on the trees fulfill the function of providing dappled sunlight, and the worms in the soil fulfill the function of aerating the soil and keeping the whole thing healthy. This norm tells us *something* about the wood; but it obviously tells us little about its history, about *why* it exists in its current form. For this we have to use the Darwinian norm. According to Darwin the function of the bluebells is to attract pollinating insects, the function of the leaves is to photosynthesise and shade out competitors, and the function of the worm is to make more worms: the Darwinian function explains why each of these entities exist.

The theory of natural selection allows us to use the function of an entity to explain its history. However, if we want to understand the history of a system there are two distinct questions we can ask. The first *how did the system get like it is now?*, i.e. we can ask about the past history of the system. But we can also ask *what will happen to the system in the future?*, i.e. we can ask about its *future* history. Now as long as there are no major disruptions to this system, and as long as all change is incremental and gradual, then the answer to the first question will also answer the second. Bluebells exist now because they attracted insects in the past, and in an undisturbed wood this situation is unlikely to change radically. But suppose that the bluebells were stolen by an avid gardener, transplanted to a greenhouse, and artificially pollinated in order to grow new commercial varieties. Now what matters for the future history of the flowers is their ability to attract customers, rather than insects. The functions of an evolved system change as soon as a new selective regime comes into place. The Proper function — i.e. the function defined by past history — has remained the same, but the current Darwinian function has changed. We are using the same norm in each case — i.e. reproductive success — but in one case to understand how the system got to how it is now, and in the other to explain how it will change in the future. The same norm yields different functions in each case.

7.2 The Function of 'Function'

What is the function of the concept of 'function'? What explanatory purpose do we want to use it for? Is it supposed to explain the past history of a system, or its future? Most of the discussion in the philosophical literature has tried to define a single, unitary, concept of function (see for example Wright (1973), Cummins (1975), Neander (1991), Bigelow and Pargetter (1987), and Millikan (1984)(1993)(1999).) But there is no reason to expect that a single concept can answer both questions.

According to Millikan (and Wright) the function of an entity *must* account for its occurrence: it is the ability of the ancestors of an entity to fulfill that function that explains why it exists in that form today. Thus Millikan's Proper Function is very similar to what we usually describe as the 'purpose' of an entity; i.e. what that entity exists *for*. In other words it is a *teleological* concept of function, similar to Aristotle's Final Cause which 'pulled' natural events towards some future goal. But Darwin showed how the same functional organisation could be explained by the force of natural selection 'pushing' from the past. Therefore, as Mayr (1982) argues, Darwin's teleology is not strictly the same as Aristotle's but is more of an *as if* teleology — what he calls *teleonomy* — i.e. an explanation of why teleological descriptions work so well, rather than a naturalisation of final cause *per se*.[2]

It is often assumed that functions *must* be teleological. However there is another way of defining norms and functions based, not on the history of an entity, but on its effects. Such approaches are generally labelled *consequentialist*. Anscombe (1958) originally invoked this term to describe theories of ethical norms, such as Mill and Bentham's Utilitarianism, or Aristotelian and Marxist ethics (Miller, 1981). To put it crudely, according to these theories an act is considered (ethically) correct if it contributes to a state of maximal happiness, *eudaimonia*, or communism, respectively. Conversely, if the act leads to a state that one considers bad (such as maximal unhappiness, dystopia, or to the continuance of capitalism) then the act was wrong. And if the act makes no positive contribution either way then it is ethically neutral. Given these norms for judging an act then, in certain circumstance, one can derive a non-teleological sense of ought. For example if one is given a number of choices then we can use consequentialist criteria for judging which is the best. That act is then what one *ought* to do. Thus consequentialist theories start from norms and then derive oughts, rather than *vice versa* as teleological theories do.

[2] Actually this is not quite true. Aristotle himself considered a proto-Darwinian theory of evolutionary teleology that had been proposed by Empedocles (see Aristotle's *Physics*, II.8). Aristotle agrees that this would be an example of teleological causation but rejects it on purely empirical grounds since he saw no evidence of descent with modification in nature. Thus purposiveness is not an *a priori* part of Aristotle's teleology, just an *a posteriori* explanation of it.

(Obviously the possibility of deriving oughts from norms depends on assumptions about agency, i.e. about the metaphysics of counterfactual assertions that an agent could have acted differently. Does it make sense to say that we ought to do something we are physically incapable of? This problem becomes relevant in section 11.3 where I discuss the role of utopias in forming ethical judgements, but does not affect the current discussion.)

The norms that consequentialism defines for an entity are derived from the norms of its consequences. Bentham and Mill believed that maximising happiness was a good thing so they concluded that acts that contribute to it are also good. Similarly it is the goodness of Aristotle's *eudaimonia* and of Marx's communism that enabled them to define norms for social acts. In short, ends justify means. The same basic strategy works for biological functions: we can define the correct function of an entity on the basis of its Darwinian consequences, rather than its Darwinian history. The historical function of a transplanted bluebell is to attract insects, but its consequentialist function is to attract customers.

Consequentialist and historical norms serve different purposes. They are complementary and so we should use them both. However it is interesting to note that in many cases the consequentialist approach is of more practical use to biologists than the teleological.

For example, what is the function of a behaviour that an animal learns in order to survive in novel circumstances that were not part of its evolutionary history? Millikan argues that all such lifetime innovations must be the products of innate learning mechanisms — such as those discovered underlying Skinneran operant conditioning in *Aplysia* (Carew et al., 1983)(Hawkins et al., 1983). These innate mechanisms will have some well-defined evolutionary history, and hence a 'direct' proper function, from which the function of the lifetime innovations may be derived. Therefore the existence of derived proper functions for learned behaviours depends on there being an innate 'end to flexibility' in the brain, as Millikan puts it (1984, p46–48).

An historical (i.e. teleological) analysis of the functional organisation of brains encourages the view that the brain is divided into discrete and innate modules with highly specialised functions (Millikan, 1993, p49) — caricatured as the 'Swiss army knife' model. Now there is certainly a growing body of evidence for such modularisation (Pinker, 1998). But the plain fact is that we should not *assume* it is universal, nor are we ever to likely to be able to determine the evolutionary history of particular brain modules in practice (Lewontin, 1998). And nor must development 'bottom out' in innate mechanisms in quite the way that Millikan requires (Elman, 1995). A consequentialist view of function offers a more practical alternative. The progress that we have so far made in investigating the functional organisation of brains has mostly come from an understanding of the role that they play in the current behaviour and development of the organism, rather than their individual evolutionary history (Kennedy et al., 1995). It is still necessary to use the Darwinian norm to make sense of the functional organisation of brains, but we usually do this without knowing much about their evolutionary history. Of course once we uncover some aspect of how brains work — i.e. how they enable the animal survive and flourish and so meet the Darwinian norm — then the next obvious question is to ask how evolution produced those mechanisms. But, in purely practical terms, a consequentialist analysis of function is of more immediate use to a neuroscientist than an historical one.

7.3 The Function of Behaviour

The consequentialist view of function is also consistent with that which underpins the investigation of animal behaviour. Tinbergen (1951) famously defined four questions that the ethologist — or any other biologist — can ask of a trait: how is it caused, how does it develop, what does it do, and how did it evolve? For example, suppose we ask 'why does our thumb move in a different way to the other fingers?' We might give an answer in terms of the anatomy of the hand (causal), or in terms of its embryology (developmental), or in terms of its evolutionary history (historical), or in terms of what it currently enables us to do (functional). 'All of these answers would be correct: no one would be complete' (Hinde, 1982, p21). Tinbergen went on to explain the functional question as follows.

> In attacking our problem, let us start from observables — i.e., from behaviour. But instead of studying its causes we shall study its effects; in other words, rather than look back in time, as we do when studying causation, we investigate what happens as a consequence of the observed behaviour. I should like to stress that this *is* an investigation of cause-effect relationships, which as such requires experimental study as well as observation and speculation; it differs from the study of behaviour causation merely by the fact that the observed behaviour is the cause, and that its effects are studied; we follow events with time instead of preceding events and we determine an animal's success. (1965, p521)

And Tinbergen explicitly contrasts this investigation with that into the *origins* of behaviour, arguing that 'they can suggest (*though no more than that*) the selection pressures that have in the past molded the species to what it is now' [p523]. Of course there is no rigid distinction in practice between the forces that originally moulded a behaviour and those that now maintain it. But Tinbergen is arguing that there is, at least, a clear *conceptual* distinction between historical and consequentialist function, and therefore we can determine the consequentialist function that a behaviour serves now in perfect ignorance of its evolutionary past.

For the ethologist the key problem in determining the function of a behaviour is to determine which of its many causal consequences actually help the organism reproduce — and so succeed in the Darwinian sense. And, as Hinde notes, to answer this question it is strictly necessary to assess reproductive success with and without the character (1975, p6). This is rarely possible in naturally occurring populations — where behaviours tend to be universal — and so ethologists experimentally intervene to test which factors are significant. A classic example of this is Tinbergen's investigation of the 'fanning' of the male stickleback which is done mainly when there are developing eggs in the nest. He found that when fanning is prevented, or when it is allowed but the nest is covered with a watch glass, the eggs die. However they develop normally in the absence of the male provided freshly aerated water is pumped over the nest, suggesting that the function of fanning is to ventilate the eggs (Tinbergen, 1951).

Millikan uses this methodological problem as an argument against Bigelow and Pargetter's consequentialist account of function, in which they attribute a function to a trait 'when it confers a survival enhancing propensity on a creature that possesses it' (1987, p192). Millikan interprets this as meaning 'a trait that enhances or would enhance survival is one where, on average over the *actual* individuals in the species, having it would produce a more fit individual than not having it' (1993, p39). She concludes that if the trait is universal in a population, with no current competitors, then the concept of 'enhanced propensity of survival' is meaningless. So if *all* male sticklebacks fan then that fanning cannot be said to enhance their fitness, since there are no non-fanning sticklebacks to be enhanced compared to.

The counter-argument to this objection depends on the fact that fitness is operationally defined as the *expected* number of offspring produced by a class of individuals; it is therefore a *disposition* to reproduce. And as with other dispositional properties problems arise if we understand them in terms of actual outcomes. Recall Carnap's argument that if the dispositional property of 'being soluble' is defined as 'dissolving when in water' then the claim that '*X* dissolved because it was soluble' is tautologous (section 3.3). Similarly, if fitness is defined as the actual rate of reproduction then we face the old canard that 'survival of the fittest' is also tautologous. Of course the answer is that 'solubility' describes a property of a substance *in virtue of which* it dissolves, and similarly 'fitness' describes properties of organisms in virtue of which they differentially reproduce. Fitness is a measure of the *reproductive power* of an organism

It may be the case that we only discover which substances are soluble by putting them in water, but we should not confuse the way that we measure a property with the facts in virtue of which an entity holds it. Similarly, the function of the stickleback fanning is not defined by the fact that their eggs died when Tinbergen covered them with a watch glass. It is the causal process of aeration (along with its effects on the eggs) that constitutes the function of the fanning, not the differences between fanned eggs and covered ones. In short, it is possible to define the consequentialist function of a trait in terms of the role that it plays in the reproduction of the organism now, even in the absence of a range of contemporary (or historical) alternatives.

7.4 Conclusion

In this chapter I have argued that we should use the Darwinian norm to judge the success of a biological trait because this norm naturalises functional explanations of natural history, and so fits the 'idea of the good' for science outlined in chapter 3. But we can use this norm in two ways. The first is to explain the origin of biological traits. And the second is to explain the future fate of the organisms that possess them.

In the final chapter I will argue that it is possible to use the same criterion to define norms for *social* traits — including language, science, and ethics — but this requires that a Darwinian-type theory of natural selection can be applied to social history. This is the subject of chapter 10. But before then it is necessary to define more precisely what we mean by natural selection. And this is the subject of the next two chapters.