

Chapter 8

Function, dysfunction, and adaptation?

Kelly Roe and Dominic Murphy

Jerome Wakefield has argued that mental disorders are harmful dysfunctions. In claiming to capture people's intuitions, however, Wakefield argues that people think of mental disorders along the lines of a two-stage model. Now, the two-stage model sees psychiatry as a branch of medicine, in that it rests on a scientific account of the normal function of the human mind/brain. If psychiatry is continuous with medicine and physiology in this way, its analysis of function and malfunction should reflect that continuity. In this chapter we argue that the way the relevant biomedical sciences determine function does not presume a selectionist concept of function. We argue that the relevant accounts of function are those drawn from mechanistic explanation rather than historical explanation; the life sciences ask all sorts of questions, but the questions which medicine asks are not those which a selectionist account of function can answer. The chapter contrasts Wakefield's (and others') selectionist (or historical) view with a causal (or mechanistic) theory of functions - the systemic capacity view, which sees the function of a system as its contribution to the maintenance of the larger system in which it is embedded. The authors argue that the systemic capacity view provides a better account of how functions are understood and functional claims tested in medicine and physiology. However, they end by wondering whether there is in fact any role for science to play in determining what the overall functional state of an organism is, and hence whether science can draw the line between health and illness in the way that the two-stage system takes for granted.

8.1 Introduction

The medical model of psychiatry sees mental disorder as a bio-medical disorder. The difference in subject matter between psychiatry and another branch of medicine is supposed to reflect the fact that the disorder is mental rather than, say, cardiovascular, but when we call a bodily system disordered we mean exactly the same thing throughout medicine. In this chapter we attend to the puzzle of what “dysfunction” means by looking at rival conceptions of function in the context of the bio-medical conception of disorder. Throughout, our concerns will be dictated by what we call (following Murphy 2006) the two-stage view of the concept of mental disorder, which distinguishes facts about dysfunction from normative judgements about the adverse consequences of dysfunction.

The problem of offering a naturalistic or scientifically respectable account of biological function and dysfunction has long been a concern in the philosophy of biology. The debate on dysfunction within psychiatry and philosophy of psychiatry has usually gone on without looking at this wider philosophical literature, but we will rely on it. After introducing the “two-stage view” and saying something about the picture of psychiatry we presuppose, we will discuss two accounts of biological function and assess the prospects of incorporating them into psychiatry. We will argue that the *systemic capacity view* of biological function and dysfunction seems better suited than the *selectionist view* to capture what bio-medical scientists take themselves to be doing. We will then consider the objection that neither account can explain dysfunction. Finally, we ask whether biological dysfunction is in fact necessary for psychiatric and bio-medical disorder. We do not offer a comprehensive overview of the function debate; we only care about the features of it that seem relevant to psychiatry.

8.2 The two-stage view

The two-stage view is the most popular account of psychiatric disorder among theorists who deny that ascriptions of mental illness are entirely normative. It was introduced by Wakefield (1992), who borrowed heavily from earlier work by Boorse (1975, 1976a). Two-stage theorists hold that there are two individually necessary and jointly sufficient conditions for disorder. First, there is a biological dysfunction. Second, the dysfunction must result in harm to the individual and/or society, as judged by prevailing social norms. “Harm” is uncontroversially a normative notion, but psychiatric dysfunction is assumed to be a matter for medicine to establish, just as it would establish that an esophagus is dysfunctional.

The two-stage view aims for a middle ground between (i) a scientism that says psychiatry has no role for values at all and (ii) a constructivist claim that our judgments that a person is disordered depend entirely on their having violated some norm. The view supposedly respects both the role of science in psychiatry and that of social norms. However, the two-stage view faces two sets of conceptual problems. First, there are the difficulties involved in justifying the intuition that science plays a role in the discovery of objective facts about disorder. In the final section of this chapter we will discuss what that role is. Second, we have the intuition that norms have a role to play in whether an individual is harmed by his or her dysfunction. This is thought to have normative implications along the lines of rights and duties to treatment. Now, unpacking the notion of harm is at least as problematic as unpacking the notions of function and dysfunction (e.g., see De Block 2008). Perhaps one kind of harm is simply physical injury, but the relevant concept of harm involves judgments about the quality of someone's life. These judgments need to be sensitive to both the individual's own needs and goals, and the ideas about well-being that feature in the wider society. Here we put "harm" aside to ask what the best way to think about dysfunction might be.

Boorse (1975) distinguished "disease" from "illness". Disease is failure to function as designed, meaning that some biological system does not conform to the "species-typical design" of humans. Illness is a matter of judgment that a disease is undesirable, entitles one to special treatment, or excuses bad behavior. Boorse proposed an account of functions as properties of systems designed to contribute to survival and reproduction (1976a, p. 62–63). Wakefield advocates a selectionist view of function, according to which the function of a psychological system is what natural selection designed it to do (see Chapter 5).

So the question we take up has to do with function talk within a medical context. We assume for the sake of argument that psychiatry is a branch of medicine in a strong sense (Murphy 2009), so that mental illnesses are caused by distinctive cognitive neurophysiological pathologies. An example of this view is Nancy Andreasen's (2001, p. 172–6) argument that an explanation of mental illness will ultimately cite destructive processes in brain systems, just as bodily diseases are explained by such processes in other organs. The process at issue need not be entirely endogenous: it can mediate the effects of cultural forces or other environmental risk factors. Nor does the cause of disorder have to completely destroy a brain system: it may be enough to put the system into a stable but chronically dysregulated state. The way to understand the scientific part of the two-stage view, then, is that there are, in psychiatry, phenomena that fit the conception of disease as a destructive process that predominates in biomedicine generally. Our question is, what does it mean to say that psychiatry

is concerned with departures from the natural functions of the human mind/brain? Put differently, we aim to understand the idea of dysfunction in biomedical contexts. In the next section, we will consider both evolutionary and systemic capacity views of function with an eye to their prospects of accounting for dysfunction as a naturalistic grounding for bio-medical norms.

8.3 Theories of function

The modern literature on function stems from two seminal papers. Wright (1973) argued that ascriptions of function to a structure are causal-historical. His analysis applies to any structure that participates in a selection process and thus is not explicitly evolutionary. Millikan (1984) and Neander (1991) built an evolutionary analysis of function on Wright's foundations. Cummins (1975) was the other key paper. Cummins's concept of function was not historical but causal. He understood the function of an entity to be the causal contribution it makes to the operation of the overall system(s) that includes it. This is the underlying idea behind the "systemic capacity" analysis of function, which we will defend in this chapter. According to Cummins, a component may have a function even if the component was not "designed", therefore parts with no selection history can be ascribed a function.

Now if psychiatry is a branch of medicine, its function concepts should be continuous with those of medicine and physiology more generally. In this section we argue that the way the relevant sciences determine function does not presume a selectionist concept of function. We will argue that the relevant accounts of function are those which find a home in mechanistic explanation (the systemic capacity view) rather than historical explanation (the selectionist view). The life sciences ask all sorts of questions, but the questions which medicine asks are not those which a selectionist account of function can answer.

8.3.1 The selectionist view

Evolutionary views of function involve causal-historical explanations of traits that we will call *selectionist*. Wakefield's argument is that we look to science for our best theory of how human traits come about, and science says that functions are fixed by the historical process of evolution by natural selection. The heart is a standard example. Millikan (1984) said that the heart is a pump because it is the heart's pumping that causally contributes to the successful reproduction of organisms with hearts: if x is a member of a biological category it is not because of "the actual constitution, powers, or dispositions" of x , but because of the "proper function" of x (Millikan 1984, p.17). X 's proper function depends on the history of x 's lineage, which explains x 's being supposed to do whatever it does. The point is quite subtle because the relevant history consists of correlations

obtained between ancestors of x having a certain character and their having been able to perform x 's function. So the structure of a heart explains why it pumps, but it does not count as a heart in virtue of having that structure.

The selectionist account of function seems to offer two big benefits. First, it promises to give a definite specification of the function of an organic system and hence a clear criterion for calling it dysfunctional. Second, it seems to offer a scientifically unproblematic way to say what a system ought to be like. If you are worried about the accusation that function talk is normative you can embrace natural selection. Teleological notions are commonly associated with the pre-Darwinian view that the biological realm provides evidence of conscious design by a supernatural creator. The point about evolutionary views is that they assuage this metaphysical concern by showing how norms are part of nature. We are not going to suggest that there is something wrong with the Darwinian picture of natural order. But we will suggest that an evolutionary concept of function and dysfunction is a poor bet for psychiatry.

We begin with a question about inquiry. On the face of it, the evolutionary view confronts a simple epistemic problem. How do we ascertain whether the relevant history is actually obtained? And even if a certain history is obtained, what is the proper assignment of evolutionary function and dysfunction? We are not saying that it is impossible to test claims that something is an adaptation, but examples of successful tests, we claim, look nothing like the tests we see in medicine or the parts of biology adjacent to medicine.

Cain and Sheppard (1950) provide an example of a test of a selectionist function. They demonstrated that the appearance of the shells of the grove snail (*Cepaea nemoralis*), which had been believed to be a neutral trait ascribed to random factors, actually varied adaptively. For example, darker and more uniform shell patterns predominated in darker and more uniform habitats like deep woodland. Populations of snails that were more heavily predated by song thrushes had a higher incidence of more conspicuous patterns. Cain and Sheppard persuaded everybody that *C. nemoralis* patterns were adaptive camouflage, and therefore one can appeal (defeasibly) to current utility to fix function. One can also measure reproductive success over several generations to see if the variant one believes to be adaptive is in fact spreading in the population in response to selection pressures. This has been done, for example, by Peter and Rosemary Grant, who spent more than 30 years tracking the responses to selection pressure of Darwin's finches (Grant 2000).

It is even more difficult to test an ascription of evolutionary function in humans. Remember, we need an account of function that will license judgments that some system is dysfunctional. One analysis would be that traits that are selected against (are becoming less prevalent) in the face of the success of other

variants are dysfunctional. To do this, we need to show that humans with that trait are reproductively less successful. That would take generations, and no branch of medicine or physiology tests claims about the functional architecture of humans by showing how possession of a trait correlates with reproductive success. The sorts of tests that would be needed to demonstrate that are almost impossible to do on humans, given the difficulty of establishing over several generations that a part of our biology does in fact enhance reproductive success relative to the competition. Of course an unhealthy heart can kill you, but we do not decide that your heart is healthy by looking at the number of offspring you have. The situation is even worse for mental illness. We have little systematic evidence on whether alcoholics or depressives or psychopaths have fewer children than control subjects. (Psychopaths, for instance, tend to have irregular, promiscuous sexual lives, but that doesn't mean that they don't reproduce successfully.)

At this point, somebody with a commitment to the evolutionary picture may object that in fact the relevant test is that of current fitness. We can work out what a trait does for an organism in the current environment and thereby show that it is adaptive. However, there are problems with this strategy too. Lloyd (2005, p.166–7) notes that to assume, based on current evidence, that mammalian fur has always functioned to help with thermoregulation we have to also buy into the claim that selection pressures have remained constant in the past. To do this we really need independent evidence from other historical sciences, and the relevant assumption about the stability of past pressures is in any case very hard to make in human biology. There seem to have been substantial changes in our environment and our relations to it, even over small time scales, like the last few thousand years.

But don't medicine and physiology routinely try to work out what a system contributes to the overall functioning of the organism? Yes, but that doesn't mean that in doing so we are trying to establish that a biological component has a selectionist function. For example, take Hubel and Wiesel's famous program of mapping the receptive fields of cells in the visual cortex and then establishing further visual information-processing channels in the brain. That program, and the research on the neurobiology of vision inspired by it, depended on a set of engineering assumptions about the way the brain is organized to process information. It did not test assumptions about the selective advantage and history of the components of the visual brain. Most physiological research is based on establishing the components, and the functional relations between components, in biological systems. It is not aimed at uncovering evolutionary relationships. It may be that the facts uncovered in physiology are evidence for evolutionary relationships, and of course all biological systems have an evolutionary history, but when we determine what normal

function is, in medicine, we do not even try to establish what something's selectionist function is.

Schaffner (1993) argued that although medicine might use teleological talk in its attempts to develop mechanistic explanations, that talk is just heuristic. It focuses our attention on entities that are useful to the organism. Schaffner suggested that as we learn more about the role a structure plays in the overall functioning of an organism, the need for functional ascriptions drops out. It is replaced by the vocabulary of mechanistic explanation: the causal relationship of parts that jointly produce phenomena of explanatory interest. Functional explanations that draw on evolutionary considerations are, he claimed "necessary, but empirically weak to the point of becoming almost metaphysical" (Schaffner 1993, p. 389–90).

In our view, biomedical ascription of function to a system makes no claims about adaptedness or selective history. It requires only that we can identify the role played by a system in the overall economy of the organism. How is dysfunction determined? By the use of a biomedical concept of normality that is an idealized description of a component of a biological system in an unperturbed state that may never be attained in actual systems. It does not rest on the failure of a biological part to function as its ancestors did, but by its failure to be close enough to the causal contribution of the analogous part in the idealized overall system.

Wachbroit (1994, p. 588) argues persuasively that when medicine or physiology says that an organ is "normal", the relevant conception of normality "is similar to the role pure states or ideal entities play in physical theories". Such an idealization represents actual organs or systems in unperturbed states (see also Ereshefsky 2009). To understand a real case we add information to develop a model that resembles actual hearts (Wachbroit 1994, p. 589). For instance, Gross (1921) was able to establish post mortem that anastomotic communication between main arteries increases over a typical lifespan, thereby establishing that we need to model younger and older hearts differently. The point of such idealizations is not to represent the statistically average heart, but to describe hearts in a way that allows departures from the ideal to be recognized and to serve as template from which more realistic models can be built. In general, physiological theories are families of such idealizations, and bodily systems are understood as functional parts of larger systems, typed unhistorically. In so far as psychiatry is a branch of medicine, the concept of function it needs will resemble the unhistorical concepts of physiology and bio-medicine. Evolutionary considerations are just beside the point.¹

¹ Furthermore, a trait's evolutionary function (and hence its dysfunction) might be quite different from what psychiatrists think of its function and dysfunction.

Here is one last argument against reading the selectionist view of function into psychiatry. Suppose we discovered that schizophrenia, bipolar disorder, depression, and psychosis were evolutionary adaptive strategies: forms of behavior evolved to further the interests of the sufferer. (Theories of this sort are reviewed by Murphy 2005.) Would we then re-think their status as mental disorders or would we be led to reject the evolutionary dysfunction view as providing an adequate account of dysfunction in psychiatry?

Horwitz and Wakefield (2007) argue that the human mind includes a system that has evolved by natural selection to respond to loss. They argue that the intuitive distinction between normal sadness and morbid depression tracks the workings of their hypothetical system, which explains why we become sad in situations where sadness seems like the right response. Major depression occurs when the system kicks in for no reason or produces excessive responses to trivial misfortunes, whereas normal sadness occurs in response to serious misfortune, for the system they hypothesize is "biologically designed to produce such responses at appropriate times" (Horwitz and Wakefield 2007, p. 25). They suppose that there is a system misfiring in cases of major depression because they are committed to an evolutionary view of function that underpins their two-stage picture. Their conjecture, accordingly, is that since depression is obviously a matter of the mind going wrong, it must be a matter of some system failing to function as natural selection has designed it. It follows that if whatever underlies depression is not an adaptation failing to function as designed, they would have to say that depression is not a mental disorder.

It is not impossible that we should be led to conclude that what we had seen as a mental illness is actually not one, and there are historical examples of that happening. (Psychiatrists famously changed their minds about the pathological nature of homosexuality in the 1970s.) However, overturning our judgments about depression would be a real revolution. Depression and psychosis are the two paradigmatic kinds of mental disorder. A theory which said, based on evolutionary considerations, that depressed people or psychotics are really healthy, would almost certainly be rejected. That is a conceptual revisionism that we would be very unlikely to accept. It would be like finding out that cancer is not a disease. Nothing is settled for ever in science, but it is very hard to imagine us siding with the revisionist, rather than simply saying that they had uncovered some interesting causal explanations of the nature of a disease. Or, if we did accept it, it is unclear whether the concept of mental disorder would survive, since its very utility would be called into question by the removal of its exemplary instances. The facts might also go the other way and point to some condition being a mental illness even though nobody had ever thought it was. An apparently normal or beneficial part of our everyday

psychology might reflect a failure of an ancient adaptation to do its job: suppose that giving to charity is caused by the failure of an adaptation designed to make us altruistic towards members of our immediate communities. In that case, charitable giving would be a candidate mental illness. Although commonsense is not definitive, we would need very good grounds indeed to accept an analysis that raises the possibility that giving to charity is pathological and paranoid delusions are not, and we do not think that the grounds to accept such revisionism are present. Actual scientific explanations of disease, psychiatric or not, are accepted routinely in the absence of evolutionary considerations. Suppose that a good explanation of the underlying neurobiology of depression were discovered. Psychiatrists would dislocate their shoulders from all the patting themselves on the back that would ensue in the light of what would be seen as a scientific triumph, and nobody would care whether there was an evolutionary rationale, or testable Darwinian hypothesis, for the system that was uncovered. Testing in biomedicine does not work that way, as we have tried to show, and our ordinary intuitions do not demand an evolutionary rationale either.

8.3.2 The systemic capacity view

Cummins (1975) claimed that when we say that the function of the heartbeat is to circulate blood through the organism we are not accounting for adaptation but *explaining circulation*. We begin by assuming that the circulatory system is what needs to be explained, and we identify a system as having a function in the context of explaining it. We may explain the *advantage* of the heartbeat by identifying the activity it facilitates. This is different from explaining the existence of the heartbeat. Cummins (1975, p. 746) thought that we could only say why a system exists by appealing to the intentions of a designer. We suppose that it can often be done for organisms by appealing to evolutionary history. But we agree with Cummins that explaining what a systemic component presently does is different from explaining why it is there. The life sciences ask lots of questions, and trim their accounts of function to fit.

Cummins argued that the basic explanatory use of function talk in the life sciences derives from a particular analytic strategy in which the biologically significant capacities of a whole organism are explained by breaking down the organism's biology into a number of "systems", the circulatory system, the digestive system, the nervous system, and so on, each of which has its characteristic capacities. These capacities are in turn analyzed into the capacities of their component organs and structures. We can reiterate the systemic capacity framework down through levels of physiology, explaining the workings of the circulatory system, the heart, certain kinds of tissue, certain kinds of cell, and so on. Much mechanistic research in biology exemplifies this approach.

Many theorists have argued that there are biological contexts in which Cummins's ahistorical analysis is broadly correct (Godfrey-Smith 1998; Kitcher 1998; Craver 2007; Bechtel 2008). *Function* in biochemistry and physiology typically refers to the contribution a structure makes to the overall organismic system containing it and therefore differs from functional talk in evolutionary disciplines. As Kitcher (1998, p. 266) already noted, "philosophical analyses reveal unresolved ambiguities in biological practice", and philosophical analysis should respect, rather than try to reform, the differing scientific usages. The view of function within physiology is, in outline, the conception of systemic functional analysis we introduced above and aim to defend as the right characterization of function talk in medicine, and hence in psychiatry.

The systemic capacity conception of function uses an analytic strategy to explain how something is able to perform a function by treating functions as dispositions of a component of a larger system. On the systemic capacity account functions are assigned to components in virtue of the role that they play in the production of a phenomenon, typically the output in some greater system. Once one has the relevant system then one proceeds to analyze the system into components and assign functions to the components in virtue of the role they play with respect to the production of the phenomena that one wants to explain.

One longstanding objection to the use of a systemic capacity account of function is that it robs us of the power to say when a system is dysfunctional. This is held to be a problem because assignments of systemic capacity function look to be relative to the interests of the researcher rather than a feature of the world, as the two-stage view requires. Davies (2001) is one recent theorist who reviews this objection. He claims that assignment of function to components is doubly relative. First, which components are relevant is going to partly depend on what phenomena the researcher is interested in. So, if we are mostly interested in the noises that the body makes, it is the sound of the heartbeat which is the relevant property of the heart. This makes functional analysis mind-dependent. There is no such thing as the natural functions of a system, just whatever the investigator finds interesting. Cummins (1975, p. 763–4) saw this objection coming. He argued there would be little point in applying a functional analysis because the disposition of the body to make a noise is just a process of the same type as the disposition of the heart to make a noise. Functional talk would not be interesting in such cases. But this is a rather pragmatic, rather than principled, reason for ruling out the sound the body makes as the disposition of interest. Davies argues that a better defense can be made by restricting functional ascriptions to hierarchically organized systems in

which lower level capacities realize upper level ones. The noise of a heart is part of the overall noise of the body, but it is at the same level, so it does not count as a functional subpart of the noise system. That gives us a characterization of function independent of our explanatory interests.

Cummins and Davies have both maintained that systems (and remember, a system from one perspective may be a subsystem of a more encompassing unit) must consist of two distinct levels. There is the level of the phenomena and the system that produces them and there is the lower level of the components and their functions. Once we hit a level at which the outputs are basic, where the "system" cannot be analyzed into further components, then we have reached the end of the systemic capacity chain of explanation and moved into a different science, which analyses the physical composition of the system. But where does it peak? Is there a principled topmost level of analysis such that all components at lower levels are functioning to realize that topmost disposition? The two-stage view seems to require that major organ systems can be assessed as functional and as making a contribution to overall health, independently of what anyone thinks. If mind-independence is established via participation in a hierarchy of natural levels, we want to know whether a human being is a level in the hierarchy. If the overall person is a level in the relevant hierarchy, we have a level of analysis of which the major organ systems are the components.

Our current question, then, is whether there is a level at which each individual counts as a functional system with the major sub-systems—respiration, circulation, and so on—as the proper parts of that system. For the two-stage view to work there needs to be a mind-independent level of analysis corresponding to the individual human. This is necessary because if there is not such a mind-independent whole-organism level we face the danger that whatever aspects of overall human flourishing we attend to in psychiatry (or elsewhere in medicine) will just be assessed relative to the interests of the investigator. If that is the case, then the evaluation of one's social cognition or self-esteem or mood cannot be conducted in terms of its principled contribution to an overall level of function, but just in terms of the particular output of interest. Bluntly, you could show how thoughts of suicide play a role in explaining self-destructive behavior if you are interested in self-destructive behavior. However, the two-stage view is interested in establishing not just that self-destructive behavior exists and is contrary to our view of a good human life, but in showing how it rests on dysfunction. So for the two-stage picture to apply there must be a natural functional level of the human person, with a set of outputs whose character is mind-INdependent. If this cannot be specified, then we can say that biological subsystems make a contribution to overall functioning, but not in the mind-independent way that the two-stage picture

needs. We will be left with an account of overall human function that is a reflection of our norms, and identify dysfunction in that light. And that is just the situation that the two-stage view is designed to keep us out of.

The problem we face is that of trying to establish what the major organ systems of the body (and, for psychiatry, the major components of our psychology) are actually for. Boorse's answer (1976b) is that they are there to help us survive and reproduce. Griffiths (2006, p. 2) agrees: the "causal functions which are of primary interest to biologists are those which contribute to an organism's capacity to survive and reproduce." We are not convinced this is right. Boorse offers no real argument for his point; he just asserts that in physiology the relevant system for functional analysis is the organism and the relevant goals are those of its survival and reproduction (1976b, p. 84). Griffiths (2006, p. 2) does have an argument. He says biology makes use of a causal conception of function to identify "adaptive traits, which increase the fitness of organisms that possess them relative to other types." Adaptive traits may not yet be adaptations (i.e., we may not explain them as products of evolution by natural selection) but they are adaptive in so far as they explain survival value; the functional analysis explains the survival value of a particular piece of behavior or physiological process. However, there are other questions about the role that biological systems play. As well as questions of survival value, we can ask questions that simply aim to find out how a system does what it does in the context of the superordinate system.

There is an idea almost as old as natural selection that may help to answer these questions. It is Claude Bernard's 1927 [1865] suggestion that major systems in the human body seek to maintain stable internal homeostatic states, in his favorite example by regulating the chemistry of the blood. Bernard argued that organisms can only explore and transform the external environment if they have sufficient internal stability. In this view the answer to the question "What is the function of the major physiological systems?" is "To keep the internal environment stable." We suggest that homeostasis, not survival value, is what guides physiological answers to questions about causal explanations of biological systems.

Physiological mechanisms are mutually integrated into an organism-level system that depends on the external environment but is not integrated into it the same way. The overall physiological system can move through different environments, for instance, in a way that its components cannot; they must remain *in situ* to do their job. Similarly, what goes on within the cell is integrated relative to the relationship between the cell and its surrounds. At every level of the hierarchy the internal environment is relatively independent of the external one, and more fully integrated internally than the higher level is.

integrated with its surrounds. If mechanisms co-operate to maintain a constant internal environment, the system can compensate for changes in the external environment.

This relative uncoupling of the overall system from its range of possible environments is well defended for psychology by Rupert (2009); in psychiatric contexts we can follow Rupert in treating the organism as a relatively enduring cognitive system. The specific problem for psychiatry is to identify failures of cognitive parts that constitute the mechanism's underlying intelligent agency.

We have argued that the problem of relativity in the systemic capacity view of function can be solved in medicine and physiology by appeal to a hierarchy of integrated systems. A further, distinct, problem concerns not the relativity but the indeterminacy of functional ascriptions. Millikan (2002, section 6) points out that we might ascribe a broader or less broad function to a trait; after all, if my ear canals can keep me upright in a gravitational field of 1G they can also work under forces of 0.9G, 0.8G and so on. If we regard these as different functions then we have an infinity of them. She proposes to solve this problem by introducing "the descriptive generality requirement", which states that functions (of whatever type) should be described according to the most general principles available. We agree, and pause only to note that the principles will typically derive from the even broader principles guiding the idealization of the system.

So we have a way to explain function, this time in terms of the contribution a system makes to the operation of a wider, integrated set of systems that maintain themselves in equilibrium. This view has a lot in common with the causal conceptions of function that philosophers have worked within since Cummins. It detaches function from natural selection and sees it working in ecological and physiological time rather than evolutionary time, to make contributions to the maintenance of a living system that are distinct from questions of survival value and reproductive potency. In the next section we will see whether the systemic capacity view of function can also account for dysfunction, and then in the final section we will discuss whether this conception of function licenses the attribution of mind-independent norms to systems of psychiatric interest.

8.4 Dysfunction

Something clearly needs to be said about dysfunction. Any theory must be able to meet the problem of distinguishing the functional from the dysfunctional. Some theorists might not insist that a theory of function has the resources to account for dysfunction, but it is clear that any account of function that purports to be relevant for general medicine or psychiatry has to meet the

dysfunction challenge. In this section, we will evaluate both the selectionist and systemic capacity views on dysfunction. If either view can cope with the challenge and the competing view cannot, we have a powerful argument in favor of the successful view. To do this, we will discuss Davies' argument against both selectionist and systemic accounts. We will argue that this argument fails, in both views, mainly because Davies cannot substantiate his key contention that the classification of entities in the life sciences is essentially functional.

Supporters of an evolutionary account of function (e.g., Neander 2002) often think that one of the virtues of their theory is the straightforward way in which an account of dysfunction (or malfunction) follows from the account. Their idea is that we can say when a system is malfunctioning by observing that it is not carrying out the job which natural selection designed it to perform.

In contrast, it is widely bruited that systemic accounts of function cannot deal with dysfunction at all. If the function of a system is relative to our explanatory interests, it seems that a putative malfunction can just be understood as a contribution to a different overall disposition of the system. We endorse Godfrey-Smith's (1993) response to this objection: a token component in a system is malfunctioning when it cannot play the role that lets other tokens of the same type feature in the explanation of the larger system.

Davies (2001, p. 212) denies this. He says that Godfrey-Smith's point works only if "incapacitated functions' tokens retain their membership in the functional types and not just the generic type" and he denies that there are any grounds for thinking that is true. That is, a dysfunctional heart just belongs to the generic type "circulatory device" because functional types are defined in terms of what they can do. Therefore, if a component cannot carry out its normal contribution to the overall system then it ceases to be a member of the type.

However, Davies also thinks that roughly the same argument works to deny that the selectionist can explain dysfunction. Evolutionary accounts, too, he argues, individuate types according to their functions. Since the functions are thought to be necessary and sufficient for membership in the type, it is thus impossible for an instance to both be a member of the type (possess the necessary and sufficient condition or function) and yet lack the function and hence dysfunction. So Davies (p. 210) maintains that instead of saying that a heart is malfunctioning, all the selectionist view lets us say is that the instance that does not pump is not a heart after all. It therefore doesn't have the function of pumping and thus it isn't malfunctioning—it just lacks the function that we wanted to assign to it. Dysfunctional hearts lack the defining capacity of historical success, and hence do not belong to the type at all. Davies' claim seems

correct to the extent that if having some function F really is both necessary and sufficient for F's being classified as a member of the functional kind K, then it follows that if F were to lack the necessary and sufficient condition for being a member of kind K then it would simply stop being a member rather than becoming a dysfunctioning member. By analogy, if we consider an instance of gold and we then apply a proton gun and remove one of the protons then the instance isn't a malfunctioning or abnormal instance of gold in virtue of having one fewer proton. Rather, it would no longer be an instance of gold.

Davies' argument, if correct, shows that no naturalistic concept of function can accommodate malfunction. He therefore concludes that there are no norms of performance in nature that a naturalist can embrace (p. 214). This seems to mean that medicine is built on purely nonnaturalistic or normative assumptions that people make about whether other people are flourishing. If that is right then it seems that the two-stage view is indefensible not just in psychiatry but in all of medicine.

However, Davies' argument can be refuted. It relies, as we just saw, on the assumption that both the selectionist and the systemic capacity theorist individuate biological types according to the function that they assign to the type. Insofar as types possess their function as a matter of necessity he seems correct that an instance of a type cannot malfunction. Davies maintains that the burden of proof is on theorists who think there can be norms of nature to provide an account of how to individuate kinds.

However, we follow the biological consensus and deny that biological components are essentially typed according to their function—a counterargument that seems to work for both the evolutionary view and the systemic capacity view. The point that traits are not typed by their function was first made by Amundson and Lauder (1994), and it has been endorsed widely for the best of reasons, that is, it appears to capture biological practice. So Griffiths (1997, p. 215–6), for instance, argues that if biology has kinds at all, they are either cladistic kinds (members of a shared lineage) or disjunctions thereof: it is common descent rather than function that determines that a heart is a heart. More recently, Griffiths (2006, section 6) has argued that homology determines that a heart is a heart. He argues that functional classifications in biology type organisms by analogy (or shared evolutionary purpose) and are logically dependent on classifications in terms of homology. An organ can still be a token of a type defined in terms of homology even if it is not currently functioning.

So Davies' objection appears to fail for contexts supplied by evolutionary questions, since it seems untrue to insist that traits are typed according to their function in biology generally. We will now argue that it is also untrue for

systemic capacity accounts. Griffiths (2006) has argued that many disciplines in experimental biology type traits by homology. In medicine there is a long tradition of identifying systems anatomically before going on to investigate their physiology. In the neurosciences, for instance, which are plainly relevant to psychiatry, there is a century-old tradition of identifying brain regions (Brodmann's areas) based on the physical architecture of cells within the region and then going on to ask what the function of that region might be. We appeal to this kind of procedure to defend our strategy of arguing that organs have functions, rather than insisting that they are essentially tokens of functional types. We will not investigate the precise relationship between this tradition and the tradition of typing by homology, which also involves identifying biological components in terms of their anatomical structure and position and relationships to other organs, and not just solely as functional types. We can therefore say that a heart, even if it has lost the pumping capacity that hearts often have and hence is malfunctioning, retains its identity as a heart because it is still in the position characteristic of hearts and it retains some of the musculature and internal anatomy of a heart. Or consider a doctor conducting an autopsy. She can identify bodily organs as tokens of organ types in order to assess the degree of pathology in each case. Reasoning like this is perfectly satisfactory scientific practice: we mentioned earlier how Louis Gross' research in cardiology depended on generalizing from post-mortem hearts to living ones. However, every system in a corpse no longer possesses its function in the sense that Davies's argument relies on. In summary, sciences type components of biological systems in nonfunctional ways. Its function is a property of a biological unit, not its essence. Thus it is that Davies' challenge can be met for both evolutionary and systemic accounts.

8.5 Dysfunction and the role of science

We now turn to ask more general questions about the role of science in discovering disorder. The two-stage view says that scientific facts play a significant role in determining whether or not a condition is a disorder. The dysfunction criterion was initially introduced to help us determine which individuals were in fact disordered, in a way that avoids subjective, mind-dependent, or culturally relative judgments. For the two-stage view to work, the science of mental disorder (and, indeed, that of disease more generally) can't be just a particular application of a nonnormative neuroscience or molecular biology, but a distinct province of that wider science, one concerned with dysfunctions rather than just unusual cognitive or physiological processes. We will end this chapter by asking whether the two-stage view can be sustained with respect to the

systemic account of function as we have sketched it, and whether it makes any difference if it can't be.

The scientific aspect of the two-stage view thus has the job of rebutting the skeptical claim that disorders are just violations of norms that currently prevail in a society. Because it must play this role in the two-stage picture, science must go beyond the role of simply determining what kinds of conditions there are, how they develop, and what interventions are effective for them. We need a definitive list of dysfunctions that justify our regarding a condition to be a disorder, not just knowledge of how the mind and body work in physiological contexts. In the remainder of this section we sketch a defense of a more modest role for science in a program that is basically normative.

Cooper (2005, 2007) and Murphy (2006) have drawn an analogy between the concept of mental disorder and that of weed. Weeds are not a scientifically relevant category of entities. We can perhaps say that a weed is a fast-growing species that negatively impacts on economically valuable crops, usually through competition for nutrients, sunlight, and space. What fixes the extension of "weed" (and similar concepts like "vermin" or "precious metal") is a set of contingent human interests that can change over time. There is nothing inherently dysfunctional about a weed; weeds are just species that we don't like because of certain interests that we have. Suppose that determining that a condition is a disorder is like determining that a plant is a weed. The judgment is determined by value judgments we have already made. So "weed" is not a technical term in ecology and the science of weeds is just the science of plants, put to special use. "Weed" rarely appears within publications in reputable ecological journals, but nonetheless there is real, explanatory mind-independent knowledge to be had about each sort of "weed". For those who are skeptical about the two-stage view, science does not uncover dysfunction in a way that is independent of our value judgments; science is directed by those value judgments.

We will consider cancer as an example. It is an obvious instance of a biomedical disorder. This example has been chosen to emphasize that our skeptic's questions aren't specific to psychiatric disorders. Theorists have often thought that the presence of normativity in psychiatry would undermine psychiatry's status as a branch of medicine, but we think that psychiatry is useful in helping us become clearer on the role and limits of science for medicine more generally.

Let's suppose that one wants to understand or explain cell development roughly along the lines of the systemic view. One way to do this is to construct an idealization of the development of cells of a given type. One kind of information that we would want our model to contain would be the causal

information inferred when we make interventions that seem to have a fairly robust bearing on the future course of the cell. An alternative would be to attempt to model subtypes of cell development. There are a number of considerations that bear on whether we should "lump" or "split" phenomena in these ways. One consideration is whether we discover differences in the response to our manipulations that result in differential outcomes that seem important. The occurrence of fairly robust responses under intervention is thus one important consideration for our individuating importantly different kinds of phenomena. Medicine has one further refinement of this that we shall consider shortly.

Suppose we want to understand cancerous cell development. We have two fairly different ways of proceeding. One way is to initially proceed as before. We build a model of cell development in general or a particular kind of cell development. We can then proceed to model cancerous cell development by explaining what "break downs" occur in our model in order to explain cancer as a "biological malfunction" of the cell in the systemic sense. An alternative would model cancerous cell development on its own terms, in much the same way as we initially modeled the development of the noncancerous cell. Our model would be constructed on the basis of some idealization of the development of particular cancerous cells. Now it seems that different groups of scientists could proceed differently on this and we could well end up with two distinct models of cancer. According to the first model cancer would be a "biological dysfunction" whereas according to the second model cancer would be a distinctive pathway that cells can take. Of course, everyone thinks cancers are pathologies, but our skeptic is asking how we establish that by scientific investigation of cell development, as opposed to merely using our prior assumptions about what seems intuitively pathological and plugging a causal model into it.

Both models seem capable of capturing precisely the same causal information with respect to providing different points at which we can intervene to disrupt the process we have modeled. We can disrupt the course of cell development and we can disrupt the course of cancerous cell development. But surely cancer can't be both a biological dysfunction and a merely unusual kind of biological development at the same time! Given the set up we have imagined, what further scientifically discoverable fact is there that tells us whether or not cancer involves a dysfunction? How much is biological dysfunction an assumption of our modeling rather than something that is to be discovered by it?

To save the two-stage view, we need to answer the question we just asked: what fact is there that science can discover that discriminates between cancer as a dysfunction and cancer as an unusual developmental pathway? If there is

no such fact then we must reconsider whether science is playing a foundational role in determining that conditions are disorders, as the two-stage view says. The skeptic's alternative is that science discovers important biological facts guided by prior normative judgments that something is a disorder.

To save the two-stage view, we must uncover natural norms.

To reply to the skeptic we cannot stay at the level of the biological system, but we must move up to ask about the role of the system in the overall economy of the organism. The answer from the two-stage theorist who adopts our version of the systemic capacity is: look at what the system you are studying does for the organism. The reason why cancer is a dysfunction is that it drives the organism out of equilibrium and into a new state in which other systems stop being able to act as we usually explain them. This approach also requires a way of differentiating normal from abnormal development; basically, it defines normal development as the set of pathways that lead to the final, functional, adult form.

The systemic capacity theorist can use the idea of a natural hierarchy in the organism to defend the claim that disease perverts the functioning that is normal for an organ system. The textbook tells you what a healthy organ is like by reference to an abstraction—an idealized organ. This concept of normality is not justified by conceptual analysis. It draws its authority from its predictive and explanatory utility: we account for variation in actual hearts (a particular rhythm, say), by citing the textbook rhythmic pattern (which may be very unusual statistically) and identifying other patterns as arrhythmic. The role of the idealizations, as we have said, is to classify real systems according to their departure from the ideal, and the ideal must be justified by an appeal to overall organismic homeostasis.

Our skeptic just says that now the problem recurs. What justifies our idealized or assumed “normal” systems? Variation in biological traits is ubiquitous, so establishing whether or not a mechanism is functioning normally depends on whether an overall picture of normality for the organism can be adumbrated in a way that doesn't depend on our prior values. The skeptic just denies that can be done. The exponent of the two-stage view will say that it is possible.

We might think that disorders can be tied to a break between normal and abnormal functioning of an underlying mechanism, such as a failure of the kidneys to conserve electrolytes. Skeptics argue that while one way of construing the phenomenon is that the kidneys “fail to conserve”, another is that they simply “don't conserve”; conserving electrolytes is not part of the model of what those kidneys are doing. The problem seems to recur at each level on the systemic analysis. Adding layers up (e.g., considering individuals as functional

or homeostatic units in a social group) or down (the organ systems that comprise them) will not determine how we idealize or assume "normal" or "homeostatic" systems to be in a way that is independent of our values. No biological system can sustain a stable internal environment if its system for filtering waste has broken down. We may think that in that case we have a clear rationale for arguing that when a kidney does not filter waste then it is "failing". We are dealing with a "problem" for the overall system, not merely an alternative pathway for a component. But the skeptic can reply that lying behind this intuition is our (entirely reasonable) valuation of organismic integrity. It is in virtue of our valuing it that we are inclined to describe many component processes that threaten to disrupt it as "dysfunctional".

Distinguishing failures to flourish from functional systemic failures will always be a hard problem for psychiatry. For example, judgments of irrationality are central to many psychiatric diagnoses, and our standards of rational thought are not based on biological findings. They reflect standards derived from normative assessment (Murphy 2006, chapter 5). The possibility of psychiatric explanation employing the methods and models of physical medicine, then, depends on how much of our psychology is like the visual system, that is, decomposable into structures with a clear natural function that can be tied in to a biological hierarchy topping out in homeostasis. Some mental processes may lend themselves to such a treatment, but it is unclear how the notion of homeostasis even applies to most of our rational and emotional lives. We may also wonder whether the notion of homeostasis itself can be rendered nonnormatively. Accordingly, our skeptic's challenge is likely to be very hard for psychiatry to overcome. If this can't be done then the same problem arises in medicine more generally. Fortunately the skeptical position doesn't undermine the considerable role that science can play with respect to modeling traits of interest, individuating kinds of conditions, discovering their etiology and course, and developing more or less effective interventions for them. However, the skeptical view does threaten the two-stage view by arguing that the science is guided by judgments of disorder, instead of providing a foundation for them. The challenge to the naturalist is to establish, in medicine generally, not just psychiatry, the mind-independence of the functional hierarchy and the relevant notion of organism-level performance.

8.6 Conclusion

Since the rise of a mechanistic conception of nature in the seventeenth century, medicine, including psychiatry, has struggled to make sense of the apparent teleology of biological systems. Without a satisfactory account of function

it is hard to see how we can have a satisfactory account of malfunction, which endangers any naturalistic perspective on disease. Many theorists have considered evolved, selected functions to be an attractive solution to this problem. In this chapter, we have argued against Darwinian accounts of function, as they do a poor job of accounting for medical practice and suffer from debilitating epistemic problems. In our view, psychiatry and medicine presume a systemic capacity view of function, in which functions are considered as components of mechanisms designed to keep a system in homeostasis.

Acknowledgments

We thank Paul Griffiths, Jack Justus, and a discerning anonymous referee for their comments.

References

- Amundson, R. and Lauder, G.V. (1994) Function without purpose: the uses of causal role function in evolutionary biology. *Biology and Philosophy*, 9, 443–69.
- Andreasen, N.C. (2001) *Brave New Brain*. Oxford University Press, New York.
- Bechtel, W. (2008) *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. Routledge, London.
- Bernard, C. 1927 [1865] *Introduction to the Study of Experimental Medicine*, Macmillan, New York.
- Boorse, C. (1975) On the distinction between disease and illness. *Philosophy and Public Affairs*, 5 (1): 49–68.
- Boorse, C. (1976a) What a theory of mental health should be. *Journal for the Theory of Social Behavior*, 6, 61–84.
- Boorse, C. (1976b) Wright on functions. *Philosophical Review*, 85 (1), 70–86.
- Cain, A.J. and Sheppard, P.M. (1950) Selection in the polymorphic land snail *Cepaea nemoralis*. *Heredity*, 4, 275–94.
- Cooper, R. (2005) *Classifying Madness: A Philosophical Examination of the diagnostic and statistical manual of mental disorders*. Springer, Dordrecht.
- Cooper, R. (2007) *Psychiatry and the Philosophy of Science*. Acumen, Stocksfield.
- Craver, C.F. (2007) *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford University Press, New York.
- Cummins, R. (1975) Functional analysis. *Journal of Philosophy*, 72, 741–64.
- Davies, P.S. (2001) *Norms of Nature*. MIT Press, Cambridge, MA.
- De Block, A. (2008) Why mental disorders are mental dysfunctions (and nothing more). Some Darwinian arguments. *Studies in the History and Philosophy of Science (Part C)*, 39, 338–46.
- Ereshefsky, M. (2009) Defining “health” and “disease”. *Studies in the History and Philosophy of Biology and Biomedical Sciences*, 40, 221–7.
- Godfrey-Smith, P. (1993) Functions: consensus without unity. *Pacific Philosophical Quarterly*, 74, 196–208.

- Godfrey-Smith, P. (1998) Functions: consensus without unity. In D.L. Hull and M. Ruse (eds), *The Philosophy of Biology*. Oxford University Press, Oxford, pp. 280–92.
- Grant, P.R. (2000) *Ecology and Evolution of Darwin's Finches*. Princeton University Press, Princeton, NJ.
- Griffiths, P.E. (1997) *What Emotions Really Are*. University of Chicago Press, Chicago.
- Griffiths, P.E. (2006) Function, homology, and character individuation. *Philosophy of Science*, 73 (1): 1–25.
- Gross, L. (1921) *The Blood Supply to the Heart in its Anatomical and Clinical Aspects*. Hoeber, New York.
- Horwitz, A. and Wakefield, J. (2007) *The Loss of Sadness: How Psychiatry Transformed Normal Sadness into Depressive Disorder*. Oxford University Press, New York.
- Kitcher, P. (1998) Function and design. In D.L. Hull and M. Ruse (ed.), *The Philosophy of Biology*. Oxford University Press, Oxford, pp. 258–79.
- Lloyd, E.A. (2005) *The Case of the Female Orgasm: Bias in the Science of Evolution*. Harvard University Press, Cambridge, MA.
- Millikan, R.G. (1984) *Language, Thought and Other Biological Categories*. MIT Press, Cambridge, MA.
- Millikan, R.G. (2002) Biofunctions: two paradigms. In A. Ariew, R. Cummins, and M. Perlman (eds), *Functions*. Oxford University Press, Oxford, pp. 113–43.
- Murphy, D. (2005) Can evolution explain insanity? *Biology and Philosophy*, 20, 745–66.
- Murphy, D. (2006) *Psychiatry in the Scientific Image*. MIT Press, Cambridge, MA.
- Murphy, D. (2009) Psychiatry and the concept of disease as pathology. In M. Broome and L. Bortolotti (eds), *Psychiatry as Cognitive Neuroscience: Philosophical Perspectives*. Oxford University Press, New York, pp. 103–117.
- Neander, K. (1991) Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science*, 58, 168–84.
- Neander, K. (2002) Types of traits: Function, structure and homology in the classification of traits. In A. Ariew, R. Cummins, and M. Perlman (eds), *Functions: New Essays in the Philosophy of Biology and Psychology*. Oxford University Press, New York.
- Rupert, R. (2009) *Cognitive Systems and the Extended Mind*. Oxford University Press, New York.
- Schaffner, K.F. (1993) *Discovery and Explanation in Biology and Medicine*. University of Chicago Press, Chicago.
- Wachbroit, R. (1994) Normality as a biological concept. *Philosophy of Science*, 61, 579–91.
- Wakefield, J.C. (1992) The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist*, 47, 373–88.
- Wright, L. (1973) Functions. *Philosophical Review*, 82, 139–168.