Against the Argument from Functional Explanation

Thomas W. Polger
Department of Philosophy, ML 0374
University of Cincinnati
Cincinnati, OH 45221-0374, USA

thomas.polger@uc.edu

voice: (513) 556-6330
fax: (919) 556-2939

Versions presented at SSPP 2001 and ISHPSSB 2001

Against the Argument from Functional Explanation

ABSTRACT

There is an argument for functionalism—and *ipso facto* against identity theory—that can be sketched as follows: We are, or want to be, or should be dedicated to functional explanations in the sciences, or at least the special sciences. Therefore—according to the principle that what exists is what our ideal theories say exists—we are, or want to be, or should be committed to metaphysical functionalism. Let us call this the *argument from functional explanation*. I will try to reveal the motivation for making such an argument, and sketch the kind of response that should be made by critics of functionalism.

Against the Argument from Functional Explanation

1. The Argument from Functional Explanation

There is an argument for functionalism—and *ipso facto* against identity theory—that can be sketched as follows: We are, or want to be, or should be dedicated to functional explanations in the sciences, or at least the special sciences. Therefore—according to the principle that what exists is what our ideal theories say exists—we are, or want to be, or should be committed to metaphysical functionalism.[1] Let us call this the *argument from functional explanation*.

I do not know of anyone who has explicitly defended the argument from functional explanation in print, though I have heard it many times in conversation. Because it is not clearly stated in the literature, I am loathe to attribute it to anyone, or to give it any very specific form.[2] So my treatment will be necessarily schematic. I will try to reveal the motivation for making such an argument, and sketch the kind of response that should be made by critics of functionalism.

---

[1] One way to interpret this argument is as supporting the interpretation of psychological theories in the *de dicto* mode.

[2] The argument from functional explanation is, I think, tacit in Shoemaker's (1981) and Lycan's (1987) arguments that if everything is functional then mental states must be functional. Shoemaker and Lycan are arguing that our theories and explanations are functional—the causal theory of properties, or scientific theories in general—and therefore that they must pick out functional properties or kinds. And this argument is behind Dennett's (1991, 1995, 1998) claims that all the differences that matter are functional differences.

2. The Autonomy Motive

The argument from functional explanation comes nearest to overt expression in debates concerning the "autonomy" of the mental.[3] Putnam is responsible for introducing the idea that the autonomy of the mental is what matters (1975 in Block 1980). He argued that it was a mistake to care about what stuff—"copper, cheese, or soul"—minds are made of; what we care about is the autonomy of minds (1975). At the same time, Fodor argued that being materialists does not require us to be identity theorists, and that as a consequence psychology can go about its business without a mind to whether psychological kinds correspond to physiological kinds (1974). Putnam and Fodor each criticize the identity theorist for making a mistake about what is important in a philosophy of mind. What is important is not the stuff, but the autonomy of psychological explanations.

According to the advocates of autonomy, "a property is real (or autonomous) just in case it is *essentially* invoked in the characterization of a regularity" (Antony and Levine 1997: 91). This means that one has to show that functional explanation in psychology is not only legitimate, but that it is compulsory. It is often supposed to be a consequence of identity theory that we can explain everything about minds without using distinctly psychological language. The worry is that we would then have to dispense with psychological theories and *ipso facto* with the bothersome psychological states, events, and properties that they name. In that case, it may seem that we have no basis for insisting on the reality of psychological states, events, and properties—or on the pertinence of psychological explanations. That is, that Identity Theory entails, or

[3] See the essays in Tomberlin (1997).

at least invites, a brand of eliminativism (Antony and Levine 1997). So the argument goes.

I have tried to state the autonomy worries without using the customary language of "reduction" because I believe that it only confuses matters.[4] Though in philosophy of mind, 'reduction' and 'identity' tend to be used almost interchangeably, it is a great mistake to conflate them. In this context "reduction" is a basically explanatory or theoretical notion, whereas identity is a fundamentally metaphysical notion. At the time that 'identity' and 'reduction' came into use in philosophy of mind, it was thought that theoretical and metaphysical "reduction" were doctrines that necessarily go together. That can no longer be assumed. The idea of explanatory "reduction" has itself come under scrutiny. One symptom of the confusion about "reduction" is that some philosophers of mind have come to argue that concerns about causal exclusion and explanatory autonomy are based on an archaic conception of theoretical "reduction" (e.g., Bickle 1998). But these new accounts are unsatisfying to those of us who are concerned about the metaphysical questions in which we have been entangled.[5] It is no improvement to replace our old concepts of "reduction" with new ones if it is a mistake to think that the status of the metaphysical claims of Identity Theory depend on views about explanatory "reduction" to begin with.

It is not my purpose to settle questions about "reduction." Rather than enter the fray, I want to reconsider what is assumed by "reductionists" and

[4] This also accounts for my annoying device of only using 'reduction' in quotes or scare-quotes.

autonomy-minded functionalists. If the argument from functional explanation is to succeed, then it must begin by securing the premise that our explanations in psychology are, or ought to be, functional in a sense that would—assuming the argument is valid—entail that mental kinds are functional kinds. But that premise is by no means certain.

The argument from functional explanation is almost invariably conducted in terms of causal-role functions. It is not just any old explanation that, on most views, underwrites the reality of its objects; that is a privilege reserved for causal explanations. (It is for that reason that Shoemaker's causal theory of properties makes all causal relations between properties into functional relations.) Every one of the autonomy arguments concerns causal-role functions. In short, most advocates of the argument from functional explanation assume that the functional properties are *all the causal properties*.[6]

Because I am particularly concerned to show that the arguments for functionalism are inadequate, I have focused on the argument from functional explanation. But there is a counter-argument, call it the *argument from "reductive" explanation*, that purports to show that identity theory must be correct. Both arguments accept that what exists is what is *essentially* invoked by our best explanations. And both are bad arguments.

---

[5] Typically this is of no concern to the proponents of new versions of "reductionism" because they are only interested in explanatory claims to begin with. They may see my metaphysical worries as just another sign that I am already headed down a garden path.

[6] Or, in Fodor's case, all the causal properties except those directly appealed to by physics: "Simply to have a convenient way of talking, I will say that a law or theory that figures in bona fide empirical explanations, but that is not reducible to a law or theory of physics, is ipso facto *autonomous*; and that the states whose behavior such laws or theories specify are *functional* states. (In fact, I don't know whether autonomous states are ipso facto functional. For present purposes all that matters is that functional states are ipso facto autonomous.)" (1997: 149)

3. Brandon on Holism, Reductionism, and Mechanism

The debate over "functionalism" and "reductionism" in philosophy of mind is similar to the debate over "holism" and "reductionism" in philosophy of science.[7] The conflict between functionalism and identity theory is typically set-up such that identity theory is said to offer "reductionist" explanations whereas functionalism plays the part of offering "holistic" explanations. "Reductionism" in this context is the idea that "for any level of phenomena (save the lowest one, if such exists) there is a lower level that explains in a causal mechanical way the focal level" (Brandon 1996: 181).[8] Holism, in contrast, contends that "the parts of wholes are interconnected in such a way that they cannot be studied in isolation—at least not if one wants to know how they behave in the whole" (Brandon 1996: 186). If mental states are essentially relational —if they are functional states in Van Gulick's (YYYY) and Shoemaker's (1981) strong sense—then functionalism is a holistic doctrine. It might then seem convenient to oppose identity theory as "reductionist." Thus, the debate between functionalists

[7] I note three differences. First, the debate over "holism" and "reductionism" in philosophy of science typically involves levels of organization, whereas the debate in philosophy of mind traditionally involves orders, or both orders and levels. Second, both Wesley Salmon (1989) and Robert Brandon (1990) take it that the distinction between top-down and bottom-up explanation in the sciences corresponds to (is just another way of describing) the distinction between inside-out and outside-in explanation. But this will not typically be true of the wholesale microphysical "reduction" that concerns philosophers such as Kim and Davidson. Giving an adequate microphysical explanation will usually involve greater spatio-temporal breadth than a macrophysical explanation. If I want to explain a creature's behavior at an organismic level, I can say that it has certain (internal) beliefs and desires, and so forth. Giving the microphysical explanation may involve near Laplacian descriptions of the paths and inertia of particles, and so forth. Finally, the functionalist is still a "reductionist" in terms of philosophy of science. Some functionalist's acknowledge this (e.g., Lycan 1987), and Kim goes to great lengths to demonstrate the point (1998).

[8] In fact, this is what Brandon calls "multi-level reductionism." A stronger view would be "single-level reductionism," according to which there is one level at which all phenomena can be explained. But I have been clear that the Identity Theory is not committed to that claim.

and identity theorists seems to recapitulate a general debate over holism and "reductionism" in scientific explanation generally.[9]

As "reduction" is out of favor generally, holism might at first seem to be a good thing for functionalists. It is not. Robert Brandon (1996) argues that the debate over holism and "reductionism" in biology,

> is based on a confusion—a confusion between reductionism and a doctrine I will call mechanism. Thus… the opposition between reductionism and holism is based on a false choice, and that neither should be endorsed. On the positive side I hope to show that contemporary biology is, and should be, largely mechanistic. (1996: 179-80)

Brandon argues that the correct distinction is between holism and mechanism, "reductionism" being an entirely implausible thesis that received illicit support due to its confusion with mechanism. He locates the confusion in the historical opposition of both "reductionism" and mechanism to vitalism—itself confused with holism (Brandon 1996).

In the choice between holism and mechanism, Brandon urges that, usually if not always, we choose mechanism. Mechanism "[a]s an ontological thesis… is vague and open-ended in comparison to either holism or reductionism.… *A mechanism is just a causal pattern*" (1996: 193-194, italics original). The explanations offered by Identity Theory are mechanistic. Functionalists also want mechanistic explanations.

What functionalists want—and in particular what the advocates of the argument from functional explanation want—is for mental states to be causally

[9] Note, however, that this debate is different from the familiar one about theory "reduction" that stems from Nagel (1961) and is the topic of Bickle (1998). Brandon comments, "Theory reduction… has little or nothing to do with the dispute among biologists concerning reductionism" (1996: 180).

efficacious. This will be the case if mental states can be explained mechanistically and can figure in mechanistic explanations. And this seems like a good idea to me, too. What I now want to suggest is that the reality of the mental does not depend on the autonomy of psychological explanations.

4. Mechanism and Psychology

Consider an anecdote from Wesley Salmon (1989): A "friendly physicist" encountered a child on a plane. The child was holding a balloon on a string, and the physicist asked the child what would happen to the balloon during the takeoff. The child, evidently a child who had not learned not to talk to strangers, answered that the balloon would move toward the back of the cabin. The physicist says that he thinks the balloon would move toward the front of the plane; lo and behold, it behaves as he predicts. As the story goes, the friendly physicist ends-up winning a little bottle of Scotch from the attendant, who did not believe his prediction. Salmon asks,

> Why did the balloon move toward the front of the cabin? Two explanations can be offered…. First, one can tell a story about the behavior of the molecules that made up the air in the cabin, explaining how the rear wall collided with nearby molecules when it began its forward motion, thus creating a pressure gradient from back to front of the cabin. This pressure gradient imposed an unbalanced force on the back side of the balloon, causing it to move forward with respect to the walls of the cabin. Second, one can cite an extremely general physical principle, Einstein's *principle of equivalence*, according to which an acceleration is physically equivalent to a gravitational field. Since helium-filled balloons tend to rise in the atmosphere in the earth's gravitational field, they will move forward when the airplane accelerates, reacting as they would if a gravitational field were suddenly placed behind the rear wall. (1989: 183)

Salmon takes this example to show the compatibility of two views of explanation whose rivalry he chronicles for four decades. The first kind of explanation Salmon calls *causal/mechanical*, the second *unificationist* (1989: 183-185). Causal/mechanical explanations describe the causal processes involved in a phenomenon, such as the balloon's motion; unificationist explanations appeal not to particular causal mechanisms but to general principles, such as the principle of equivalence (1989: 183-85).

According to Salmon's way of categorizing explanations, both functionalism and identity theory provide causal/mechanical explanations of mental states. Both explain what makes a state a mental state by appealing to the mechanisms that produce the phenomena of interest. Both, that is, are broadly mechanistic; they differ over the nature of the mechanisms. Cast in Salmon's terms, the unificationist alternative to functionalism and identity theory is psychology itself, which attempts to subsume mental events under general psychological laws. Salmon and Brandon argue that both causal/mechanical and unificationist explanations can be correct; of course there is still a question about how the different sorts of explanations are related, but each can be correct without excluding the other (Salmon 1989: 183-185; Brandon 1990 160-161). If Salmon and Brandon are correct, autonomy is too strong a requirement on the acceptability of a causal explanation.

The friendly physicist example bears striking resemblance to Putnam's example of the square peg and round hole. Putnam imagines a board with square and round holes cut in it ("region 1" and "region 2"), and a square peg ("system *A*") just smaller in height and width than the two holes: "We have the

following very simple fact to explain: *the peg passes through the square hole, and it does not pass through the round hole.* (1975 in Block 1980: 137)

Putnam considers two explanations. First: a microphysical deduction "from just the laws of particle mechanics or quantum electrodynamics that system *A* never passes through region 1, but that there is a trajectory which enables it to pass through region 2" (1975 in Block 1980: 137). Second:

> that the board is rigid, the peg is rigid, and as a matter of geometrical fact, the round hole is smaller than the peg, the square hole is bigger than the cross-section of the peg. The peg passes through the hole that is large enough to take its cross-section, and does not pass through the hole that is too small to take its cross-section. (1975 in Block 1980: 137-38)

Putnam contends that if we want to explain why the peg goes through one hole and not the other, then we must or should use the second sort of explanation. He claims that the microphysical deduction is not an explanation at all, or "is just a terrible explanation, and why look for terrible explanations when good ones are available" (1975 in Block 1980: 138). Putnam thinks that this example demonstrates that microphysical explanation cannot replace the macrophysical explanation—that macrophysical explanation of the peg's behavior is autonomous. He intends to generalize this point to show that psychological explanation is also autonomous.

Does Putnam really want to deny that there can be microphysical explanations of the behavior of pieces of inert matter? Perhaps we should conclude that he says this only because he believes that there must be some facts left unaccounted for by the microphysical account such that a proper explanation *essentially* involves the macrophysical objects, thus validating their reality and causal power. Otherwise, Putnam fears, square pegs and mental states will turn

out to be unreal. But Salmon and Brandon are telling us that the correctness of different explanations does not depend on the unavailability of alternative explanations.

In philosophy of mind, mechanism is confused with functionalism—they are both opposed to the evils of "reductionism." Functionalists also misconstrue Identity Theory as "reductionist," which it is not. (To be fair, some Identity Theorists contribute to the error, as does the confusion between different "reductionist" doctrine, viz., theory "reduction" and "reductive" explanation.) The purported "reductionist" who denies that the difference between conscious and non-conscious states is a matter of functional role, is then accused of harboring residual Cartesian intuitions of one sort of another (e.g., Dennett 1991, 1995). This is because once one assimilates all causal relations into functional relations, then denying that functional relations are what matter in naturalistic theory seems to be ridiculous. But the mistake is to think one can import all causal relations into functionalism. That is exactly the mistake of confusing functionalism with mechanism. Unfettered "functional" relations have none of the distinctive qualities of more robust—and more restricted—notions of function:

> Without such relativization and addition of detail, the claim that psychological states are functional states is incredibly trivial. Every psychological state has causes and effects, and if we look carefully enough, we can always find a causal difference between any two different psychological states. (Sober 1985: 190)

The functionalist, by loosening the notion of a functional relation until it encompasses all causal relations, receives illicit support from the plausibility of mechanism.

5. A Plea for Mechanism

Please, take a minute (or more) to do some soul-searching. Why not eschew explanatory and metaphysical functionalism in favor of mechanism and Identity Theory? Are you really committed to the view that all explanation should be functional explanation? If so, what sort of function? Is it a causal notion of function? If so, is it constrained by some theory, or does it potentially include every causal relation?

Here is a litmus test. Elsewhere (CITATION REMOVED) I have suggested that to instantiate a functional state is *to have a function*. What do you think of that claim? (Make sure you are using the notion of function univocally in "functional state" and "have a function.") If that seems wrong to you—if you do not think that functional instantiation is *having a function*—then you are not a metaphysical functionalist. Still think you're a metaphysical functionalist rather than just a mechanist? Why, if your theory does not involve having a function in any robust sense?

Think about carburetors. It is true that carburetors have the causal-role function, crudely put, of mixing air and fuel, relative to some explanation of the capacities of a combustion engine to power a vehicle. Does that mean that to be a carburetor is to instantiate the function, crudely put, of mixing air and fuel? No. Fuel injectors mix air and fuel but they are not carburetors. Some carburetors, broken and defective carburetors, do not or are not apt to mix air and fuel. Carburetors are mechanical devices *par excellence*.

Isn't it important, you might wonder, that carburetors are *supposed* to mix air and fuel? Ah, but now we have a different notion of function on our hands—not a causal function, but one that explains how the carburetor came to

be part of the car, how it came to have its form, and so forth. This is an teleological notion. That it is supposed to be a carburetor, however, plays no role in the explanation of what the device, the mechanism, in fact does for the car of which it is a part. But we do not in general think that everything teleological explanations assure causal powers, or are required to underwrite them.

The argument from functional explanation is something like a default argument for functionalism. It is a fallback argument that comes out at dinner tables and bars where it is difficult to assess. But if I am right, it is a mistake. (Don't forget, I condemn the "reductionist" counterargument as well. Confused Identity Theorists are as culpable as functionalists on this matter.) What both sides want, I believe, is mechanism. Whether all explanations should be mechanistic, I do not know. But mechanism, in general, is a fine idea. Identity theory is thoroughly mechanistic.

5. A Plea for Mechanism

Antony, L. and Levine, J. 1997. "Reduction With Autonomy." In Tomberlin (1997).

Bickle, J. 1998. *Psychoneural Reduction: The New Wave*. Cambridge, MA: The MIT Press.

Block, N. 1980. *Readings in Philosophy of Psychology, Volume One*. Cambridge, MA: Harvard University Press.

Brandon, R. 1990. *Adaptation and Environment*. Princeton: Princeton University Press.

Brandon, R. 1996. "Reductionism versus Holism versus Mechanism." In R. Brandon, *Concepts and Methods in Evolutionary Biology* (New York: Cambridge University Press, 1996).

Dennett, D. 1991. *Consciousness Explained*. Boston: Little, Brown, and Co.

Dennett, D. 1995. "The Unimagined Preposterousness of Zombies." *Journal of Consciousness Studies*, 2, 4: 322-26. Reprinted in Dennett 1998.

Dennett, D. 1998. *Brainchildren*. Cambridge, MA: MIT Press.

Fodor, J. 1974. "Special Sciences, or The Disunity of Science as a Working Hypothesis," *Synthese* 28: 97-115. Reprinted in Block 1980.

Fodor, J. 1997. "Special Sciences: Still Autonomous After All These Years," in Tomberlin (1997).

Kim, J. 1998. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press.

Lycan, W. 1987. *Consciousness*. Cambridge, MA: The MIT Press.

Nagel, E. 1961. *The Structure of Science*. London: Routlege and Kegan Paul.

Putnam, H. 1975. "Philosophy and Our Mental Life," in Putnam (1975) and Block (1980).

Putnam, H. 1975. *Mind, Language and Reality: Philosophical Papers, Volume 2*. New York: Cambridge University Press.

Salmon, W. 1989. "Four Decades of Scientific Explanation." In Kitcher and Salmon (eds), *Minnesota Studies in the Philosophy of Science XIII: Scientific Explanation* (Minneapolis, MN: University of Minnesota Press.

Shoemaker, S. 1981. "Some Varieties of Functionalism." *Philosophical Topics* 12, 1: 83-118. Reprinted in Shoemaker (1984).

Shoemaker, S. 1984. *Identity, Cause, and Mind*. New York: Cambridge University Press.

Sober, E. 1985. "Panglossian Functionalism and the Philosophy of Mind." *Synthese* 64: 165-193. Reprinted in Lycan (1990).

Tomberlin, J. (ed). 1997. *Philosophical Perspectives 11: Mind, Causation, and World*. Boston: Blackwell Publishers.

Van Gulick, R. 1983. "Functionalism as a Theory of Mind." *Philosophy Research Archives*: 185-204.