Why Biology still needs Teleology:

A Modern Aristotelian account of Life, Knowledge, and Health

Robert C. Koons

Professor of Philosophy, University of Texas - Austin

Visiting Research Fellow, James Madison Program, Princeton University

Abstract

Since the Galilean Scientific Revolution in the 17th century, and especially since the death of vitalism and the rise of Darwinian evolutionary theory in the late 19th and early 20th centuries, biologists have treated the Aristotelian notion of 'teleology' (the existence of functions or purposes in nature) as utterly outdated or, at best, as a heuristic device or useful fiction. This is however an untenable position, since biological inquiry exists primarily for the sake of biological knowledge, and biological knowledge is inextricably bound up with teleological concepts, like that of *gene* or *enzyme*. Moreover, the very possibility of rational thought and knowledge depends upon a teleological foundation. Alexander Pruss and I have argued that the Functionalist theory of mental states can only work in a teleological setting. In addition, attempts to reduce teleology to a supposedly non-teleological basis (exemplified by the work of Ruth Garrett Millikan) have come to nought. Finally, modern quantum physics points the way to a robust sort of emergence of new

causal powers at both the chemical and biological scale, enabling a revivification of scientific teleology, properly understood. If this is indeed the case, this fact has great implications for biomedical ethics and the vocation of the physician.

1 Is Teleology Merely Heuristic?

Some skepticisma about the place of teleology had begun already in the high middle ages with Duns Scotus and William of Ockham, but it was the influential figure in the English Renaissance, Francis Bacon, who explicitly banned all teleology and final causation from true natural philosophy, or "science" as it came to be called. This turn from teleological explanations to quantitative explanations in terms of momentum, force, and energy were at the heart of the Scientific Revolution in astronomy and physics, including the work of Descartes, Galileo, Kepler, and Newton.

For quite some time, biology continued to explicitly and unapologetically teleological in character, but over time biologists came more and more to emulate modern physics and to seek to find a place for living things within a purely quantitative and bottom-pick explantory picture. This indisputably led to great advances in biochemistry, from the synthesis of urea to the discovery and mapping of DNA. In my view, the reduction of living things to chemistry should be thought of as a merely useful fiction, not to be taken literally true. But in fact, most biologists have taken exactly the opposite view: treating physical reductionism as the sober truth, and the teleological element in biology as a mere "heuristic", a useful fiction.

The eventual acceptance of Darwin's theory of evolution seemed to many to clinch the matter, since Darwin could be taken as a way of explaining how it is that things seem to have purposes and functions, even though they are in reality mere concresences of matter, driven into repeating patterns by physical and chemical forces alone. T. H. Huxley made the point with his characteristic bluntness:

> "That which struck the present writer most forcibly on his first perusal of the 'Origin of Species' was the conviction that Teleology, as commonly understood, had received its deathblow at Mr. Darwin's hands. For the teleological argument runs thus: an organ or organism (A) is precisely fitted to perform a function or purpose (B); therefore it was specially constructed to perform that function."[1]

Huxley and many others took Darwin's theory as a way of explaining the apparent fitness of organs for purposes without appealing to the purposes themselves.

S. H. P. Madrell writes that "the proper but cumbersome way of describing change by evolutionary adaptation [may be] substituted by shorter overtly teleological statements" for the sake of saving space, but that this "should not be taken to imply that evolution proceeds by anything other than from mutations arising by chance, with those that impart an advantage being retained by natural selection."[2]

---

[1] From "Criticism on 'The Origin of Species'", *Natural History Review*, 1864, p. 7.
[2] S. H. P. Madrell(1998) "Why are there no insects in the open sea? "*The Journal of Experimental Biology* 201:2461–2464.

The famous 20th century biologist J. B. S. Haldane is supposed to have quipped, "Teleology is like a mistress to a biologist: he cannot live without her but he's unwilling to be seen with her in public."

Haldane's witticism points to an important fact: teleological language and concepts are ubiquitous and ineliminable in biology. If we suppose that they are merely 'heuristic', we have to ask, heuristic for what? To what further discoveries do teleological models lead? Only to still more teleological knowledge. It would be crazy to suppose that all of biology is merely a fiction, useful only as a tool for additional chemical and physical discoveries. In fact, physics and chemistry can do quite well on their own: they stand in no need of biology. Biology exists for its own sake, and biological inquiry never escape from the teleological domain.

As Georg Toepfer has put it in a recent essay:

> "…teleology is closely connected to the concept of the organism and therefore has its most fundamental role in the very definition of biology as a particular science of natural objects…. The identity conditions of biological systems are given by functional analysis, not by chemical or physical descriptions…. This means that, beyond the functional perspective,

> which consists in specifying the system by fixing the roles of its parts, the organism does not even *exist* as a definite entity." [3]

This was recognized by the Neo-Kantians of the early 20th century:

> "We even have to define this science [biology] as the science of bodies whose parts combine to a teleological 'unity'. This concept of unity is inseparable from the concept of the organism, such that only because of the teleological coherence we call living things 'organisms'. Biology would therefore, if it avoided all teleology, cease to be the science of organisms as organisms."[4]

Evolution itself presupposes a strong form of teleology in the very idea of 'reproduction'. No organism ever produces an exact physical duplicate of itself. In the case of sexual reproduction, the children are often not even close physical approximations to either parent at any stage in their development. An organism successfully *reproduces* itself when it successfully produces another instance of its biological kind. This presupposes a form of teleological essentialism.

---

[3] Georg Toepfer (2012), "Teleology and its constitutive role for biology as the science of organized systems in nature," *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 43: 113-119, at 113, 115, 118.

[4] H. Rickert (1929), *Die Grenzen der naturwissenschaftlichen Begriffsbildung* (Tübingen: Mohr), p. 412; cited and translated by Toepfer, op cit. p. 114.

Richard Dawkins has suggested that we think of organisms as mere "robots" that our DNA molecules have "designed" for reproducing themselves. In fact, DNA molecules never succeed in producing perfect physical duplicates of themselves, and even if they did, the mere physical duplication of the molecule would not constitute reproducting. Suppose, for example, that an extrinsic billionaire builds a chemcial factory that does nothing but fill barrels with copies of his own genome, launching them into deep space. No one would think that such a man had succeeded in procreating trillions of descendants. A DNA molecule counts as a copy of one of one's genes only when it is successfully fulfilling the function of a gene within a living organism, indeed, within a living organism of the appropriate teleologically-defined kind.

I defend an all-out realism about teleology. Teleological statements in biology are both literally true and indefinable in non-teleological terms. Anti-realism about any matter in philosophy takes one of two forms: reductive or eliminative. According to a reductive account of teleology, teleology is a real phenomenon, but it is in reality identical to or wholly constituted by certain non-teleological facts. The world is fundamentally non-teleological, but certain complex facts about that non-teleological world can be fittingly described in teleological language or using teleological concepts. *Eliminative* anti-realism about teleology has a harder edge: it simply denies that that there are any truths couched in teleological terms. All teleology is at best a useful fiction, at worst an error or confusion. I will argue against eliminative teleological anti-realism in sections 2 and 3, and against

reductive anti-realism in section 4. I will then sketch an Aristotelian, hylomorphic conception of living things in section 5, and conclude in section 6 with a discussion of the ethical implications of a robust teleological realism.

2 Is Teleology Eliminable?

We've already seen that biology gives us good reason to reject the eliminability of teleology. Teleology is essential to the practice of the biological sciences. However, some hard-core anti-teleologists might be willing to bite the bullet and insist that biology itself is merely a useful fiction. On this view, all real science is fundamental micro-physics; everything else is merely "stamp collecting," as Ernest Rutherford is supposed to have said.

I have argued in several places[5] that teleological realism is needed to account for the possibility of science of any kind, including physics. The reason for this is that we must suppose that the human mind has a natural predisposition to gravitate toward true scientific theories. It is a commonplace of the philosophy of science to recognize that the choice of theory is always under-determined by the data: there are always going to be an infinite number of mutually inconsistent theories and models that agree with any body of empirical data. Theory choice is in fact guided by quasi-aesthetic considerations, like simplicity, elegance, and symmetry. This can be

[5] "The Incompatibility of Naturalism and Scientific Realism," in *Naturalism: A Critical Appraisal*, ed. William Lane Craig and J. P. Moreland (Routledge, London, 2000), pp. 49-63; reprinted in *The Nature of Nature*, edited by Bruce L. Gordon and William A. Dembski (ISI Books, 2011); "Epistemological Problems with Materialism," in *The Waning of Materialism*, edited by Robert C. Koons and George Bealer (Oxford University Press, Oxford: 2010), pp. 281-308.

indicative of theoretical truth only if there is a kind of pre-established harmony between the aesthetic dispositions (both innate and acquired) of practicing sciences and the deep structure of the laws of nature. For reasons of time, I won't rehearse those arguments here.

Instead, I will focus on the necessity of teleology for the understanding of the human mind.

3 Why Teleology is Needed in Accounting for the Mind

3.1.1 The Classical, Aristotelian Model of the Mind

In the Aristotelian picture of the world, popular until the Great Disruption of the seventeenth century, the natural world is suffused with teleology, functionality, and normativity. Each real entity has its own natural end or purpose. Sensory knowledge and intentional agency are fully natural processes, processes of a kind that is ubiquitous in nature.

Once Bacon and Descartes succeeded in overturning the Aristotelian model, a new picture emerged of the so-called "physical" world. The physical world was one of essentially inert quantities of matter, moved hither and thither according to blind forces governed by strict "laws of nature". Sensation, thought, and agency are clearly processes of a very different kind. Consequently, most early modern philosophers followed Descartes into a kind of metaphysical dualism of mind and matter. This

proved to be an unstable, ultimately inhospitable halfway house on the road toward all-out materialism.

The modern materialist is left with two critical problems: where to locate the qualitative aspects of reality, as encounted in sense experience (the so-called secondary qualities or qualia), and how to ground the reality of intentionality, the aboutness of thought. Both are insoluble for the materialist. On the first, see Bealer and Koons, *The Waning of Materialism*. I will focus here on the problem of intentionality.

3.1.2 The Modern Alternative: Functionalism

I am going to draw in the remainder of section 2 from a joint paper written by Alexander Pruss and me, "Must a Functionalist be an Aristotelian?" to appear next year in an OUP anthology, *Causal Powers*, edited by Jonathan Jacobs.

The philosophical idea called 'Functionalism' remains the most promising strategy for "materializing" or "physicalizing" the mind, that is, for accounting for the intentionality of the mind in a physical world devoid of Aristotelian teleology. Physicalizing or "naturalizing" (as many materialists prefer to put it) the mind demands that the fundamental vocabulary of psychology must be wholly physical (for description of inputs and outputs), plus the language of causation, dispositions, conditionals, or function, as well as the terms of logic and mathematics, achieving as a result a thoroughly non-teleological language. British philosopher and logician

Frank Ramsey[6] offered the logical tools needed to express mature Functionalism, describing a logical process that has come to be known as "Ramsification". We start with the true theory of psychology, one including explicitly mental terms and predicates (like 'pain' or 'conscious of'). This theory is supposed to capture the one Pattern of interactions that is definitive of having a mind. We then replace the mental terms and predicates with variables, which can be taken to stand for some unknown physical entities and states that do in fact realize the relevant Pattern. The naturalized theory of the mind consists of the resulting "Ramsey sentences".

Clauses of the Ramsey sentence will have a form something like one of these:

(1) If the system *x* is in internal state S_n and in input state I_m at time *t*, then *x* at the next relevant time *t*+1 is in internal state S_k and output state O_j. (Indicative conditional)

(2) If the system *x* were in internal state S_n and in input state I_m at time *t*, then *x* would at time *t*+1 be in internal state S_k and output state O_j. (Subjunctive conditional)

(3) Whenever the system *x* is in internal state S_n, *x* has a disposition to enter immediately into output state O_j and internal state S_k in response to input state I_m. (Dispositional state)

[6] Ramsey, F.P. (1929), "Theories," In R.B. Braithwaite (ed.), *The Foundations of Mathematics and Other Logical Essays*, pp. 212–236, Paterson, NJ: Littlefield and Adams.

(4) System x's being in internal state S_n confers upon it the power to produce output state O_j and internal state S_k immediately in response to input state I_m. (Causal power)

Here, x is either the whole mind or a subsystem. This could in principle be a very low level subsystem, say one that takes two truth value inputs and returns their disjunction, or a very high level one, say one that takes desires and beliefs and outputs motor activation signals.

The project of Ramsifying psychology raises a number of questions, including the following:

• What sort of language is involved in the specification of the links between inputs, internal states and outputs?
• Does the theory make use of conditional statements of some kind, or does it make reference to causal powers or intrinsic dispositions?
• If it does make use of causal powers or intrinsic dispositions, how are these to be understood?

- As equivalent to the truth of a conditional statement, whether absolute or probabilistic?
- Or in an Aristotelian way, according to which powers are intrinsic properties of substances, definable in terms of their inputs and outputs,

and conferred on substances by essential or accidental natures, and dispositions are powers together with a teleological directedness towards their exercise?

First we will argue that conditional views are untenable, and then we will evaluate the dispositional and powers views. We will argue that the only plausible form of functionalism requires that the connections between inputs, outputs and mental states be described as causal powers, in accordance with the assumptions of standard Aristotelian metaphysics. Thus, ironically, the best attempt to escape teleology leads us inexorably back into its clutches.

3.1.3 Conditional functionalisms

Standard problems with conditional accounts of dispositions apply just as well to all the non-material conditional forms of the accounts. We can imagine, for instance, that the individual has strapped to her a bomb that explodes if system x is in internal state S_n and receives input I_m at time t, but that in fact this condition does not obtain. Then, the conditionals (1) and (2) will both be false. Yet having such a bomb that never goes off strapped to one, while unfortunate, does not make one not have a mind.

One might try to use context-sensitivity to ward off such worries, for instance by saying that in evaluating conditionals or conditional probabilities we only should consider those causal factors that are internal to the system *x*. But we can replace

the bomb by a fatal disease, and the distinction between "internal" and "external" causal factors will become untenable.

What if the antecedents of the conditionals are strengthened to include the claim that the whole system survives until the next relevant time? Here we borrow an idea from Harry Frankfurt: the introduction of a purely hypothetical neural-manipulator.[7] The manipulator wants the subject to follow a certain script. If the subject were to show signs of being about to deviate from the script, then the manipulator would intervene internally, causing the subject to continue to follow the script. Moreover, if by some near-miracle the subject succeeded in deviating from the script for a step, the manipulator would push the subject right back to the script. We are to imagine that the subject spontaneously follows the script, and as a consequence, the manipulator never intervenes.

Frankfurt introduced such a thought experiment to challenge the idea that freedom of the will requires alternative possibilities. We use it to show that the existence of mental states is independent of the truth of conditionals linking the states to inputs, outputs and each other. It is obvious that the presence of an inactive manipulator cannot deprive the subject of his mental states. However, the manipulator's presence is sufficient to falsify all of the usual conditionals and conditional probabilities linking the states. If the manipulator's script says that at time t+1 the

[7] Harry Frankfurt (1969), "Alternate Possibilities and Moral Responsibility," *Journal of Philosophy* 66: 829–39.

subject is to be in state S_n, then that would happen no matter what state the subject were in at time *t*.

Again, it won't do to say that the conditionals need to hold on the assumption of no external interference.[8] For we can always replace an external intervener by an internal one—say, an odd disorder of the auditory center of the brain that causes it to monitor the rest of the brain and counterfactually intervene.

Moreover, cognitive malfunctioning is surely possible as a result of injury or illness. The theory to be Ramsified cannot plausibly incorporate the effects of every possible injury or illness, since there are no limits to the complexity of the sort of phenomenon that might constitute an injury or illness. Injury can prevent nearly all behavior – so much so, as to make the remaining behavioral dispositions so non-specific as to fail to distinguish one internal state from another. Consider, for example, locked-in syndrome, as depicted in the movie *The Diving-Bell and the Butterfly*. Therefore, the true psychological theory must contain postulates that specify the *normal* connections among states.

Without resorting to Aristotelian or evolutionary teleology (an option we will discuss later), our only account of normalcy will be probabilistic. Thus, a system *normally* enters state S_m from state S_n as a result of input I_m provided it is *likely* to do this. However, serious injury or illness can make a malfunctioning subsystem rarely

[8] Martin Smith (2007), "Ceteris Paribus Conditionals and Comparative Normalcy," *Journal of Philosophical Logic* 36 (1):97 - 121.

or never do what it should, yet without challenging the status of the subsystem as, say, a subsystem for visual processing of shapes. And, again, a merely counterfactual intervener. whether external or internal, can change what the system is likely to do without manipulating the system in any way.

Alternately, one might try to define normalcy in terms of what systems *of the same type* are likely to do. Thus, a system *normally* enters state S_m from state S_n as a result of input I_m provided that most of the time systems of *this type* do this. A serious problem here is that we are giving the functional claims in order to characterize the *type* of system. But it is then circular in the functional claims to refer to other systems of the same type. One might try to Ramsify over types to solve this problem, but one will still have problems with one of a kind minds.

Moreover, the probabilities of state transitions in systems of a given kind depend deeply on the environment the systems are in. A plausible account would have to say that a normal transition is one that is likely to occur in systems of the given type *in a normal environment*. But, again, it does not appear possible to specify a normal environment without resorting to something like teleology or proper function.

4 Is Teleology Reducible? Two Theories of Normativity

Thus, any Functionalism with a hope of success must have an account of the normal state of the organism, where the 'normality' involved is a normative notion, not

merely a matter of averages or actual frequences. There are two prima facie plausible accounts of the natural basis of normativity: Aristotelian powers and evolutionary accounts.

4.1 Aristotelian Normativity

An Aristotelian can give a straightforward account of normativity: a substance is supposed to produce *E* on occasions of *C* if and only if its nature includes a *C*-to-*E* power (one might also prefer more active terms like "tendency" or "striving").

This account may appear insufficient in the light of the possibility of indeterministic powers. Could not a substance have both a *C*-to-*E* and a *C*-to-non-*E* power, in which case it would neither be supposed to produce *E* in *C* nor to produce non-*E* in *C*? One might complicate the account by excluding such cases of competition in some way, or positing higher order powers that decide between the competing powers. But there are also two simpler moves. One move is to say that in such cases, the substance is in the "unhappy" position of being supposed to do incompatible things—it will necessarily fail at one of them.

A more complex move is to say that it cannot happen that a substance has both a *C*-to-*E* and a *C*-to-non-*E* power. Rather, the substance has a *C*-to-*E* and a *C*'-to-non-*E* power, and if it happens that both *C* and *C*' obtain, then the substance will fail to do one of the things it should do. This move fits with a natural metaphysical interpretation of quantum indeterminacy. Take an electron in the mixed spin state

|up>+|down>, and measure the electron's spin, thereby forcing the electron's state to collapse indeterministically to |up> or to |down>. Suppose the electron ends up going to |up>. What explains its going to |up> is not that the electron used to be in state |up>+|down>. Rather, what explains its going to |up> is that the electron used to be in a state that had a |up> component (or had a significant such component). That the state also had a |down> component is true but does not help to explain the electron's transitioning to |up>. Thus, the electron has two powers with incompatible outcomes and different, but potentially co-occurring, activating conditions: (a) being in a measurement situation with a state with an |up> component and (b) being in a measurement situation with a state with a |down> component.

Functionalism can then be put in an Aristotelian mode, referring to the presence of powers to produce outputs and internal states (including other powers). The result would be a non-reductive and non-physicalist version of functionalism, since the form of the theory would rule out the states' realizers being merely physical states of constituent particles.[9]

4.2 Evolutionary accounts of normativity

The other potential source of normativity is evolutionary selection. If a system *x* belongs to a reproductive family *F*, then *x* is supposed to produce *E* under circumstances *C* if and only if doing so is one of *F*'s adaptations. This seems to be the

[9] See George Bealer, "The Self-Consciousness Argument," in Robert C. Koons and George Bealer (2010), *The Waning of Materialism: New Essays on the Mind-Body Problem*, Oxford: Oxford University Press.

most promising alternative to the Aristotelian account, since there doesn't seem to be any vicious circularity or regress.

Ruth Garrett Millikan developed such an account in considerable detail (in *Language, Thought and Other Biological Categories*).[10] Here is a simplified version of her definition, which will be a paradigm of such accounts of normativity:

(5) A thing *x* is supposed to produce *E* in circumstances *I* if and only (i) *x* belongs to a reproductive family *R* in which some feature *C* occurs non-accidentally with finite frequency (between 0 and 1), (ii) there has been a positive correlation between having feature *C* in *R* and producing *E* in circumstances *I*, and (iii) this positive correlation has been in part causally responsible for the successful survival and proliferation of family *R* (including *x* itself).[11]

Similar proposals have been made by Larry Wright, Karen Neander, Nicholas Agar, Kim Sterelny, David Papineau, and Fred Dretske.[12]

[10] Ruth Garrett Millikan (1984), *Language Truth and Other Biological Categories*, Cambridge, Mass.: The MIT Press.

[11] Millikan 1984, p. 28. Millikan's actual definition requires that *C* be a "Normal" or reproductively established characteristic of *R*. Instead of requiring that *C* be positively correlated in R with the function *F*, she requires only that the positive correlation hold in some set *S* which includes *x*'s ancestors, together with "other things not having *C*." Her exact wording of clause (3) is:

> One among the legitimate explanations that can be given of the fact that x exists makes reference to the fact that *C* correlated positively with *F* [i.e., the function of producing *E* in circumstances *I*] over *S*, either directly causing reproduction of *x* or explaining why *R* was proliferated and hence why *x* exists.

None of these variations would make any difference to our objection.

[12] Larry Wright, "Functions," *The Philosophical Review* 82 (1973):139-168; Karen Neander, "The Teleogical Notion of a Function," *Australasian Journal of Philosophy* 69 (1991):454-468; Karen Neander, "Misrepresnting and Malfunctioning," *Philosophical Studies* 79 (1995):109-141; Kim Sterelney, *The Representational Theory of Mind* (Oxford: Blackwell, 1990); Nicholas Agar, "What do

There are a number of objections to these evolutionary accounts.

Objection 1: Can 'reproduction' be defined naturalistically and without reference to function or teleology? Complex organisms (especially ones that reproduce sexually) never produce exact physical duplicates of themselves. Conversely, since everything is similar to everything else in some respects, every cause could be said to be 'reproducing' itself in each of its effects. Real reproduction involves the successful copying of the essential features of a thing. For living organisms, these essential features consist almost entirely of biological functions. Hence, we cannot identify cases of biological reproduction without first being able to identify the biological functions of things. Yet Millikan's account requires us to put the reproductive cart before the functional horse.

A Millikanian version of functionalism would have the consequence that a thing has a mind only if it belongs to a reproductive family *R* for which the standard Pattern of dispositions has successfully contributed to the survival of *R*. Thus, whether a thing has a mind depends on the evolutionary history of its kind. This engenders a second problem.

Frogs Really Believe?" *Australasian Journal of Philosophy* 71 (1993):162-185; David Papineau, *Philosophical Naturalism* (Oxford: Oxford University Press, 1993); Fred Dretske, *Naturalizing the Mind* (Cambridge, Mass.: MIT Press, 1995).

Objection 2: Millikanian functionalism has the implausible consequence that mental functioning is one generation behind neural functioning. For a mutation can never be normal on her account in the generation in which it first occurs—it only becomes normal in their descendants. For instance, on this view, presumably one of our distant vertebrate ancestors, call it Sim, evolved the first form of those neural structures that are responsible for consciousness. But it was Sim's children, not Sim, that were conscious if we use Millikanian functions as the backing for functionalism. For on Millikanian views, the structures as found in Sim did not function normally. It was only once their non-normal functioning helped Sim reproduce that they functioned normally in Sim's descendants and hence made them conscious. Not only is this an implausible claim, but it has an undesirable epiphenomalist consequence. Consciousness as such is useless to us—it does not affect our action or fitness. Assuming Sim's children had no relevant new mutations, their behavior was much like Sim's, but they were conscious while Sim was not.

Objection 3: What does it mean for a particular disposition to 'cause' or to 'contribute' to a particular instance of R-reproduction? There are two possible answers. First, we could say that the disposition contributed to the act of reproduction just in case some exercise of the disposition by the parent occurs in the actual causal history of the creation of the child. Second, we could instead require that the disposition be part of a *contrastive* explanation of the reproduction: part of a minimal explanation of why in this instance reproduction or survival occurred, as opposed to not occurring.

The first answer would greatly over-generate adaptations. Any feature of the parent that is both the product of some disposition of the parent and that influences in any way the process of reproduction would count as one of the kind's essential adaptations. For example, suppose that rabbits are disposed to twitch their left rear leg whenever a cosmic ray strikes the spinal cord at a single point, and suppose that this disposition was actually exercised by some rabbit in the past as it was successfully locating a bunch of carrots. Even if the twitch played no role in explaining the rabbit's survival, it would still count as adaptive, so long as it was part of the total cause of this rabbit's survival in this concrete instance.

Thus, we'll need to turn to the second answer, contrastive explanations. The use of contrastive explanation fits standard biological practice, which identifies adaptations with the results of natural selection, and selection is inherently contrastive in nature.

Now to our third objection. Say that a region R of spacetime is *impotent* provided that nothing in R can affect what happens in spacetime outside R. Consider first the following principle:

(6) (Almost global supervenience of physical minds.) Suppose worlds w_1 and w_2 are exact physical duplicates, except in an impotent region R of spacetime. Then w_1

contains an instance of mindedness outside of R if and only if w_2 contains an exactly similar instance outside of R.

Imagine a world w_1 which contains a planet much like earth, where history looks pretty much like it looks on earth, and which also contains a Great Grazing Ground (GGG), which is an infinite (we only need: potentially infinite) impotent region. Moreover, by a strange law of nature, or maybe the activity of some swamp aliens, whenever an organism on earth is about to die, it gets hyperspatially and instantaneously transported to the GGG, and a fake corpse, which is an exact duplicate of what its real corpse would have been, gets instantaneously put in its place on earth. (We will call it "earth" for convenience but we shan't worry about its numerical identity with our world's earth.) Furthermore, there is no life or intelligence outside of earth and the GGG.[13] Moreover, the organism dies as soon as it arrives in the GGG.

Our world's earth has organisms with real minds, and the earth in w_1 has a history that is just about the same. The only difference is that in w_1 all the deaths of organisms occur not on earth but in the GGG, because they get transported there before death. But this does not affect any selective facts. Thus, the evolutionary theorist of normativity should say that the situation in w_1's earth is similar enough to that on our earth that we should say that w_1's earth contains organisms with exactly the same minds.

[13] Assume that any swamp aliens who created the GGG and the transport system don't count as alive or intelligent.

The hard work is now done. For imagine a world that is exactly like w_1 outside of the GGG, but inside the GGG, immortal and ever-reproducing aliens rescue each organism on arrival, fixing it so it doesn't die, and even make the organism capable of reproduction again. Furthermore, they do the same for the organism's descendants in the GGG. The GGG is a place of infinite (at least potentially) resources, with everybody having immortality and reproduction, with the aliens shifting organisms further and further out to ensure their survival.

Now in w_2, there is no selection: Nobody ever dies or ceases to reproduce. Thus, by Millikan's definition (5), on the contrastive reading, there is no mindedness outside the GGG in w_2—all the earthly critters are functionless zombies. But, by principle (6), there must be instances of mindedness outside the GGG in w_2, because w_2 is an exact duplicate of w_1 outside of the GGG. Hence we have absurdity. This same result obtains in the case of the forward-looking definition: since every member of every population has a perfect propensity to survive and reproduce, no specific trait contributes causally to that propensity.

Suppose our evolutionary theorist of mind denies (6). Then we have the following absurdity: It is up to the aliens in the GGG to determine whether or not there are instances of teleology (including cases of mindedness) outside the GGG, by deciding whether to rescue the almost dead organisms that pop into the GGG. But how can

beings in an impotent region bring about that there are, or are not, minds outside that region? That would be worse than magic (magic is presumably causal).

In the GGG story with post-transportation rescue, there is no natural selection, but surely there is mindedness. This shows that not only are Millikan-type stories insufficient for functionalist purposes, but *no* story on which the normativity of mental functioning is grounded in natural selection facts has a chance of succeeding.

5 Aristotelian (Hylomorphic) Emergence

Aristotelian metaphysics to the rescue! Aristotelians never faced the problem of 'naturalizing' the mind that has so bedeviled modern philosophers. What we must do is reverse engineer the Aristotelian solution to the problem, re-discovering the elements that are essential to locating all of the human mind (including our capacity for science) within the natural world.

The fundamental difference between Aristotelian and modern materialist metaphysics lies in their differing conceptions of causation. Aristotle argued that we must understand change in terms of action, action in terms of causal powers, and powers in terms of essences or natures of things. As we have seen, it is a robust conception of causal powers that is needed to ground normativity in nature. Moreover, causal powers are inherently teleological. To have the power to produce E in circumstances C is to have the C-to-E transition as one of one's natural functions. Indeed, as George Molnar has pointed out, the ontology of causal powers

builds intentionality into the very foundations of natural things. To have a power is to be in a kind of intentional state, one that is in a real sense "about" the effects one is pre-disposed to produce.

Modern philosophy took its most decisive step away from Aristotle when David Hume attacked the very idea of causal powers. Although Hume himself was a kind of eliminative anti-realist about causation, many so-called "Neo-Humeans" have taken his work instead in a reductive direction. The result has been the Mill-Ramsey-Lewis account of causal laws. According to this account, "causation" occurs whenever a transition is deducible from the "laws of nature", and a generalization is a law of nature if it is an axiom in the simplest and most powerful representation of the actual course of the world's history.

In the last twenty years or so, there has been a significant revival of an Aristotelian conception of causal powers within Anglophone philosophy. Dissatisfaction with the Mill-Ramsey-Lewis account has come primarily from four sources:[14]

1. Small worlds objections. Causal connections should be intrinsic, while the MRL account makes them all extrinsic to each transition.

[14] Some important *loci*: Michael Tooley (1987), *Causation: A Realist Approach* (Oxford: Clarendon Press); Phil Dowe (2000), *Physical Causation* (Cambridge: Cambridge University Press); Brian Ellis (2001), *Scientific Essentialism* (Cambridge: Cambridge University Press); George Molnar (2003), *Powers: A Study in Metaphysics*, Oxford: Oxford University Press. See also Chapter 3 of Robert C. Koons and Timothy H. Pickavance (2014), *Metaphysics: The Fundamentals* (New York: Wiley-Blackwell).

2. The inability of the MRL account to solve the problem of causal direction or asymmetry.

3. The extreme anthropocentricity of the MRL account.

4. The inability of the MRL account to solve Hume's problem of induction.

When later medieval thinkers like John Duns Scotus and William of Ockham began to remove the teleological element from their accounts of sub-rational and inorganic nature, they did so on the basis of a misunderstanding of what natural teleology really amounts to on the conception of Aristotle and Aristotelians like Avicenna or Thomas Aquinas. Aristotle did not suppose that non-living or non-sentient entities were somehow consciously pursuing some end, nor did he think that the postulation of real teleology required by definition the introduction of a conscious designer or user of the teleologically ordered system. The mere possession causal powers, in the full-blooded Aristotelian sense, suffices for teleology.

Many scientists and philosophers of science have assumed that the Galileo-Newton revolution in physics has consigned teleological explanation to the dustbin. However, this overlooks the continued vitality of teleological explanations in physics in the form of so-called 'variational principles', such as least action

principles.[15] Both classical and quantum mechanics can be formulated in terms of integral equations, which prescribe a path or trajectory that satisfies a holistic requirement, like the local minimization of action. In most cases, the same physical theory can be case either in terms of differential equations (with associated notions like the composition of forces and the conservation of energy or momentum) or in terms of integral equations (corresponding to teleological explanation).[16] Many scientists assume that since the integral form of a theory can be transformed into a differential form, this means that the differential form represents the more fundamental mode of explanation. This is a non sequitur, however, since differential forms can similarly be transformed into integral forms. From an Aristotelian point of view, the two forms of explanation are simply complementary.

Biological teleology requires two things: a causal powers metaphysics, and *emergent* powers at the level of organs and organisms. The argument in section 2 gives us good reason to attribute real causal powers of a sentient and rational sort to whole human organisms. Given the importance and success of biology, it is reasonable to extend this attribution to all living things. Thus, reproduction, nutrition, metabolism, growth, development, sensation, perception, and behavioral responsivity are all plausible candidates for explanation in terms of causal powers that are emergent in

[15] Wolfgang Yourgrau and Stanley Mandelstam, *Variational Principles in Dynamics and Quantum Theory* (Dover Publications, New York, 1979), pp. 19-23, 164-167; Cornelius Lanczos, *The Variational Principles of Mechanics* (4th edition, Dover Publications, New York, 1986), xxvii, 345-6; Robert Bruce Lindsay and Henry Morgenaw, *Foundations of Physics* (Dover Publications, New York, 1957), pp. 133-6.

[16] See Val Dusek, "Aristotle's Four Causes and Contemporary 'Newtonian' Dynamics," in *Aristotle and Contemporary Science*, vol. 2, D. Sfendoni-Mentzou, J. Harriangadi and D. M. Johnson, eds. (Peter Lang, New York, 2001), pp. 81-93.

the strong or ontological sense. That is, such powers are not to be identified with the mere conglomeration of the powers of the constituent particles and fields but instead have a fundamental reality and activity of their own.[17]

This emergence of new powers at the macroscopic, biological scale should be unsurprising, given the fact that, according to our most recent quantum mechanical models, we see strong or ontological emergence at the mesoscopic scale in solid-state physics and chemistry. Mesoscopic systems, like ferromagnets, superconductors, and Bose-Einstein condensates, all exhibit dynamical behavior, in the form of spontaneous symmetry breaking and thermodynamic irreversibility, that are irreducible to the microstates of the constituent particles: irreducible not just in practice but in principle, since it can be proven that microscopic models for N particles – no matter how large N is – cannot account for these observable features.[18]

6 Practical Consequences of Teleology

If biological teleology is real and irreducible, does it have practical consequences?

[17] In a recent article, I have developed an account of how such emergent powers could be realized in a world like ours: "Staunch vs. Faint-hearted Hylomorphism: Toward an Aristotelian Account of Composition," *Res Philosophica* 91 (2014):1-27.

[18] See P. Anderson, P. (1972), "More is Different," *Science* 177, 4047 (1972): 393-6; Geoffrey L. Sewell (1985), *Quantum Theory of Collective Phenomena* (Oxford: Clarendon Press); Geoffrey L. Sewell (2002), *Quantum Mechanics and its Emergent Macrophysics* (Princeton: Princeton University Press); Margaret Morrison (2006), "Emergence, Reduction and Theoretical Principles: Rethinking Fundamentalism," Philosophy of Science 73:876-887; Robin Hendry (2010), "Ontological Reduction and Molecular Structure," *Studies in the History and Philosophy of Modern Physics* 41: 183-191; Richard Healey (2011), "Reduction and Emergence in Bose-Einstein Condensates," *Foundations of Physics* 41: 1007-1030.

Yes – as the famous function argument from Book I of Aristotle's *Nicomachean Ethics* (Book 1, chapter 7) demonstrates. If human beings have a teleologically structured essence, then our faculty of *will*, our capacity to make rational choices, is itself a biological function, with a built-in ordering to some fixed end. We can call that fixed end 'happiness' ('eudaemonia' in Aristotle's Greek).

We know by empirical investigation that the biological functions of a given naturally kind are largely harmonious, in the sense of being mutually supportive, for the most part. This generalization can also be supported a prioiri, since biological kinds with disharmonious essences will fail tend to fail in competition for scarce resources.

So, we should expect the human will to be naturally concordant with human life and health, with the proper functioning of our other organs and natural capacities. The rightly ordered will is directed toward the natural end of health. Not everything is subject to legitimate human choice: the properly-functioning human will is directed toward the ultimate end of genuine happiness, which includes the integrity of the human body and its organs.

Medicine's function should be the service of health and the healthy (that is, rightly ordered) will of the patient, not the *arbitrary* will of the patient. Given that there is such a thing as health, an objective ideal toward which all rightly-ordered volition is compatible, it is appropriate that there be a profession – a craft or practice – that is devoted exclusively to the protection and promotion of health. Historically, that has

been the defining mission of medicine. Medicine should not be thought of as biochemical engineering or as an industry whose end is the satisfaction of consumer demand for organic change. Not all desires and inclinations are equally worthy of respect. The ethical physician should not indulge desires for inhuman or 'transhumanist' manipulations of the body: instead, the physician should count such desires as pathologies to be treated.

The ethical physician must always ask: am I treating a disease – a defect in the natural functioning of the body? Am I helping my patient to return to the path of true health, as defined by human nature, that is, by the natural package of emergent causal powers that constitutes the shared structure of human life? Do I refuse to collaborate in my patient's self-destructive choices, even when doing so would be legal and even fashionable?

If we do think of medicine in these terms, then there many practices and procedures that currently exist or that may exist in the near future that simply do not belong within the practice of medicine (whether or not they should be considered morally or legally permissible). Such practices include:

1. The termination of healthy pregnancies.
2. Sterilization or artificial contraception, except as part of the treatment or prevention of some disease (and a healthy pregnancy does not count as a disease).

3. Plastic surgery that is based on preferences that are a matter of contingent taste, and not on the restoration of essential human functioning (including social functioning).

4. Sex change operations, in which sexual organs that are consonant with the patient's genetic make-up are destroyed or mutilated.

5. Prosthetic or pharmacologial enhancements, designed not to cure disease or disability but rather to improve the performance of muscles, skeleton, sense organ, or brain.