Things Oughta Make Sense[1]

Louise Antony
UNIVERSITY OF MASSACHUSETTS AMHERST

Presidential Address delivered at the one hundred twelfth Eastern Division meeting of the American Philosophical Association in Washington, DC, on January 8, 2016.

When I first started doing philosophy, I didn't really know what I was doing. Some will say that not much has changed. And it is true that I don't have a worked-out metaphilosophy. It's still the case that I'm flummoxed by innocent questions about what I do. (On airplanes, I used to sometimes lie and say that I was a nurse or a lawyer or something else people could recognize. But I eventually found that if all I want is for my neighbor to leave me alone, it's far more effective to tell him the truth: "I'm a philosopher. Let me explain my theory of mental causation: it all starts with Descartes. . . .") But even if there are many big questions about philosophy that I can't answer, I still do think that I've gained some insight into what philosophy is. Or, at least, what it could be. Or what it should be.

I had the enormous good fortune as a beginning student of philosophy to fall under the tutelage of some extraordinary philosophers (as well as some extraordinary teachers who were not philosophers—for example, the flutist John Oberbrunner.) My very first philosophy teacher, and perhaps my harshest critic even to this day, was/is Peter van Inwagen. He paid me the immense compliment of taking me—my questions, my objections, my arguments, and even my confusions—seriously. He validated my inchoate conviction that things ought to make sense, and that when they didn't, you had every right to ask questions. A lot of times in my early life I felt like a freak for asking questions no one else seemed to think needed answers: Why should unbaptized babies go to Limbo? Why does Santa Claus give more presents to rich kids than to poor kids? (My interest in this particular matter, I must admit, was not purely intellectual.) In graduate school, I was fortunate to have such luminaries as Hilary Putnam, Jerry Fodor, and Ned Block as interlocuters. These philosophers demonstrated to me the power of philosophical

imagination and the unexpected richness of everyday phenomena. I was also privileged to be able to attend lectures and seminars by Noam Chomsky, who surely stands to contemporary cognitive science as Einstein stands to physics.

Influential as these figures were, however, I can see now, looking back over my career, that the dynamic that was to drive almost all of my philosophical projects was set by my reading of the work of Willard van Orman Quine. In particular, I was influenced mightily by Quine's arguments for "naturalizing" epistemology. I did not, as some do, take Quine's view to be a rejection of normativity in the theory of knowledge. Nor did I take it to be a brief for the methodological primacy of the natural sciences. He did want to put philosophy under empirical constraint—that's what he meant by saying that there is no "first philosophy"—but that is not the same thing as saying that philosophy should be e*liminated* in favor of something else.

What affected me deeply, I think, was the idea that philosophizing should be *continuous* with other ways of coming to know the world. If philosophers are studying knowledge, and psychologists are studying knowledge, surely each group's findings are pertinent to those of the other. This idea of continuity and mutual constraint brought together, in my mind, not only philosophy and science, but philosophy and everyday beliefs. As I just mentioned, I had worried as a child about things like Limbo and Santa Claus—stuff I had been told was true, but that didn't make any sense to me. It irritated and saddened me that adults seemed to dismiss my questions, that they gave me the impression that there was something wrong with trying to *make* sense of all the things that I was being told I ought to believe. I didn't understand—if justice was treating people fairly, ensuring that they did not suffer for things they did not do, and if God was just, why did he deny heaven to innocent babies just because they were not baptized? It wasn't their fault they weren't baptized—so how could it be just that they be punished? Are there two kinds of justice—justice for us, and justice for God? Why are they both called justice, when they seem to be so different?

OK—I sort of slipped that in. I am not going to pretend that I've just produced the definitive version of the argument from evil. I do think, in fact, that the extent and distribution of suffering constitutes a decisive argument against the existence of God, but that's not the point I'm trying to make right now. There are responses to the challenge I just raised, and if my topic were actually the problem of evil, I would consider those responses and tell you my responses to *them*. But my topic is really

continuity—the mutual relevance of claims of truth in one domain to claims of truth in any other.

With your indulgence, then, what I'd like to do is describe three different ways in which continuity—the central idea of naturalized epistemology—has played out in the determination of my own philosophical interests. (And I do ask your indulgence—I have to confess that whenever someone lists the various topics that I've worked on, I worry that I just have a very short attention span. So I'm hoping, with this exercise, to convince myself as well as you that there is some kind of unity to my work.)

I'll start with some work I've done in language and mind. To explain it, I need to go into somewhat more detail about my reading of Quine.

Quine's arguments in "Two Dogmas of Empiricism"[2] are directed toward the positivist program of "rational reconstruction." We start with the question, what justifies scientific theories? A reasonable first answer is that scientific theories are justified by observations. Why are observations capable of justifying beliefs? Because they are composed out of sensory data that reflect extra-sensory reality. They are "impingements" (as Quine liked to call them) upon our sensory surfaces, and the character of these impingements is determined by the character of the external world. Insofar as our observations are determined by these impingements, and insofar as our theories are determined by our observations, we have an explanation of how the theories can be expected to correspond to the nature of the world.

But of course this answer needs a lot of elaboration. There are two big problems. To being with, there is the problem of underdetermination—the amount of sensory data available to any scientist, even throughout the course of her entire lifetime, is paltry compared to the richness of the theories she develops on its basis. And this problem for the scientist is just the problem, in second intention, that every human being faces when constructing his or her own "theory"—his or her set of commonsensical beliefs—about the external world. The second big problem is that the actual construction of scientific theories does not usually, if ever, follow the neat and direct causal path suggested by our opening considerations. Sometimes hypotheses are triggered by occurrences that have no evidential bearing on the matter at hand; sometimes evidence that does bear is neglected or discounted. The actual route from sensory evidence to scientific theory is meandering at best, full of false starts and dead ends, with progress often dependent on dumb luck.

It has to be conceded, then, that the actual story about the relation between sensory data and scientific theory is not going to be epistemologically ideal. Still, the thought is (was) that if the content of scientific hypotheses could be shown to be constructible *in principle* from sensory evidence that is available *in principle*—that would show how *in principle* science could correctly mirror the world.

Quine, famously, had telling criticisms of this program. He asserted that its viability depended upon principles that could not be defended, namely, the two "dogmas" of empiricism: the analytic/synthetic distinction and reductionism. I'm not going to rehearse *those* criticisms here, though. I want to focus on a different criticism—still Quinean—but one that's independent of the technical problems with positivism. This is the criticism that the program of rational reconstruction could not possibly fulfill the purpose that inspired it. If the question is how might theory be built out of sensory data, *in principle*, then the only issues are the technical ones. But if the question is what *justifies* belief in *actual* science, then the project of rational reconstruction is *irrelevant*. In a criminal trial, it matters little if the prosecutor *would be* justified *if* the defendant's fingerprints had been found on the murder weapon; the apposite question is, given that the fingerprints were *not actually* found on the murder weapon, what warrants the prosecutor in pursuing the case? Similarly, it matters little if we can characterize an ideal way for theory to be justified on the basis of sensory evidence, if what we want to explain is *how nonideal science* works so well. What we need to know is how, on the fragmentary and paltry evidential basis on which we must rely, do we ever manage to produce the science that we do? As Quine said, in one of my favorite passages (I do my best to quote it in almost every paper I write): "Why all this creative reconstruction; why all this make-believe?"[3]

Since the challenge of scientific theorizing was, as I said earlier, simply a special case of the problem of empirical knowledge, Quine's recommendation was to transform the study of empirical knowledge generally, from a project of a priori, idealistic construction into a project of realistic explanation. Start with our most obvious cognitive achievements—science, for example—and figure out how these were possible, given the constraints that define the human epistemic condition.

Not what seemed evident to me at the time (the late 1970s) was that the theories of language and mind that were then coming out of MIT—and really, out of Harvard, given the fundamentality of Hilary Putnam's work on functionalism—were doing *exactly* what Quine was urging us

all to do. Chomsky, for example, began his theorizing about language with the naturalistic question how a human child could acquire a complicated system of symbolic communication under the conditions in which children routinely performed this feat. The central point to be explained, for Chomsky, as for Quine in his critique of positivism, was underdetermination: the child's linguistic data were fragmentary relative to the rich content of the linguistic system the child managed to acquire, typically within the first four years of life. How, then, was language acquired? Linguistics naturalized!

Quine, of course, did not see things this way. He argued that the underdetermination of language by data was different in kind from the underdetermination of physical theory. The underdetermination of language was, he maintained, a second *level* of underdetermination, above and beyond the underdetermination of physical theory by sensory evidence. Fix the physical facts, he said—or at least, fix the theory that one used to capture those facts—and the linguistic "facts" floated free. Or so he argued. But his argument for this, ironically, contained the same fundamental flaw he had identified in the positivists' program of rational reconstruction—it was built on make-believe. Quine's argument for this doctrine, the "Indeterminacy of Translation," relied vitally on the following assumption: whatever facts about language there are, they must be rationally reconstructible on the basis of facts about the environmental contingencies of language use. This assumption, however, is just as unmotivated, and just as problematic as the positivists' assumption that scientific theory can be reconstructed on the basis of sensory data.

From the fact that the positivists were unable to reconstruct science under the strictures they had imposed on themselves, Quine did not conclude that there was no "fact of the matter" as to whether scientific theories were true or false. Physical theory might be underdetermined with respect to sensory *evidence*, but if we fixed all the physical *facts*, we would also fix physical theory. Not so with linguistic theory, Quine said; fix all the physical facts, and linguistic theory is still undetermined. The argument seems persuasive, I think, until you realize that Quine is operating with a proprietary conception of "all the facts." You'll recall that his radical translator was given a constrained database to work with. The would-be gavagalogist was allowed to record the environmental circumstances in which the native informant asserted, affirmed, or denied sentences in her language. But the translator was not allowed access to lots of other information that one might have thought could be pertinent—in particular, the translator was denied evidence about the native's mental states, not to mention their neurological realizations.

One important, and telling, exception to the ban on psychology—Quine granted the gavagologist knowledge of the behavioral signs of agreement and disagreement. I say that this exception is telling, because it highlights the in-principle dependency of any attributions of meaning on assumptions about the speaker's mental states. If the gavagalogist is not spotted this assumption—viz., that he can recognize when the native agrees or disagrees—the gavagologist's project would not get off the ground. Once we recognize the importance of this bit of psychology, we may also begin to see how much Quine is taking for granted about the *motives* of the native informant. Anthropologist Derek Freeman[4] charged that much of the polymorphous perversity reported by Margaret Mead in her study of the sex lives of Samoans was based on *bullshit*—and I use the term in its strict, Frankfurtian sense[5]—pitched to her by her native subjects for their own amusement. I suppose it did not occur to Mead that her informants would be anything but sincere and forthcoming. Of course, we are not told the ethnicity or race of Quine's "native informants," but his narrative does rather encourage the colonialist picture of indigenous peoples as simple folk, unfettered by their own projects and responsibilities, and warmly cooperative with the unintelligible project of this perfect stranger. I invite you all to see if it makes a difference to your assessment of Quine's argument to imagine that it's *English* that's being radically translated, and that you yourself are the native informant. How would *you* react to a guy with a notebook following you around all day, periodically pointing at animals and grunting "Rabbit? Rabbit?"

But I digress. The point is this: the evidence that Quine permits his radical translator to work with falls very far short of "all the physical facts." It does not, as I said, include any facts about psychology or neurology, facts that it's not implausible to assume have something to do with determining the properties of the language the native is speaking. The only justification Quine offers for his particular selection of facts is the thought that these *must* be the facts that determine the course of language acquisition. They are, he assumes, the only ones available to a human child when she is learning her native language, and the only ones available to the adult tutors who are shaping the child's emerging linguistic abilities. Quine writes:

> The sort of meaning that is basic to translation, and to the learning of one's own language, is necessarily empirical meaning and nothing more.

(Quine said "necessarily"!!!)

> A child learns his first words and sentences by hearing and using them in the presence of appropriate stimuli. These must be external stimuli, for they must act both on the child and on the speaker from whom he is learning. Language is socially inculcated and controlled; the inculcation and control turn strictly on the keying of sentences to shared stimulation.[6]

But if the criterion for inclusion in the data base is really what's available to a child learning language, then surely the author of "Naturalized Epistemology" would do a little empirical investigation to see what actually *is* available to the human child as she acquires linguistic competence. ("Why all this make-believe?") Surely Quine would do what Chomsky did—study the actual circumstances in which children pick up their native tongue. But no. He did not.

When we do look at the results of empirical study, we will notice something surprising. I've been complaining that Quine unjustifiably restricted the body of fact on which his radical translator could draw. But, it turns out, not only did Quine give his translator too *little* information (because Quine is right that behavioral dispositions do not fix the linguistic facts), he also gave his translator *too much*.

Quine gives his radical translator access to all *possible* behavioral data. If the translator is deliberating between two alternatives for a translation of a native construction, and if the matter *could* be settled by empirical test, Quine generously provides the results of such test to the translator's body of evidence. So consider, for example, (from Lila Gleitman) the pair of verbs "chase" and "flee." For a very large range of circumstances, it's the case that if one of these verbs is appropriate, so is the other: whenever the hounds are chasing the fox, the fox is fleeing the hounds. This sort of situation poses an in-principle problem for the real-life language learner: the observed contingencies of use for these two verbs are apt to be identical. It doesn't help to point out (as Steven Pinker[7] does) that there *are* contexts in which one but not the other is suitable (e.g, I can flee the city, but the city doesn't chase me when I do so). The *counterfacts* cannot help the child; if the child never actually encounters such circumstances, the evidence they *would* provide cannot be part of the child's data base.[8] If Quine were consistent in his rationale for excluding information about neurobiology and cognitive psychology, he ought also to have excluded evidence that the translator might have gathered but did not actually obtain.

There are other bits of generosity on Quine's part that have no grounding in the actual circumstances of acquisition. (1) The human child does not have the pedagogical equivalent of the translator's native informant; children receive virtually no explicit feedback on their linguistic efforts, and when they do, they ignore it. (Aside: it is a matter of controversy whether the child's environment does or doesn't contain enough information for the child to learn the grammar of her native language purely on the basis of environmental stimuli. But the current point is that Quine does not even ask the *question* what data are actually available to the child.) (2) A large percentage of the child's early linguistic data consist of questions and imperatives; it is of the nature of these sorts of sentences that they are uttered in the *absence* of their subject matter. (Generally, it's *open* doors, windows, and mouths that we ask people to shut.) (3) The verbs that occur most commonly in adult speech, and which are among the first for children to acquire, are also the most general; it is hard to think of a circumstance in which it would *not* be appropriate to use some form of "do," "make," "have," or "be."[9]

These considerations are, of course, all part of Chomsky's "poverty-of-the-stimulus" argument for the existence of innate linguistic knowledge. But that's not my point at the moment. I'm making an *ad hominem* against Quine—I'm charging him with being insufficiently naturalistic. He draws ontological conclusions about the domain of language from a thought-experiment that proceeds from theoretical suppositions about how language acquisition *ought* to proceed. He did not conclude from the failure of Carnap's project in the *Aufbau* that empirical science is indeterminate, that there is no fact of the matter which of two (apparently) conflicting scientific theories is correct; he concluded only that if we want to understand how we do manage to make sense of the empirical world, we must do so by working within the constraints that define the human epistemic condition. But focus on the actual characteristics of the human epistemic condition is the whole point of naturalizing epistemology!

Chomsky used the poverty of the stimulus to argue for the existence of innate information that the human child brings to the task of language acquisition. Quine, himself, certainly appreciated the need for substantive pre-theoretic, or, as we might as well say, *a priori* elements in explaining the development of human knowledge in general—he just didn't see it in the case of language. Quine pointed out, for example, that induction cannot be performed unless there is in the learner an "innate similarity space" that determines the parameters of generalization. Moreover, he (with Joseph Ullian, in *The Web of Belief*)[10] acknowledged the vital role played by "extra-empirical" assumptions in scientific

theorizing: we prefer the simpler explanation to the more complicated, the more conservative alterations to the more radical ones, the more general claims to the more *ad hoc*. Of course, as Quine also noted, these preferences frequently come into conflict—as when a radical alteration promises ontological economy and increased generality—and it is unlikely that we will ever have a comprehensive normative account of how such trade-offs should be decided.

These *a priori* elements, we should acknowledge, are a type of *bias*. They incline us in certain directions, and they do so in advance of evidence. They must do so in advance of evidence because we need them in order to *get* evidence. Without an innate similiarity space, we would have to consider every objective parameter of variation equally important. Our empirical conjectures would be tests of our principles of projection as much as they would be tests of our generalizations. So it is our innate biases to treat some properties as more important than others that makes empirical knowledge possible. (And one note: there is a kind of ultra-empiricism that holds that such biases, though handy, are really not necessary for the achievement of knowledge. Defenders of this view appeal to the existence of programs that permit huge and crazy fast computers to pull statistical regularities from enormous data sets. This is interesting if what one is interested in is the power of mechanical computation. But it tells us nothing about the strategies that human beings employ when they successfully infer statistical patterns in naturalistic conditions. If they have neither the power, the time, or the data that are available to the statistical inference engines, then *they* must be doing it a different way.)

The idea that bias has a *constructive* role to play in the generation of human knowledge really resonated with me and informed not only my work in the philosophy of mind, but also my work in feminism. Feminist epistemologists have been critical of abstract and individualistic models of human knowledge and have called for accounts that "situate" knowledge in the natural and social environment. Quine's naturalized epistemology seemed to me to answer this call. The naturalized approach to the study of individual knowers, the approach that seeks to explain how an individual comes to know what she comes to know in the circumstances in which she manages to do so, reveals that she is aided by a host of native inclinations—biases—and, in some cases, by highly specialized learning mechanisms. But insofar as our epistemic achievements rely on these inclinations and mechanisms, they rely as well on a hospitable environment. Our humanly native language-acquisition device might not enable us to acquire, say, Klingon under conditions of casual exposure;[11] the biases that that device embodies

might even impede our efforts to learn Klingon in any other way as well. Maybe there is some argument from evolutionary considerations to the conclusion that we cannot be *too* poorly attuned to any environment that is nomologically possible—this would be a good thing to talk about over drinks—but we do know that our intuitive physics and kinematics lead us astray when things get much bigger or much smaller or much faster than what we're used to in everyday life. I think the most likely thing is that whenever creatures evolve who can extract information from their environments, they can only do so because they utilize some "merely" satisficing epistemic mechanism, one that yields a lot of information—or a lot of *useful* information—while trading off some degree of accuracy. It is in this sense that real, actual knowledge is "situated." I conjecture that the perfectly "*safe*" epistemic strategy does not really exist—at least not for creatures who rely on their bodies to find things out.

The idea that we human beings rely a lot on fallible biases and heuristics is quite familiar today due partly to popular books written by social scientists like Daniel Kahneman, Gerd Gigerenzer, and Dan Ariely.[12] This is a glass-half-empty-half-full sort of situation. A fallible heuristic like Gigerenzer's "familiarity heuristic" (when in doubt, pick the one you've heard of) is great when it's applied in the right domain (one where familiarity actually co-varies with the property you're interested in) but dangerous if applied in the wrong one. The same is true of more general cognitive biases that human beings tend to display, like confirmation bias, and one-trial "learning." If one is a member of a prey species, the cost of an epistemic bias that yields lots of false positives for the presence of predators may be more than offset by the benefit of quick judgment when the predator is actually approaching. Sarah-Jane Leslie, in her work on generics in natural language, points out that we human beings readily accept a number of generic claims that are in fact true of only a tiny fraction of the group they cover—"ticks carry Lyme disease"—when these generics involve what she calls "dangerous properties." She theorizes that this natural cognitive bias, allowed to operate freely in certain social conditions—conditions that encourage ingroup/outgroup sentiments—may be the source of racist prejudice.[13]

So, from a naturalized epistemological point of view, individual cognitive biases are neither inherently bad nor inherently good—it all depends on the environment in which the biases are deployed. But I think that we can see that there is the same relativity even when we talking about epistemic strategies that might seem more robustly reliable than the processes that, for example, generate "dangerous-property generics." What I have in mind are cognitive strategies for extracting genuine regularities, when they are deployed in social environments that have

been shaped by injustice. There are two kinds of cases I want to describe briefly.

The first kind involves our accurately tracking the causal consequences of systematic disadvantage. It is not racist to believe that most professional basketball players are black. It is simply a fact: nearly 75 percent of NBA players are African-American.[14] Nor is it racist to believe that there are not many black classical musicians: less than 3 percent of positions in U.S. symphony orchestras are held by African-Americans.[15] And it is not sexist to believe that most philosophers are men—the numbers are clear: 70 percent or more of academic philosophy jobs are held by men.[16] It is perfectly proper, epistemically, to believe such things on the basis of observation. The problem with such generalizations is not their content, but the significance we assign to them, or the inferences we make from them, especially when these extra cognitive steps occur without our conscious notice.

Our capacities for generalization often connect us to natural kinds; the reliable co-occurrences of observable properties often signal the existence of a "real essence," a property or set of properties that are explanatorily or causally fundamental. But properties that have no inherent or deep natural connection with each other can be brought into concordance by extrinsic, contingent forces. In human life, these forces are often social, and they are often reflective of power relations.

We go wrong when we uncritically attribute properties like athletic ability, musical ability, philosophical ability to racial or gender *essences*. (A word on group essences: one could argue—I would argue—that the belief that there are group essences involves some kind of confusion—whether it's a mistake in reasoning, or a false element in folk metaphysics, I don't know. But some race theorists are willing to allow that belief in group essences is innocent on its own. Kwame Anthony Appiah, for example, distinguishes the view he calls "rac*ialism*"—the belief that there are group essences—from "rac*ism*"—the belief that these group essences lead to differences in moral character or moral worth.[17]) In any case, to draw the conclusion that facts like the ones I've cited support claims about inherent ability, or any other inherent property, neglects the degree to which these regularities reflect systematic inequities in opportunity. Maybe white men can't jump, but they can play tennis. Only 4 out of the top 100 male players ranked by the Association of Tennis Professionals[18]—and this is an international organization—are non-white. Is it just possible that the difference in accessibility between basketball courts and tennis courts explains most of the demographic variance?

The second sort of case where innocent and epistemically legitimate generalizations can lead us astray involves social feedback loops. Feminist theorists like Marilyn Frye, Catherine MacKinnon, and Sally Haslanger, as well as philosophers of social science like Ian Hacking,[19] have described the dynamics of these phenomena in which a message about the character of a dominated class becomes, in a way, self-verifying. Consider the claim that women have no facial hair. Is it true or false? When I ask my undergraduate students, most of them say that it's true—almost all the men say it's true, and many of the women do too. Now, of course, many women do have facial hair. Many women have quite a lot of facial hair. But this will not be evident from observation. There is a powerful gender norm for women in contemporary American society that says that women *ought not* to have facial hair. The gender norm against facial hair on a woman is so strong that woman who have abundant or dark facial hair go to great lengths to remove it, often at great expense, and almost always despite a lot of discomfort. The result: look around—you will not see many women, if any, with noticeable mustaches, sideburns, or beards. So the generalization seems to be confirmed: women have no facial hair. And then, the more strongly it is confirmed, the more it seems—to men and women alike—that a woman who *displays* facial hair is some kind of statistical outlier. Or, to put it less kindly, is some kind of freak, not a "real" woman, or not a "good" woman. The expectation that women will have hairless faces (and armpits and legs, now that we're getting down to it) becomes normative. Women who "refuse" to "deal with" their facial hair have lost jobs over this; they have been scolded, ridiculed, and internet-shamed.[20]

There are similar stories to tell about a great many of the properties that are taken to be "natural" consequences of being biologically male or biologically female There is an interplay among normative expectations, the structure of reward contingencies, the behavior that results from these structures, and the observable regularities, that finally feed back to vindicate the normative expectations. My own view of gender—one that used to be considered radical but that is now regarded by some as reactionary—is that gender is the output of a system of social constraints that creates and sustains artificial concordances among properties that would otherwise vary independently—a system that we would be better off without.[21]

Now, I should say that, although I endorse the idea that much of human knowledge is situated—again, where I mean that the success of our epistemic strategies depends on their being exercised in the right sort of environment—I also disagree with many feminist epistemologists about the utility of abstraction and idealization in the

study of knowledge. Insofar as human knowledge involves an interaction between a cognitive or perceptual mechanism and an environment, it is appropriate, and indeed, necessary to study the operation of the mechanism in abstraction from the environment in which it is deployed. I see no problem in studying the human mind as if it were a computer—provided we tether our investigations to the naturalistic conditions in which human mentality arises.

I've been pushing an interpretation of "naturalistic epistemology" according to which the main idea is that all forms of human inquiry are continuous. But this continuity is a matter of mutual constraint, not a matter of methodological identity. There are lots of ways in which things we come to believe through one epistemic mode conflict with things we believe through some different epistemic mode. Common sense tells us that there are material objects, some of which are solid and colored; physics tells us that these "objects" are mostly empty space, and (maybe) that color is a merely subjective unity. Our linguistic practice allows us to form names that seem to refer to properties; our logic tells us that if we take this appearance at face value, we will run into contradictions. We have strong "intuitions" that certain things are morally wrong, but these intuitions seem to float free of empirical test.

I myself have been grabbed by apparent conflicts between our everyday, ordinary view of ourselves as rational agents (or at least as agents capable of acting on the basis of reason) and the picture that seems to emerge from the biological and psychological sciences, that we are a bundle of causal mechanisms. This is, of course, the conflict between—as Wilfred Sellars so beautifully put it—the "manifest" image and the "scientific" image. Not that Sellars thought that there *was* a conflict—he believed that the two images were not reconcilable, but that that was OK. He was, on my accounting, a methodological dualist. According to him, there is no mutual constraint, no need to "reconcile" the two images. We are, according to him, simply engaged in different practices when we consider matters as falling within the "space of reasons," versus considering them as falling within the "space of causes." This is absolutely not OK in my opinion. If our practice of assigning and giving reasons serves us well in everyday life, then that is a fact about the world that must be explained. Human beings, whatever else we are, are material beings, subject to the same physical laws as govern everything else in the universe. Somehow, the reasons we have for doing things *literally* move us—they issue in bodily motions, and that must be explained. Moreover, our practice of regarding human beings (and some non-human animals) as rational agents is productive. We can readily make, in an instant, predictions on the basis of assumptions of

(very minimal) rationality that it would take a LaPlacian demon an eon to compute. My favorite example is one from Georges Rey: construct a pathetically easy True-False quiz, and administer it to a bunch of undergraduates who you promise to pay $100 if they get a perfect score. Have them record their answers in No. 2 pencils on an optical scan sheet. On the basis of this set-up, you will be able to predict the distribution of graphite on 100 sheets of paper with extreme precision.[22] Notice that this little experiment spans the "gap" between the space of reasons and the space of causes. "Knowing the answers to the questions," "understanding how to use an op-scan sheet"—these are intentional characterizations of things, a placing of matters within the space of reasons. But graphite marks on paper—that's space-of-causes stuff. Yet we are making predictions about the latter on the basis of the former. And so it is with much of our everyday figuring-out.

The most promising idea for explaining this, this coordination between reasons and causes, seemed to me, early on, to come from Donald Davidson. Reasons could move us, he argued, because reasons were a kind of cause. A reason was a type of mental state, a belief-desire combo called a pro-attitude, that was token-identical with some brain state. (I prefer to say "realized as" rather than "is token-identical with," but let that pass.) Qua mental state, the pro-attitude *rationalized* the action—the action was revealed to be reasonable in light of the agent's belief and desire. Qua brain state, the pro-attitude *caused* the action—our friendly neighborhood neurologist could presumably—someday—trace the physical path between the brain event that was the onset of the pro-attitude and the eventual set of muscle contractions that constituted the action.

What a beautiful picture! But then Davidson goes all squooshy, and announces that there can be no regular, nomically reliable connections between events typed in mental terms and events typed in physical terms. "Nomological slack between the mental and the physical," Davidson wrote (breaking my heart, not that he cared—or knew) "is necessary as long as we to conceive ourselves as rational beings."[23]

Since I was a graduate student in Cambridge in the late 1970s, one of the towering figures I was fortunate enough to encounter was, as I mentioned earlier, Jerry Fodor. Fodor was—and is—all about reconciling our ordinary "folk psychology" with scientific accounts of the mind, and had a perfectly lovely theory that could do everything that needed doing on this score. According to Fodor, propositional attitude states were (a) functional states involving (b) relations to sentence-like mental representations, items in a "language of thought" as it were. Positing

such representations explains a lot: why there were Frege cases, why we don't immediately recognize the truth of every mathematical theorem as soon as we learned that 2 + 2 = 4, not to mention how Oedipus wanted to marry Jocasta without wanting to marry his mother.[24]

But returning to the issue at hand—it seemed to me to be a no-brainer. If we took a realistic view of mental representation, we could explain how mental states can have both rational/semantic properties *and* causal properties. Moreover, we could explain how a mechanism that looks merely causal from the vantage point of physics can nonetheless be "disciplined" to respect the rational relations among thought contents. Because (1) syntax can encode rationally important structure, and (2) syntactic structures can be realized in physical structures, and finally (3) there can be physical causal mechanisms that are sensitive to just those physical differences that correspond to syntactic differences, we can—wait for it—mechanize rationality! How could anyone not love this picture?

Well, Davidson didn't. And many other philosophers didn't. And many still don't. But if I have any vestiges of my Catholic upbringing, they reside in my evangelical zeal to bring the good news of the Language of Thought to all humankind. I find that too many of my philosophical brethren and sisteren struggle unnecessarily with problems that realism about mental representation just—solves. I mentioned Frege problems. Many philosophers—I won't name names; you know who you are—think that Frege problems show the need for complicating our notion of semantic *content*. Reference, they say, is too *coarse-grained*. We need finer-grained contents, or we need ways of constructing contents that mirror the structure of the sentences that express them. Or we need epistemic contents in addition to metaphysical contents. But all this is to misdiagnose the problem: we don't need more contents, or more kinds of content; we need a whole other dimension, something *orthogonal* to content.

Here's the troublesome truth: all of the puzzles about intensionality derive from the fact that there is no guarantee whatsoever that there will be a one-one correspondence between things and representations of things. It doesn't matter if your "things" reside in metaphysically possible worlds, epistemically possible worlds, all possible worlds, or the planet Tatooine.[25] As soon as there's a *thing*, BINGO—there's the possibility of two distinct representations of the thing. And because representations are objects in their own right, I can have a relationship with one of them that I don't have with the other. So I can pick up one representation, but not the other. Or I can pick up both representations and think that they

refer to different things. None of this is an accident, nor does it indicate representational malfunction. It is a *feature* of representation, not a bug, that representations are ontologically, causally, epistemically, and every other way distinct from the things they represent. In fact, they are *so* completely independent of the things they represent that there don't even need *to be* things in order for there to be representations that represent them. Yes! They solve the problem of empty reference too! Ask me later about the exciting details!

But getting back to the issue of reasons and causes: Why was Davidson so adamantly dualistic about the mental and the physical? He surely didn't think that there was an unbreachable gap between, say, the biological and the chemical. Why did he think that there not only was but *had to be* a gap preserved between the mental and the physical? I think that what he was afraid of was this: Suppose that we were to discover a robust correlation between a particular belief-type and a particular neurological state-type, relativized, if you like, to a particular person. So suppose that, so far, whenever I believe that there's a spider in the vicinity, there has been a distinctive pattern of firing in a specific region in my brain. It used to be that in order to tell whether I believed there to be a spider in the room, you would have to observe my behavior—you'd ask me if I thought there was a spider, for example, or, knowing that I am severely arachnophobic, you'd see if I was sidling toward the door or (if escape was impossible) screaming. But now it looks like you have a new means at your disposal: you can check the brainoscope and see if the characteristic pattern was occurring. All good, right? Wrong! What if the brainoscope says that I am in the state of believing that there's a spider there, but I am not screaming, not sidling, and when you ask me if I believe there's a spider, I calmly tell you no. That is, the brainoscopic data are in conflict with the behavioral data—the ordinary marks of belief on which we normally rely and which enable us to make sense of each other's actions (and inactions). If we allow the neurological facts to override the behavioral facts, we violate the precepts that define the very notion of belief!

This would be disconcerting, I concede. Just as disconcerting as it would be if a surgeon opened the head of an intelligent human being and found that her skull was completely empty. If such a thing were to happen, I think it would be epistemically appropriate to reconsider our commitment to the physicalistic idea that human intelligence depends on our having brains. But it's not going to happen. Because human intelligence *does* depend on our having brains. So we don't have to worry that this nightmare scenario will come to pass. (That's why they call it "science *fiction*.") The conceivability of such a thing happening is

not, in itself, an argument against materialism. (I think even Joe Levine[26] agrees with me about this.)

My response to Davidson, then, is that he needn't worry. I can *modus ponens* his *modus tollens*: (1) Our folk-psychological criteria for the possession of mental states are reliable. (2) We have excellent reason to think that mental states are realized by neurological states (remember all the stuff about explaining the success of mental state attributions?), so therefore (3) our folk psychological criteria will *not* come into conflict with our neurological data. As Thomas Burnet wrote to Isaac Newton in another context, "The truth cannot be an enemy to truth." (Funny story: Burnet was trying to reassure Newton that the non-cyclical geology that he, Burnet, was pushing would not in the end conflict with Scripture.[27]) I'm prepared to put money down on this, and so should anyone who really believes that folk psychology works. For extra confidence, let me throw in this additional consideration: the development of our reductive account of mental states is going to be *constrained* by folk psychology. We won't believe a theory that makes the folk psychology come out wrong any more than we would believe a reductive physical theory that tells us that boiling water is colder than ice. Now, there should have been *ceteris paribus* clauses all over the place back there—but a little slippage is OK—even Davidson agrees that we *sometimes* have to treat someone as irrational—if the constitutive principles of belief-attribution can be thwarted a little, now and then, without our having "changed the subject," then a little slippage, now and then, between neurological and behavioral evidence shouldn't be a problem either.[28]

CONCLUSION

I started with the admission that for some time I did philosophy without really knowing what it was I was doing. So what have I learned about philosophy that I didn't know before? Well, I've argued that philosophy is continuous with science, and that philosophy is continuous with common sense. But what *is* philosophy? My meta-philosophy, such as it is, is this: Philosophy is distinctively synoptic. Philosophy is the discipline that surveys all the things that we think are true and tries to figure out how they can be true together. It might not work—we might have to give up some cherished beliefs, we might end up in places we have wanted like anything to avoid—but then again, we might succeed.

A well-known philosopher I was talking to once told me that he admired my work for its "optimism." At the time, I took it as damnation with faint praise, and that might have been the intention. But now I think it's maybe

just a fair description. My mother often used to say to me when I was a child, "Louise, you just can't expect things to always make sense." Well, that's exactly what I expect, and that's why I do philosophy.

NOTES

1. My title is an allusion to a song of the same name, written by my daughter, Rachel Lark, as a present me to celebrate my presidency of the Eastern Division of the APA. https://www.youtube.com/watch?v=Te4JILBXLiI
2. W. V. O. Quine, "Two Dogmas of Empiricism" *Philosophical Review* 60, no. 1 (1951): 20–43.
3. W. V. O. Quine, "Epistemology Naturalized," in *Ontological Relativity and Other Essays* (New York: Columbia University Press: 1969).
4. Derek Freeman, *The Fateful Hoaxing of Margaret Mead: A Historical Analysis of Her Samoan Research* (New York: Basic Books, 1998).
5. Harry G. Frankfurt, *Bullshit* (Princeton, NJ: Princeton University Press, 2005).
6. W. V. O. Quine, *Word and Object* (Cambridge, MA: The MIT Press, 1960), 81.
7. Cited by Gleitman, in Gleitman 1990, below.
8. Lila Gleitman, "The Structural Sources of Verb Meanings," *Language Acquisition* 1, no. 1 (1990). Reprinted in Paul Bloom (ed.), *Language Acquisition: The Core Readings* (Cambridge, MA: The MIT Press, 1993).
9. Gleitman, op. cit.
10. Quine and Ullian, *The Web of Belief* (New York: Random House Books, 1970).
11. Klingon is the language spoken by members of the Klingon race, "a fictional extraterrestrial humanoid warrior species in the science fiction franchise Star Trek." Wikipedia, https://en.wikipedia.org/wiki/Klingon; accessed July 27, 2016.
12. Daniel Kahneman, *Thinking, Fast and Slow* (New York: Ferrar, Straus, and Giroux, 2011); Gerd Gigerenzer, *Gut Feelings: The Intelligence of the Unconscious* (New York: Viking Press, 2007); Dan Ariely, *Predictably Irrational: the Hidden Forces that Shape Our Decisions* (New York: Deckle Edge/Harper Collins, 2009).
13. Sarah-Jane Leslie, "The Original Sin of Cognition," *Journal of Philosophy* (forthcoming).
14. Richard Lapchick and Angelica Guiao, "The 2015 Racial and Gender Report Card: National Basketball Association," July 1, 2015. tidesport.org.
15. League of American Orchestras: "Welcome…" by Aaron Dworkin. Accessed July 28, 2016. http://www.americanorchestras.org/todays_news/604.html
16. See the report by Eric Schweitzgebel and Carolyn Dicey Jennings, "Women in Philosophy: Quantitative Analyses of Specialization, Prevalence, Visibility, and Generational Change," available through the American Philosophical Association's Committee on the Status of Women website: http://www.apaonlinecsw.org/data-on-women-in-philosophy
17. Kwame Anthony Appiah, "Racisms," in *Anatomy of Racism*, ed. David Goldberg, 3–17 (University of Minnesota Press, 1990).
18. Prasant Tangirala, "Race and Color in Tennis: Whither Diversity?" *Bleacher Report*, http://bleacherreport.com/articles/181775-race-and-color-in-tennis-whither-diversity
19. For example, see Marilyn Frye, *The Politics of Reality: Essays in Feminist Theory* (Berkeley, CA: The Crossing Press, 1983); Catherine MacKinnon, *Feminism Unmodified* (Cambridge, MA: Harvard University Press 1987); Sally Haslanger, "On

Being Objective and Being Objectified" in *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, ed. Louise Antony and Charlotte Witt, 8–25 (Boulder, CO: Westview Press, 1993); Sally Haslanger, "Ontology and Social Construction," Philosophical Topics 23, no. 2 (Fall 1995): 95–125; Ian Hacking, "The Looping Effects of Human Kinds," in *Arguing About Human Nature*, ed. Stephen Downes and Edouard Machery (New York and Florence, KY: Routledge, 2013).

20. In one case, a young Sikh woman was photographed and subjected to censure because she did not shave her beard, mustache, and sideburns. The incident and its (inspiring) sequelae are explained in this posting on Jezebel: http://jezebel.com/5946643/reddit-users-attempt-to-shame-sikh-woman-get-righteously-schooled

21. See my "Back to Androgyny: What Bathrooms Can Tell Us about Equality," *Journal of Contemporary Legal Issues* 9 (Spring 1998): 1–20.

22. Georges Rey, "A Not 'Merely Empirical' Argument for the Language of Thought," *Philosophical Perspectives* 9 (1995): 201–22.

23. Donald Davidson, "Mental Events," in *Experience and Theory*, ed. L. Foster and J. W. Swanson, 79–101 (Humanities Press, 1970).

24. Jerry Fodor, "Propositional Attitudes," *The Monist* 61 (October 1978): 501–23.

25. "Tatooine is a fictional desert planet that serves as the setting for many key scenes throughout the Star Wars saga." Wikipedia, https://en.wikipedia.org/wiki/Tatooine, accessed July 28, 2016.7. Joseph Levine, "Conceivability and the Metaphysics of Mind," *Noûs* 32, no. 4 (1998): 449–80.

26. Joseph Levine, "Conceivability and the Metaphysics of Mind," *Noûs* 32, no. 4 (1998): 449–80.

27. Quoted in Stephen Jay Gould, *Time's Arrow, Time's Cycle* (Cambridge, MA: Harvard University Press, 1987).

28. I lay out these points in more detail in my Louise M. Antony, "Anomalous Monism and the Problem of Explanatory Force," *Philosophical Review* 98 (April 1989): 153–87.

Photograph courtesy of Paul Shoul, Northampton, MA.